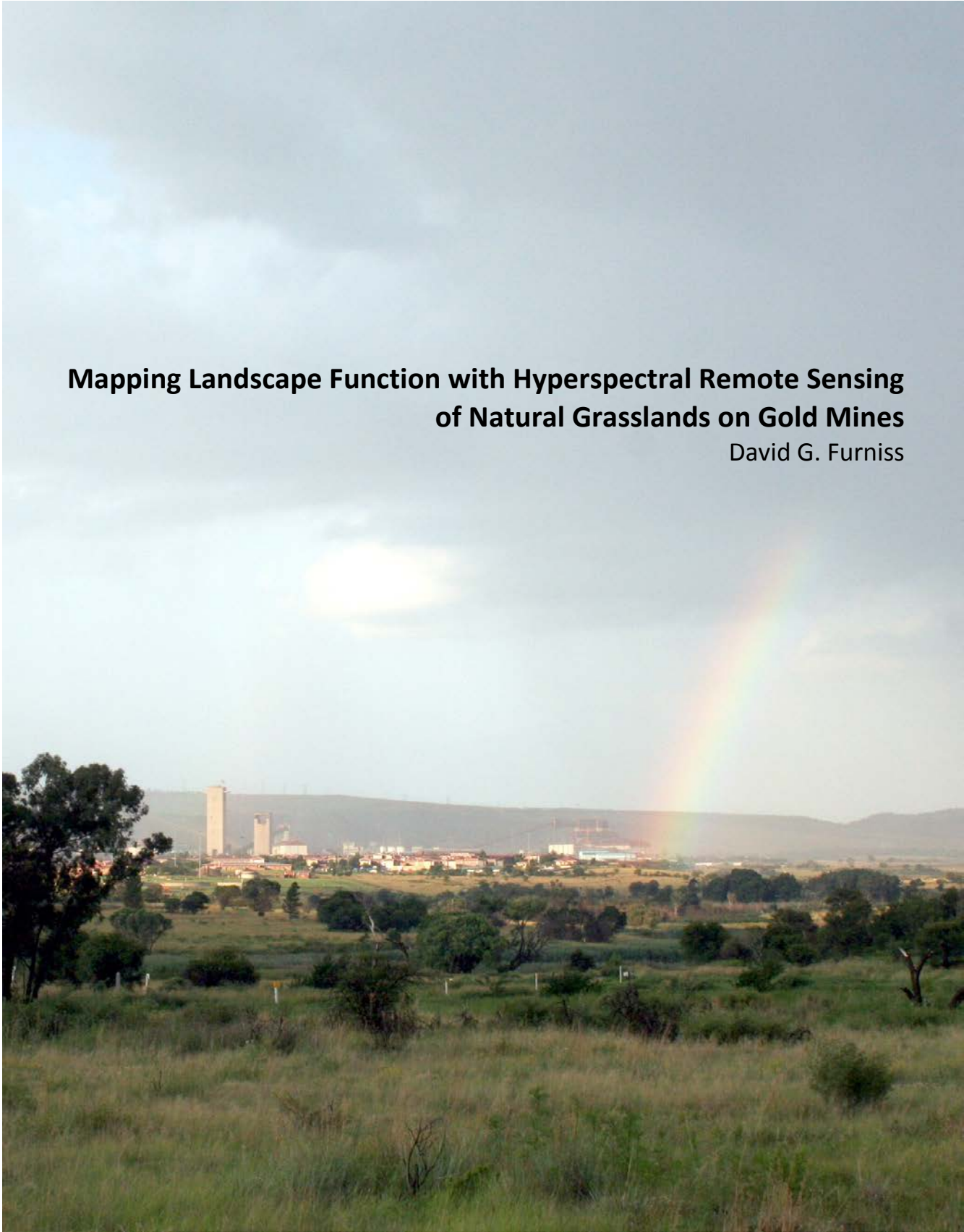


A landscape photograph showing a grassy field in the foreground, a town with a tall tower in the middle ground, and a rainbow in the sky. The text is overlaid on the upper part of the image.

Mapping Landscape Function with Hyperspectral Remote Sensing of Natural Grasslands on Gold Mines

David G. Furniss

Mapping Landscape Function with Hyperspectral Remote Sensing of Natural Grasslands on Gold Mines

David G. Furniss

Student No.: 0711251J

APES9002

Thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy.

School of Animal, Plant and Environmental Science,
University of the Witwatersrand, Johannesburg, South Africa

October 2016

Supervisors:

Prof Michael Sears: School Of Computer Science and Applied Mathematics

Prof Deanne Drake: School of Animal, Plant and Environmental Sciences

University of the Witwatersrand

DECLARATION

I declare that this thesis is my own, unaided work. It is being submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.



Signed:

David Gordon Furniss

On 7 day of October, 2016

This study was conceived and designed by D.G. Furniss with guidance from M. Sears and D. Drake and grew from a pilot study designed by I.M. Weiersbye with input from E.T.F. Witkowski, D.J. Tongway and D.G. Furniss. D.G. Furniss collected the LFA and ground-based hyperspectral data, took all photographs, did all statistical analyses, modelling and interpretation, and wrote this thesis.

ABSTRACT

Mining has negative impacts on the environment in many different ways. One method developed to quantify some of these impacts is Landscape Function Analysis (LFA) and this has been accepted by some mining companies and regulators. In brief, LFA aims at quantifying the organization of vegetative and landscape components in a landscape into patches along a transect and quantifying, in a relative manner, three basic processes important to landscape functioning, namely: soil stability or susceptibility to erosion, infiltration or runoff, and nutrient cycling or organic matter decomposition. However, LFA is limited in large heterogeneous environments, such as those around mining operations, due to its localized nature, and the man hours required to collect a representative set of measurements for such large and complex environments. Remote sensing using satellite-acquired data can overcome these limitations by sampling the entire environment in a rapid and objective manner. What is required is a method of connecting these satellite-based measurements to LFA measurements and then being able to extrapolate these measurements across the entire mine surface.

The aim of this research was to develop a method to use satellite-based hyperspectral imagery to predict landscape function analysis (LFA) using partial least squares regression (PLSR). This was broken down into three objectives: (1) Collection of the LFA data in the field and validation of the LFA indices against other environmental variables collected at the same time, (2) validation of PLSR models predicting LFA indices and various environmental variables from ground-based spectra, and (3) production of risk maps based on predicting LFA indices and above-ground biomass using PLSR models and Hyperion satellite-based hyperspectral imagery. Although the study was based in grasslands at two mining regions, West Wits and Vaal River, a suitable Hyperion image was only available for Vaal River.

A minimum of 374 points were sampled for LFA indices, ground-based spectra, above-ground biomass and soil cores along 2880 m of LFA transect from both mine sites. Soil cores were weighed fresh before sieving with a 2 mm sieve to separate root and stone fractions. The sieved soil fraction was tested for pH, EC, SOM, and for the West Wits samples, organic nitrogen and total extractable inorganic nitrogen. There was one modification to the LFA method where grass patches were collapsed into homogenous units as it was deemed not feasible to sample 180 m transects at grass tuft scales of 10 – 30 cm, but other patch definitions followed the LFA manual (Tongway and Hindley, 2004). Evidence suggested that some of the different patch types, in particular the bare/biological soil crust – bare grass – sparse grass patch types, represented successional stages in a continuum although this was not conclusive. There also was evidence that the presence or absence

of cattle play a role in some processes active in these grasslands and erosion is mainly through deflation, rain splash and sheet wash. Generally the environmental variables supported the LFA indices although the nutrient cycling index was representative of above-ground nutrient cycling but not below-ground nutrient cycling.

Models derived with PLSR to predict the LFA indices from ground-based spectral measurements were strong at both mine sites (West Wits: LFA stability $r^2 = 0.63$, $P < 0.0001$; LFA infiltration $r^2 = 0.75$, $P < 0.0001$; LFA nutrient cycling $r^2 = 0.73$, $P < 0.0001$; Vaal River: LFA stability $r^2 = 0.39$, $P < 0.0001$, LFA infiltration $r^2 = 0.72$, $P < 0.0001$, LFA nutrient cycling $r^2 = 0.54$, $P < 0.0001$), as were PLSR models predicting above-ground biomass (West Wits above-ground biomass $r^2 = 0.55$, $P = 0.0003$; Vaal River above-ground biomass $r^2 = 0.79$, $P < 0.0001$) and soil moisture (West Wits soil moisture $r^2 = 0.45$, $P = 0.0017$; Vaal River soil moisture $r^2 = 0.68$, $P < 0.0001$). However, for soil organic matter ($r^2 = 0.50$, $P < 0.0001$) and EC ($r^2 = 0.63$, $P < 0.0001$), Vaal River had strong prediction models while West Wits had weak models for these variables ($r^2 = 0.31$, $P = 0.019$ and $r^2 = 0.10$ and $P < 0.18$, respectively). For EC, the wide range of soil values at Vaal River in association with gypsum crusts, and low values throughout West Wits explained these model results but for soil organic matter, no clear explanation for these site differences was identified. Patch-based models could accurately discriminate between spectrally well-defined patch types such as *S. plumosum* patches but were less successful with patch types that were spectrally similar such as the bare/biological soil crust – bare grass – sparse grass patch continuum. Clustering similar patch types together before PLSR modelling did improve these patch-based spectral models.

To test the method proposed to predict LFA indices from satellite-based hyperspectral imagery, a Hyperion image matching 6 transects at Vaal River was acquired by NASA's EO-1 satellite and downloaded from the USGS Glovis website. LFA transects were partitioned to match and extract pixel spectra from the Hyperion data cube. Thirty-one spectra were separated into calibration (20) and validation (11) data. PLSR models were derived from the calibration data, tested with validation data to select the optimum model, and then applied to the entire Hyperion data cube to produce prediction maps for five LFA indices and above-ground biomass. The patch area index (PAI) produced particularly strong models ($r^2 = 0.79$, $P = 0.0003$, $n = 11$) with validation data, whereas the landscape organization index (LOI) produced weak models. It is argued that this difference between these two essentially similar indices is related to the fact that the PAI is a 2-dimensional index and the LOI is a 1-dimensional index. This difference in these two indices allowed the PAI to compensate for some burned pixels on the transects by "seeing" the density pattern of grass tufts and patches whereas the linear nature of the LOI was more susceptible to the changing dimensions of patch structure due

to the effects of fire. Although validation models for the three LFA indices of soil stability, infiltration and nutrient cycling were strong ($r^2 = 0.72$, $P = 0.004$; $r^2 = 0.66$, $P = 0.008$; $r^2 = 0.70$, $P = 0.005$, $n = 9$ respectively), prediction maps were confounded by the presence of fire on some transects. The poor quality of the Hyperion imagery also meant great care had to be taken in the selection of models to avoid poor quality prediction maps. The 31 bands from the VNIR (478 – 885 nm) portion of the Hyperion spectra were generally the best for PLSR modelling and prediction maps, presumably because of better signal-to-noise ratios due to higher energy in the shorter wavelengths.

With two satellite-based hyperspectral sensors already operational, namely the US Hyperion and the Chinese HJ-1A HSI, and a number expected to be launched by various space agencies in the next few years, this research presents a method to use the strengths of LFA and hyperspectral imagery to model and predict LFA index values and thereby produce risk maps of large, heterogeneous landscapes such as mining environments. As this research documents a method of partitioning the landscape rather than the pixel spectra into pure endmembers, it makes a valuable contribution to the fields of landscape ecology and hyperspectral remote sensing.

ACKNOWLEDGEMENTS

This study would have been so much more difficult without the support of many people and impossible without some. The first of these is Isabel Weiersbye to whom I am immensely grateful, and AngloGold Ashanti who allowed me access to the grasslands on their two mining properties. Funding for this research came from Isabel's Ecological Engineering and Phytotechnology Programme (EEPP) and AngloGold Ashanti. Thank you to Michael Sears who has patiently provided superb guidance. Thank you also to Deanne Drake who's penetrating questions around nutrient cycling made me think deeply about my fieldwork. Thank you to David Tongway who has tirelessly answered all questions and provided much tutoring in landscapes and LFA during weeks in the field together. Thank you also to Jenny Botha, Sashnee Raja and Innocent Rabohale without whom, I would surely have lost my marbles a long time ago. Thank you also to Benjamin Oageng who helped with soil cores on some transects at Vaal River. Without the support of two very special people, Mary Scholes and Kay Moss, I definitely would have found it much harder to have reached this point in the journey. Also thank you to the staff and students of the School of Animal, Plant and Environmental Science, you have all taught me an immense amount in my time with the school. And finally a very big thank you to my mother, who tackled the first draft with great relish and is mightily relieved to see this long-held goal come to fruition, and to my father who has always supported me.

CONTENTS

	Page
Abstract	iii
Acknowledgements	v
List of Figures.....	ix
List of Tables	xv
List of Acronyms	xviii
1. Introduction.....	1
1.1 General Introduction	1
1.2 Pilot Study.....	4
1.3 Aims and Objectives	6
2. Landscape Function Analysis and Ecological Results	9
2.1 Aims and Objectives	10
2.1.1 Landscape Function Analysis and Ecological Considerations.....	10
2.1.2 Landscape patches	14
2.1.3 Soil Stability	15
2.1.4 Soil Infiltration	16
2.1.5 Nutrient Cycling.....	16
2.2 Methods and Materials	21
2.2.1 Research Sites, Climate, Vegetation, Geology and Soils	21
2.2.2 Landscape Function Analysis Transects	24
2.2.3 Measurement and Sampling of Patch Types.....	25
2.2.4 Vegetation and Soil Measurements	26
2.2.5 Statistical Analysis	31
2.3 Results	33
2.3.1 Patch Types.....	33
2.3.2 Transects and Transect Indices	34
2.3.3 Patch Measures from Vaal River	37
2.3.4 Patch Measures for West Wits.....	42
2.3.5 Spatial Autocorrelation of Patch Measure	43

	Page
2.3.6 Correlations between Patch Variables	46
2.3.7 Principal Components Analysis	48
2.4. Discussion	54
2.4.1 Patches Types	54
2.4.2 Transects Measurements and Transect Indices	61
2.4.3 Patch Indices and other Variables at Patch Scale	63
2.5 Conclusion	82
 3. Predicting LFA from Ground-based Hyperspectral Remote Sensing	 84
3.1.1 Hyperspectral Remote Sensing	84
3.1.2 Aims and Objectives	86
3.2 Methods and Materials	88
3.2.1 Introduction.....	88
3.2.2 Spectral Sampling in the Field	88
3.2.3 Spectral Preparation.....	88
3.2.4 The Ecological Data	89
3.2.5 Partial Least Squares Regression Modelling.....	89
3.3 Results	92
3.3.1 Patch Models for West Wits.....	92
3.3.2 Patch Models for Vaal River	114
3.3.3 Predicting the LFA Stability Index.....	135
3.3.4 Predicting the LFA Infiltration Index.....	143
3.3.5 Predicting the LFA Nutrient Cycling Index.....	151
3.3.6 Predicting Ecological Variables with PLSR and Spectral Data	159
3.4 Discussion	169
3.4.1 Patch Models	169
3.4.2 Predicting Soil Stability, Infiltration and Nutrient Cycling.....	172
3.4.3 The PLSR Models Predicting the Ecological Variables.....	174
3.5 Conclusion	180

	Page
4. Predicting LFA indices from Hyperion Hyperspectral Imagery	182
4.1.1 The Hyperion Hyperspectral Sensor.....	183
4.2 Methods and Materials	185
4.2.1 Preparation of the LFA Data	185
4.2.2 Preparation of the Hyperion Hyperspectral Data	186
4.2.3 PLSR Statistical Modelling and Mapping	190
4.3 Results	192
4.3.1 The Pixel Spectra Used in PLSR Modelling	192
4.3.2 Predicting the LFA Soil Stability Index	196
4.3.3 Predicting the LFA Infiltration Index.....	207
4.3.4 Predicting the LFA Nutrient Cycling Index.....	214
4.3.5 Predicting the LFA landscape Organization Index.....	219
4.3.6 Predicting the LFA Patch Area Index	226
4.3.7 Predicting Above-Ground Biomass.....	231
4.4 Discussion	238
4.4.1 The Hyperion Hyperspectral Image.....	238
4.4.2 The LFA Indices: Stability, Infiltration and Nutrient Cycling.....	243
4.4.3 The LFA Landscape Organization and Patch Area Indices.....	255
4.4.4 Above-Ground Biomass.....	261
4.5 Conclusion	263
5. Conclusions and Directions	265
5.1 Conclusions.....	265
5.2 Recommendations going forward	268
6. Reference.....	271

LIST OF FIGURES

Figure	Page
1.1 A flow diagram showing the steps required to achieve the objectives of this research	8
2.1 The eleven Soil Surface Assessment Indices (SSAI) measured in the field using the ranking scales shown	12
2.2 Map showing the two mining regions relative to Johannesburg and the provinces of Gauteng, North West and Free State	22
2.3 Vaal River Mining Region showing the distribution of transects across a portion of the mining region	23
2.4 The position of transects and patches on the upper Varkenslaagte area in the West Wits Mining Region.....	24
2.5 Transect 10 and 11 at Vaal River	27
2.6 Transect 3 at West Wits	28
2.7 Bird's eye view of patch types described at West Wits	29
2.8 Bird's eye view of patch types described at Vaal River.....	30
2.9 A correlogram of summary and index variables for each transect.....	37
2.10 Correlograms showing correlations between variables for patches from (a) Vaal River and (b) West Wits.....	49
2.11 Biplots of the first two components from the PCA for (a) Vaal River and (b) West Wits.....	50
2.12 A bare patch from Transect 15 just before it was burnt.....	58
2.13 The different bare grass patches described in the text	60
2.14 (a) <i>Seriphium plumosum</i> burning as a fire swept through transect vr15. (b) Remnant burnt grass stubble and bare patch in the vicinity of an <i>S. Plumosum</i> shrub on ww3 transect	66
2.15 The pedestalled nature of many grass tufts at Vaal River	72
3.1 Cross-validated RMSEP using LOO for all components of a PLSR calibration model with five patch types and ground-based hyperspectral measurements collected from West Wits.....	94

Figure	Page
3.2 Calibration (left column) and validation (right column) predictions from an 18-component model for five patch types at West Wits	98
3.3 The average spectrum for each patch type from West Wits	100
3.4 Loadings for the first 3 components of a five-patch model from West Wits	101
3.5 Pearson correlation coefficients for wavelengths from spectra for all sample points at West Wits	103
3.6 Pearson correlation coefficients for spectra from individual sample points from West Wits organized into patch types	103
3.7 Score plots for the first three components of a five-patch model from West Wits which together account for 98.6% of the spectral variation.....	105
3.8 Regression coefficients from an 8-component model predicting patch types from ground-based hyperspectral measurements	106
3.9 Regression coefficients from an 18-component model predicting patch types from ground-based hyperspectral measurements	107
3.10 Calibration (left column) and validation (right column) predictions from a 6-component model for three patch types at West Wits	112
3.11 Calibration (left column) and validation (right column) predictions from a 9-component model for three patch types at West Wits	113
3.12 (a) Correlogram for all spectra from Vaal River. (b) Correlogram for spectra from Vaal River with two gypsum patches removed from the data set	116
3.13 Pearson correlation coefficients for wavelengths from spectra from all sample points at Vaal River	117
3.14 Mean spectra for the seven patch types sampled at Vaal River.....	118
3.15 Mean spectra for the three patch types that have some samples with efflorescent crusts	119
3.16 Scores plot for the first three components of a PLSR model predicting seven patch types with all points in the calibration dataset included.....	121
3.17 Calibration (left column) and validation (right column) predictions for three patch types from Vaal River using a 5-component model	124
3.18 Predictions for a 14-component model for Vaal River with no gypsum affected patches and four patch types	131

Figure	Page
3.19 Bare grass patches at Vaal River that overlap with predictions for other patch types	132
3.20 Predictions for a 13-component model for Vaal River with all sample points and four patch types	134
3.21 PLSR model predictions for LFA stability index with calibration (left column) and validation (right column) data from West Wits (top row) and Vaal River (bottom row)	137
3.22 The first three loadings for the models predicting stability at (a) West Wits and (b) Vaal River	140
3.23 Pearson correlogram between the mean spectra for each patch type and (a) the first four loadings for the LFA stability model for West Wits, and (b) the first four loadings for the LFA stability model derived from Vaal River	141
3.24 PLSR model predictions for the LFA infiltration index with calibration (left column) and validation (right column) data from West Wits (top row) and Vaal River (bottom row)	145
3.25 Loadings from the PLSR model predicting infiltration at (a) West Wits and (b) Vaal River	148
3.26 Pearson correlations between the mean spectra for each patch type and (a) the first five loadings for the LFA infiltration model for West Wits, and (b) the first twelve loadings for the infiltration model derived from Vaal River	149
3.27 PLSR model predictions for LFA nutrient cycling index with calibration (left column) and validation (right column) data from West Wits (top row) and Vaal River (bottom row)	153
3.28 Nutrient cycling loadings for (a) West Wits and (b) Vaal River	156
3.29 Nutrient cycling at a) West Wits and (b) Vaal River	157
3.30 Calibration (left column) and validation (right column) predictions of (a) fresh above-ground biomass for West Wits and (b) Vaal River, and (c) fresh root biomass for Vaal River	162
3.31 Calibration (left column) and validation (right column) predictions for SOM from (a) West Wits and (b) Vaal River	164

Figure	Page
3.32 Calibration (left column) and validation (right column) predictions for soil moisture from (a) West Wits and (b) Vaal River.....	165
3.33 Calibration (left column) and validation (right column) predictions for (a) soil electrical conductivity at Vaal River, (b) organic soil nitrogen at West Wits, and (c) dry gravel content of soil at Vaal River.....	167
4.1 Partitioning of LFA transects into sub-transects corresponding to pixels	186
4.2 November 2013 Hyperion image of a portion of Vaal River mining region displayed as a SWIR image using the bands at 2194, 1649 and 1074 nm	187
4.3 Surface reflectance spectra from a variety of features identified in the Hyperion image in Figure 4.2	189
4.4 The effect of applying the Colibri smoothing algorithm at 1 to 4 iterations on two spectra extracted from the Hyperion image.....	190
4.5 Spectra from pixels in the Vaal River 2013 Hyperion hyperspectral image corresponding with the LFA transects and used in the PLSR modelling	193
4.6 (a) Correlogram showing correlation between 128-band unsmoothed Hyperion spectra from pixels associated with the LFA transects. (b) Correlogram showing correlation between wavelengths (bands) in the 128-band unsmoothed Hyperion data cube.....	194
4.7 Correlation between the response variables used in the PLSR modelling	195
4.8 The distribution of LFA soil stability index values for each transect at Vaal River	196
4.9 Measured LFA stability index values versus predicted stability values using a 5-component PLSR model using 40 VNIR bands.....	199
4.10 LFA stability index predicted by a 5-component PLSR model using 40 unsmoothed bands from the VNIR Hyperion spectrum.....	200
4.11 LFA stability index predicted by a 5-component PLSR model using 40 VNIR bands after one iteration of the smoothing algorithm	202
4.12 Regression coefficients for (a) a five-component PLSR model with no spectral smoothing, and (b) with one iteration of the smoothing algorithm prior to PLSR modelling	203
4.13 Loading values for the VNIR PLSR models.....	204

Figure	Page
4.14	Distribution of LFA infiltration index values for pixels from each transect at Vaal River..... 207
4.15	LFA infiltration index predicted by a 3-component PLSR model 210
4.16	Measured versus PLSR predicted LFA infiltration index values using 35 bands from the VNIR block..... 212
4.17	LFA infiltration prediction maps 213
4.18	Distribution of LFA nutrient cycling index values for pixels from each transect at Vaal River..... 214
4.19	Measured versus PLSR predicted LFA nutrient cycling index values using 40 bands from the VNIR block..... 217
4.20	LFA nutrient cycling index prediction maps 218
4.21	Distribution of LFA landscape organization index (LOI) values for pixels from each transect at Vaal River 219
4.22	Measured versus PLSR predicted LFA landscape organization Index 222
4.23	LFA landscape organization index prediction map 223
4.24	Measured versus predicted LFA LOI values..... 224
4.25	LFA LOI index prediction map..... 225
4.26	Distribution of LFA patch area index (PAI) values 226
4.27	Measured versus PLSR predicted LFA patch area index (PAI)..... 229
4.28	LFA patch area index (PAI) prediction maps 230
4.29	Regression coefficients for (a) a 2-component unsmoothed full spectrum PLSR model 232
4.30	Distribution of dry biomass values (kg.m^{-2}) for pixels matching five transects at Vaal River 233
4.31	Measured versus PLSR predicted dry biomass (kg/m^2) 235
4.32	Dry biomass (kg/m^2) prediction maps..... 236
4.33	Drainage (red) derived from an ASTER GDEM overlaid on the prediction map for the nutrient cycling index 250
4.34	Occurrence and extent of fire during 2013 253
4.35	The layout of patches along a transect in a pixel space (vr14 pixel 4)..... 255

Figure	Page
4.36 (a) Fire approaching transect vr15 in 2011. (b) Transect vr15 after the fire.	
(c) Post-fire recovery during September 2011 in a fire break near transect vr7.	
(d) Recovery of grassland during October 2011 from an undated fire east of vr13, vr14 and vr15	259

LIST OF TABLES

Table	Page
2.1 Names, codes and definitions of patch types described on transects at West Wits and Vaal River sites	26
2.2 The total number of each patch type recorded on transects at each mine site.....	34
2.3 A summary of indices and measures calculated from the LFA data for each transect.....	36
2.4 Comparisons of transects from Vaal River showing the mean and standard error calculated from measurements collected from individual patches within each transect	40
2.5 Comparisons between patch types from Vaal River showing the mean and standard error calculated from individual patches.....	41
2.6 Comparisons of transects from West Wits showing the mean and standard error calculated from individual patches within each transect	44
2.7 Comparisons between patch types from Vaal River showing the mean and standard error calculated from individual patches.....	45
2.8 Spatial autocorrelation using Moran's I Autocorrelation index for shoot biomass at Vaal River as a whole and for individual transects from Vaal River	46
2.9 Principal Components Analysis for the Vaal River data showing the first five principal components.....	51
2.10 The first five components from the Principal Components Analysis of the West Wits variables	52
3.1 Wavelengths removed from the spectra prior to PLSR modelling	89
3.2 Cross-validated (CV) RMSEP using leave-one-out (LOO) estimates for calibration models predicting patch types at West Wits from ground-based hyperspectral measurements.....	93
3.3 Number of objects in calibration and validation datasets for five patch types from West Wits.....	97
3.4 Cross-validated (CV) RMSEP using leave-one-out (LOO) estimates for calibration models predicting revised patch types at West Wits from ground-based hyperspectral measurements	109
3.5 Number of samples in each patch type from Vaal River.....	115

Table	Page
3.6 Cross-validated RMSEP using LOO estimates for calibration models predicting patch types at Vaal River from ground-based hyperspectral measurements using all points.....	120
3.7 Cross-validated RMSEP using LOO estimates for calibration models predicting patch types at Vaal River from ground-based hyperspectral measurements	130
3.8 Prediction calibration and validation results for three forms of the four-patch combination model	133
3.9 PLSR models predicting LFA stability for West Wits with calibration data (n = 57) and validation data (n = 19)	136
3.10 PLSR models predicting the LFA stability index for Vaal River from calibration data (n = 180) and validation data (n = 59)	138
3.11 PLSR models predicting LFA Infiltration index for West Wits with calibration data (n = 57) and validation data (n = 19)	144
3.12 PLSR models predicting the LFA infiltration index for Vaal River from calibration data (n = 180) and validation data (n = 59).....	146
3.13 PLSR models predicting LFA nutrient cycling index for West Wits with calibration data (n = 57) and validation data (n = 19).....	152
3.14 PLSR models predicting the LFA nutrient cycling index for Vaal River from calibration data (n = 180) and validation data (n = 59).....	154
3.15 PLSR Model calibration and validation statistics for various ecological variables for above-ground biomass and soil sampled from West Wits	160
3.16 PLSR Model calibration and validation statistics for various ecological variables for above-ground biomass or soil sampled from Vaal River	161
4.1 Number of transects and sample points falling on Hyperion imagery from Vaal River, the maximum number of bands in the Hyperion data cube used for PLSR modelling and prediction mapping.....	191
4.2 Exploratory models predicting the LFA stability index from the Vaal River Hyperion hyperspectral data cube	198
4.3 PLSR model results using Hyperion spectra to predict the LFA stability index using 31 bands from the VNIR region of the spectra (488 – 610, 630 – 722, 773 – 834, 854 nm).....	206

Table	Page
4.4	Exploratory models predicting the LFA infiltration index from the Vaal River Hyperion hyperspectral data cube 209
4.5	PLSR Model predicting the LFA infiltration index using 35 bands (508 - 610, 630 – 864 nm) from the VNIR portion of the Hyperion spectrum 211
4.6	Exploratory models predicting the LFA nutrient cycling index from the Vaal River Hyperion hyperspectral data cube 216
4.7	Exploratory models predicting the LFA landscape organization index from the Vaal River Hyperion hyperspectral data cube 220
4.8	PLSR models predicting the LFA landscape organization index using 13 select bands from the near-SWIR region (1518 – 1679 nm) of the Hyperion spectra 221
4.9	Exploratory models predicting the LFA patch area index (PAI) from the Vaal River Hyperion hyperspectral data cube 227
4.10	Exploratory models predicting the dry biomass (kg/m ²) on each pixel from the Vaal River Hyperion hyperspectral data cube 234

LIST OF ACRONYMS

AGA – AngloGold Ashanti
AMD – Acid mine drainage
ASD – Fieldspec Pro FR spectroradiometer built by Analytical Spectral Devices
CV – Cross-validated
DAR – Data Acquisition Request
DCA – Detrended Correspondence Analysis
EC – Electrical conductivity for soils
EMP – Environmental Management Plan
EnMAP – Environmental Mapping and Analysis Program
EO-1 – Earth Observation 1 satellite
EROS – Earth Resources Observation and Science
FLAASH – Fast Line-of-sight Atmospheric Analysis of Hypercubes
FOV – Field of view
GV – Green vegetation
HRS – Hyperspectral remote sensing
HSI – Hyperspectral imager
HyspIRI – Hyperspectral InfraRed Imager
LAI – Leaf area index
LFA – Landscape function analysis
LOI – Landscape organization index
LOO – Leave-One-Out cross-validation
MASL – Metres above sea level
MESMA – Multiple endmember spectral mixture analysis
MSEP – Mean sum of squares
NASA – National Aeronautics and Space Administration
NDVI – Normalized difference vegetation index
NIR – Near infrared wavelengths from 700 to 1300 nm,
NPV – Non-photosynthetic vegetation
PAI – Patch area index
PC – Principle component

PCA – Principal Components Analysis

pH – Acidity or alkalinity scale

PLSR – Partial least squares regression

PRESS – Predicted residual sum of squares

PRISMA – Precursore IperSpettrale della Missione Applicativa hyperspectral imager

RMSEP – Root mean square error of prediction

RPD – Ratio of performance to deviation

RPIQ – Ratio of performance to interquatile range

SMA – Spectral mixture analysis

SOC – Soil organic carbon

SOM – Soil organic matter

SSAI – Soil surface assessment indicators

SSI – Soil surface indices

SVD – Singular value decomposition

SWIR – shortwave infrared wavelengths from 1300 to 2500 nm

TSF – Tailings storage facility

USGS – United States Geological Survey

VI – Vegetation indices

VNIR – Visible and near-infrared

VNIRA – Visible and Near Infrared Analysis

VR – Vaal River Mining Region near Klerksdorp

WW – West Wits Mining Region near Carletonville

1. Introduction

1.1 General Introduction

Mining has negative impacts on surface environments through acid mine drainage (AMD) and ground-water contamination (Tutu *et al.*, 2008, Akcil and Koldas, 2006, Winde and Sandham, 2004, Naicker *et al.*, 2003), salinization of soil and water (Lucas *et al.*, 2010, Carvalho *et al.*, 2005, Schaaf *et al.*, 1999), and aeolian transport and deposition of contaminants (Aide, 2009, Moreno *et al.*, 2007). A landscape exposed to stress may show changes in plant physiology (Menon *et al.*, 2007, Weiersbye and Witkowski, 2007, Weiersbye and Witkowski, 2003, Weiersbye and Witkowski, 2002, Horler *et al.*, 1980), a reduction in leaf area index (LAI) or canopy cover (Skousen *et al.*, 2006, Steven *et al.*, 1992), or a change in species composition (Fernández *et al.*, 2009, Fischer and Lindenmayer, 2007, Guerrero-Campo and Montserrat-Martí, 2004, Lottermoser *et al.*, 1999). Furthermore, the stress on a landscape may express as changes in biogeochemical processes such as infiltration, nutrient cycling, soil erosion and deposition, and a reduction in vegetation patch size (Kakembo, 2009, Maestre and Cortina, 2004). Therefore there is a need to understand landscape-scale responses to stress and develop techniques for monitoring and mapping environmental impacts around mining operations that will help direct management interventions.

Landscape Function Analysis (LFA) is one technique that has been proposed for monitoring ecological function, degradation and rehabilitation processes in mining environments (Tongway and Hindley, 2004) and has been accepted by some regulators and mining companies (Lacy *et al.*, 2008). Briefly, LFA involves portioning a landscape into patch and inter-patch types and measuring the length and width of these along a suitable representative transect (Tongway and Hindley, 2004). Thereafter various aspects of the soil surface, composition and cover is assessed on 11 indices known as soil surface assessment indicators or SSAIs (Tongway and Hindley, 2004). These assessments are used to calculate relative indicators of soil stability, infiltration or nutrient cycling on a percentage scale where low values indicate lower functionality for that index (Tongway and Hindley, 2004). Other indices are calculated from the patch and inter-patch sizes, such as the number of patches per unit length of transect, the patch area index (PAI) and the landscape

organization index (LOI) (Tongway and Hindley, 2004). These indicators should be compared to an analogue site which acts as a reference, or to a time series of measurements to indicate whether a site is increasing or decreasing its functionality and ability to retain or trap resources (Tongway and Hindley, 2004). However, many of these techniques for ecological monitoring, including LFA, are limited by the heterogeneity of an environment, the available man-hours, and the localized nature of monitoring techniques (Pickup *et al.*, 1998). Furthermore, in South Africa, commercial mine operations generally occupy large properties where a significant portion of the area within a mining lease is still, to a greater or lesser degree, in a natural state. These mine lands may include a variety of ecosystems and biomes on one site such as grassland, savanna, riverine forests, streams and wetlands which occupy the spaces between mine activities, and thus both the large scale of the operations and the heterogeneity make it difficult to apply LFA in a comprehensive manner.

Hyperspectral remote sensing (HRS) techniques may be an excellent alternative solution for identifying and monitoring ecological degradation on these extensive and complex landscapes. Data can be collected in a single flight or satellite overpass, and include the whole expanse of the mine property. HRS techniques have high spectral resolution which allows the production of detailed maps which are often more easily understood than tables and descriptions of ecological outputs. However, HRS techniques have some limitations. The processes or features to be monitored need to directly or indirectly have a spectral signature (Stenberg *et al.*, 2010). For example, chlorophyll has strong absorption features in the visible portion of the spectrum which has been related to plant physiological status (Naumann *et al.*, 2010, Naumann *et al.*, 2008, Carter and Knapp, 2001) and leaf area index (LAI) (Wu *et al.*, 2010, Clevers, 1989, Asrar *et al.*, 1986), whereas pH is a soil property with no spectral absorption features but can often be correlated with soil organic carbon (SOC) and clay content (Stenberg *et al.*, 2010). In addition, spatial resolution is a critical property of remote sensing imagery, especially when using it to interpret or predict biogeochemical processes and structure (Chabrillat *et al.*, 2002). For example, each pixel in a Hyperion image represents 900 m² (approximately 30 x 30 m depending on the elevation of the surface) on the ground. The spectrum for this pixel is a combination of all the features in that area that have a reflectance signature. However, biological structure and biogeochemical processes may have major variation over short distances (Ludwig *et al.*, 2000, Bolton Jr *et al.*, 1993,

62 Forman and Godron, 1981). The issues of spatial resolution and variability will be discussed
63 in more detail in later chapters. Finally, although HRS techniques are real physical
64 measurements, mathematical functions are being used to relate these to processes active in
65 the landscape, so the output from any modelling analysis needs to be carefully interpreted.
66 It is critically important that the ecological context of the data is accurately defined and
67 measured if meaningful ecological information is to be extracted from HRS data.

68 The Earth Observation 1 (EO-1) satellite was launched by NASA in 2000 and carries the
69 Hyperion hyperspectral sensor which can record 220 wavebands in the visible, near-infrared
70 (VNIR) and shortwave infrared (SWIR) with a ground pixel resolution of approximately 30 x
71 30 m (Ungar *et al.*, 2003, Pearlman *et al.*, 2001). Having completed its science mission, the
72 EO1 satellite and its sensors are available on request to the USGS. This study focuses on the
73 use of hyperspectral imagery acquired from the Hyperion sensor to predict LFA indices
74 (Tongway and Hindley, 2004) derived from a natural grass land using partial least squares
75 regression (PLSR) models (Mevik and Wehrens, 2007, Wold *et al.*, 2001, Geladi and
76 Kowalski, 1986a, Geladi and Kowalski, 1986b). The sites for this research were two deep-
77 level gold mining environments belonging to AngloGold Ashanti: West Wits near
78 Carletonville, and Vaal River near Klerksdorp, approximately 80 km and 160 km respectively
79 to the west of Johannesburg, South Africa. As these two gold mines are underground with
80 shafts accessing the ore bodies, the surface environments have extensive grasslands and
81 savannas interspersed with wetlands and streams, and in the case of Vaal River, a major
82 river with extensive riverine forest and flood plain ecosystems within the property
83 boundaries. This study focused on the grassland components because these are essentially
84 two dimensional systems whereas the savanna and riverine forests increase the complexity
85 in the landscape through multiple canopy layers, increased shade effects, and difficulty in
86 acquiring ground-based spectral data. Although ground measurements were collected from
87 both mining regions, only 2004 Hyperion imagery was available for West Wits so the
88 satellite-based analysis was restricted to Vaal River as the USGS did provide a current image
89 for this mining region.

90 1.2 Pilot Study

91 In 2007/2008, a pilot study was conducted at Vaal River (VR, Klerksdorp) and West Wits
92 (WW, Carletonville) Mining Regions to test if LFA indices could be predicted from ground-
93 based HRS data using PLSR modelling (Furniss *et al.*, 2009, Furniss, 2008). This work formed
94 an independent test of work carried out in Australia by Ong *et al.* (2008). During this pilot
95 study, twenty-three spectral vegetation indices (VI) were tested for their potential to predict
96 LFA indices using linear regression. These indices were selected on the basis of their
97 potential to indicate various plant properties such as chlorophyll, carotenoids, anthocyanins,
98 cellulose, lignin and plant water content. Data was collected during the winter of 2007 and
99 late summer of 2008. Sites for this pilot study included four vegetation types at VR: wet
100 grasslands, non-rocky grasslands, rocky grasslands and woodland or woody shrub; and a
101 single vegetation type at WW which was non-rocky grassland (Furniss *et al.*, 2009, Furniss,
102 2008). Within each vegetation type six plots were identified at two levels of disturbance:
103 high and low disturbance, on the basis of presence or absence respectively of erosion
104 and/or exposed soil. The LFA methodology was adapted on the basis of the pixel resolution
105 of an airborne campaign conducted in 2005 although the airborne data were not available
106 for this pilot study. Pixel size in the airborne data covered 3 x 3 m on the ground and
107 therefore plots of 10 x 10 m were used which would include 9 pixels per plot from the
108 airborne data. One adaptation of the LFA method was not to use transects and therefore no
109 patch identifications or measurements were recorded. Instead, within these 10 x 10 m plots,
110 25 circular quadrats were laid out in a regular network. These circular quadrats were the
111 sampling unit. Each quadrat was 50 cm in diameter based on the field-of-view (FOV) of the
112 spectrometer sensor at 1 m above ground-level (42 cm diameter). Of the full LFA
113 methodology described earlier, only the eleven SSAI were measured on each quadrat (Table
114 1). These eleven measurements were resolved into the three soil surface indices (SSI) of
115 stability, infiltration and nutrient cycling. It was these three indices that were modelled
116 against the ground-based HRS data and the VIs (Furniss *et al.*, 2009, Furniss, 2008).

117 The VIs consistently produced poor results with the best coefficient of determination for the
118 winter data obtained with the lignin index (Serrano *et al.*, 2002) predicting nutrient cycling
119 ($R^2 = 0.36$), and for summer data, the NDVI predicting infiltration ($R^2 = 0.40$). PLSR models
120 derived for the winter data performed slightly better with the best R^2 value (0.54) obtained

121 from a 15-component model when predicting the nutrient cycling index. The best PLSR
122 predictions obtained with summer data were models from non-rocky grassland that had $r^2 =$
123 0.70 when predicting the stability index. However, closer examination of the PLSR models
124 showed some weakness in these models. In all PLSR models derived with both the summer
125 and winter data, the first two components explained in excess of 95% of the variability in
126 the spectral data (Furniss, 2008). These same two components explained less than 30% of
127 the variability in the LFA data in all models and generally many more components were
128 required in a PLSR model to explain a reasonable amount of the variability in the LFA data.

129 The results from this pilot study suggest that there was little or weak correlation between
130 the three LFA indices and HRS data, and that the relationship between reflectance
131 measurements and ecosystem processes in these two environments are not well
132 represented by LFA indices under the experimental design employed. LFA is a method in
133 which the landscape components structure the sampling layout on the basis of patch
134 definitions defined by the biogeochemical processes active in the landscape (Tongway and
135 Hindley, 2004). The SSIs are then weighted according to their respective patch contribution
136 to the transect, thus resulting in two SSI measurements: one for each patch type, and one
137 for each patch type weighted by the contribution of that patch type's length to the total
138 transect length. Furthermore, using the full LFA transect method gives a number of metrics
139 of the landscape structure that were not available under the 2007/8 experimental design.
140 These include patch types, patch widths, lengths, patch area index (PAI), landscape
141 organization index (LOI), and inter-patch lengths (Tongway and Hindley, 2004). Thus in the
142 LFA measurements, there is a wealth of contextual information which was not recorded
143 under the experimental design applied in the pilot study. As mentioned previously the LFA
144 technique employed in this pilot study was adapted to the spatial dimensions of the 2005
145 HRS data. From the above discussion, it is suggested that, if the purpose is to develop
146 techniques for mapping the effects of mining on ecological functioning in natural systems,
147 the ecological sampling should be based on an experimental design more appropriate to the
148 ecological processes being mapped, and the techniques being employed to characterize
149 these ecological processes. The HRS methods should then be adapted to the experimental
150 design underlying the ecological sampling. To this end, this new study uses the LFA transect
151 to structure the spatial arrangement and unit of sampling and then applies the

hyperspectral measurements, both ground-based and satellite based, to predict the LFA measurements.

1.3 **Aims and Objectives**

The primary aim of this research is to test a method of producing ecological risk maps based on LFA indices of the surface grasslands of two deep-level gold-mining properties in a Southern African Highveld context. These maps are derived from Hyperion hyperspectral imagery acquired from the EO-1 satellite and predict five LFA indices and above-ground biomass. The five LFA indices are soil stability, infiltration, nutrient cycling, landscape organization (LOI) and patch area (PAI) indices. The first three indices focus on landscape processes and the last two on landscape organization. Partial Least-Squares Regression (PLSR) modelling (Mevik and Wehrens, 2007, Wold *et al.*, 2001, Geladi and Kowalski, 1986a, Geladi and Kowalski, 1986b) was used to derive calibration models using spectral measurements to predict the indices. Models were then tested against a validation data set before the final model was used to derive the maps from a subset of the Hyperion image which covered two extensive grasslands around the Vaal River mining operations. To achieve this, the research was broken down into the following three objectives:

- (i) Collection and evaluation of the LFA data, above-ground biomass and various edaphic properties measured from soil core samples. As LFA has a large component dependent on the practitioner's skill in allocating a variety of observations to their respective scale, the LFA measurements were evaluated against the above-ground biomass measurements and various edaphic measurements derived from the soil cores. Chapter 2 focuses on this analysis of the LFA measurements and environmental data.
- (ii) Modelling with PLSR to predict LFA indices and other environmental measurements from ground-based hyperspectral measurements. Ground-based hyperspectral measurements were collected from each sample point during the field campaign and these were used to explore the LFA measurements and environmental data using PLSR models at a scale related to the ground-based spectral measurements, i.e. the 42 cm diameter field-of-view of the spectral measurements when the sensor was hand-held at approximately 1 m above

182 ground level. The results for all these models are presented in Chapter 3
183 although the discussion for those indices repeated with the Hyperion imagery
184 (the five LFA indices and biomass) is included in Chapter 4.

185 (iii) Derivation of maps for five LFA indices and above-ground biomass using PLSR
186 models and the Hyperion hyperspectral data. The LFA and above-ground biomass
187 data was rescaled to fit the Hyperion pixels matching those points on the
188 transects. The pixels matching the transects were split into two data sets with
189 one used to derive PLSR calibration models. The calibration models were then
190 validated against the second data set and the final model was then applied to the
191 whole Hyperion image to produce the prediction maps. The results of this
192 modelling and the respective prediction maps are presented in Chapter 4.

193 The processes followed in achieving these objectives are summarized in the flow diagram
194 presented in Figure 1.1. The scale of ground-spectral measurements is vastly different from
195 the size of Hyperion spectral measurements where a pixel covers approximately 900 m².
196 Ground-spectral measurements were collected from a circular footprint with an
197 approximate diameter of 42 cm covering a surface area of 0.139 m². Thus, models derived
198 from ground measurements of LFA and spectra were regarded as different models to
199 models derived from Hyperion pixel spectra (Figure 1.1).

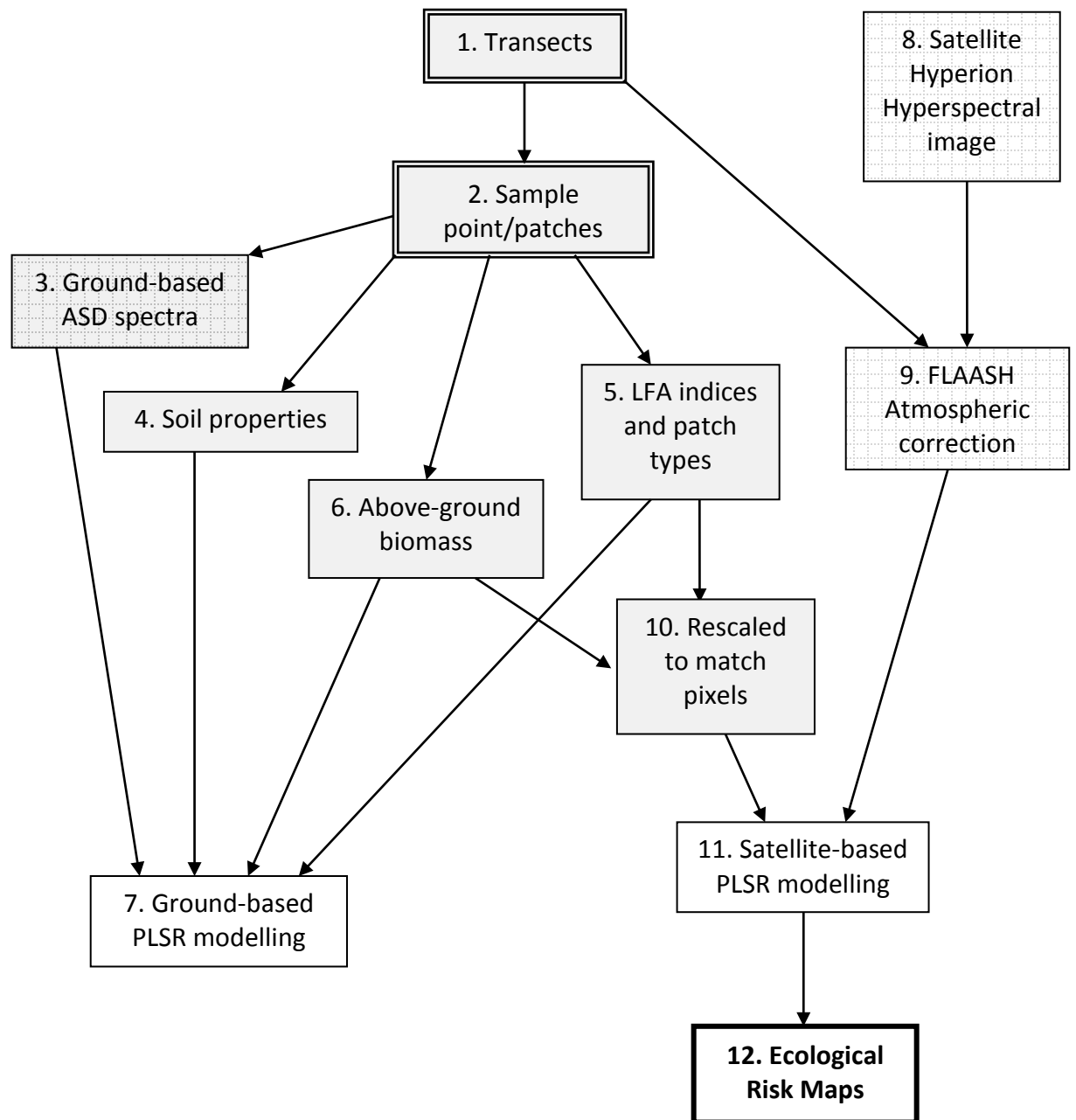


Figure 1.1 A flow diagram showing the steps required to achieve the objectives of this research. The double boxes represent transects (1) along which sample points (2) were identified. Ground-based spectra (3), soil cores (4) LFA indices and above-ground biomass (6) were sampled at each point and used to derive PLSR models at a 42 cm diameter scale. The sample points (2) were also used to select pixels from the Hyperion image (8) after it had been atmospherically corrected (9). The LFA indices (5) and above-ground biomass (6) were rescaled to match the Hyperion pixels (10) and used in PLSR modelling (11) to derive the final risk maps (12). Grey boxes represent samples collected and/or measured in the field or laboratory. The checked boxes reflect spectral data collected either in the field or selected from pixels in the Hyperion image.

247

2. **Landscape Function Analysis and Ecological Results**

248



249 2.1. Aims and Objectives

250 The aim of this chapter is to explore the ecological results of this study, and in particular, to
251 evaluate the LFA indices in the context of other empirical measurements and available
252 literature where these are available. This can be broken down into the following two
253 objectives:

- 254 (i) To evaluate, using field-derived and literature resources, whether the spatial
255 measures of landscape organization, as measured in this study using LFA
256 techniques with a modified definition of patches, reflect the natural landscape
257 organization in terms of patch and inter-patch structure resulting from both
258 natural and unnatural processes active in the landscape.
- 259 (ii) To test the three LFA indicators of soil stability, infiltration and nutrient cycling
260 against literature resources and sampled empirical measures of soil and plant
261 characteristics such as pH, SOM or above-ground biomass obtained during this
262 study.

263 2.1.1 Landscape Function Analysis and Ecological Considerations

264 Spatial resolution is an important aspect of HRS data but is an even more critical issue when
265 quantifying landscapes and ecological processes. This is due to the inherent heterogeneity in
266 processes and structures which often vary, not only spatially and temporally within the
267 same hierarchical level of organisation in a landscape, but may also vary across levels of
268 organisation in a landscape (Lovett *et al.*, 2005). A landscape has been defined as “an area
269 that is spatially heterogeneous in at least one factor of interest” (Lovett *et al.*, 2005).
270 Forman and Godron (1981) defined a landscape as a “kilometres-wide area where a cluster
271 of interacting stands or ecosystems is repeated in similar form.” However, in their scheme,
272 processes and structures of a few metres up to hundreds of metres are at a finer level of
273 scale than a landscape (Forman and Godron, 1981). These structures at a finer scale
274 correspond with the definition of patch given by Lovett *et al.* (2005) which is “an area that
275 differs from its surroundings in structure and function.” Forman and Godron (1981) were a
276 little more explicit, defining patches as “communities or species assemblages surrounded by
277 a matrix with a dissimilar community structure or composition.” Underlying these
278 definitions of patches and landscapes is scale which Lovett *et al.* (2005) defined as the

279 “spatial or temporal dimension of an object or process, characterized by both grain and
280 extent.” Ludwig *et al.* (2000) applied power laws to landscape structures and processes and
281 found that at different scales different power functions could be derived and that this
282 change in power function implied scale-dependent thresholds. These authors suggested
283 that at local scales, biological, geochemical and microtopographic redistribution processes
284 are responsible for patch heterogeneity, whereas at landscape scales, vegetation and soil
285 heterogeneity are driven by geological, hydrological and pedogenic processes (Ludwig *et al.*,
286 2000). They further argue that scale is not just a concern in defining landscape structure and
287 processes but also influences our observation processes. The effect of scale in our
288 observations shapes and limits the landscape structures and processes that may be
289 observed, and therefore our awareness of the context within which these structures and
290 processes occur may be incomplete (Ludwig *et al.*, 2000). This is particularly relevant when
291 the HRS level of observation with Hyperion imagery is from 705 km distant from the target
292 where 900 m² of the Earth’s surface is resolved into a single observation. Whereas, the
293 ecological scale of field measurement as discussed below is collected from a maximum
294 distance from the Earth’s surface of two metres and ranges in area from point measures of a
295 few cm² to transects covering 180 m.

296 Definitions of patches are at the heart of LFA as the underlying philosophy is that processes
297 active in a landscape structure how that landscape organises itself spatially which in turn
298 feeds back into the processes that are active in that landscape. Therefore by defining the
299 patch types and their organisation relative to each other, the processes active in a landscape
300 can be deduced and evaluated. Landscape Function Analysis (LFA) is a monitoring procedure
301 that assesses how effectively a hill slope is operating as a biophysical system based on the
302 use of visual and tactile assessment of indicators of landscape organisation and soil surface
303 characteristics (Tongway and Hindley, 2004). In brief, the LFA technique requires the
304 establishment of a transect along the major gradient (sometimes referred to as a gradsect)
305 and quantifying the patches and inter-patches along this transect (Tongway and Hindley,
306 2004). These patches are quantified in terms of type, length along transect, and width to a
307 maximum arbitrary distance of 5 m either side of the transect (Tongway and Hindley, 2004).
308 Thereafter, the eleven soil-surface assessment indicators (SSAI) are quantified (Figure 2.1)
309 for each of five replicates for each patch type. The practitioner uses sensory clues, mostly

visual, to assign values to each SSAI ranging from 1 – 4 to 1 – 10 depending on the particular SSAI (Figure 2.1). These SSAIs are algorithmically resolved into three soil surface indices (SSI), namely stability, infiltration and nutrient cycling, using specific SSAIs for each index (Figure 2.1). From all the above measurements a number of metrics of landscape

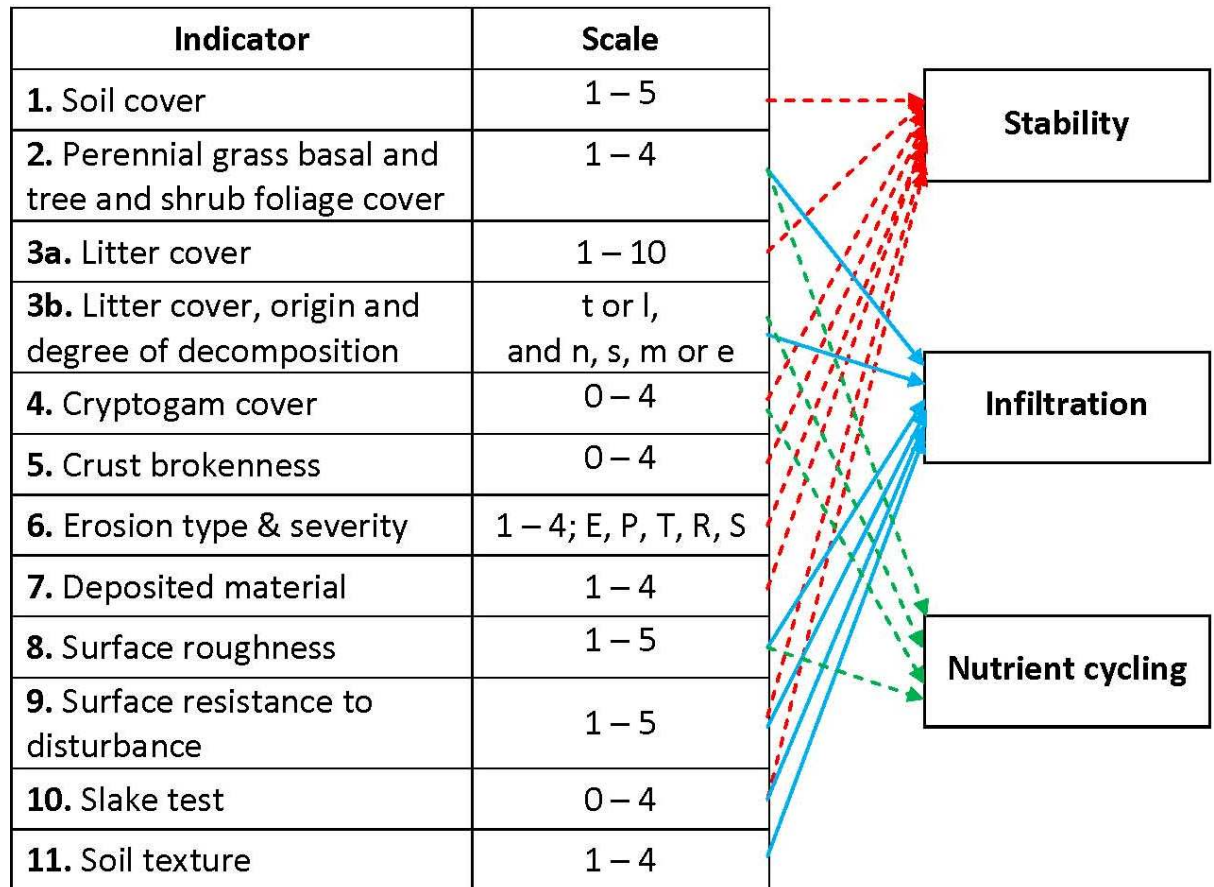


Figure 2.1 The eleven Soil Surface Assessment Indices (SSAI) measured in the field using the ranking scales shown. The SSAIs are then partitioned and resolved into the three Soil Surface Indices (SSI) of stability, infiltration and nutrient cycling, (from Tongway and Hindley, 2004). t = transported, l = local, n = nil, s = slight, m = moderate, e = extensive, E = sheeting, P = pedicles, T = terracettes, R = rills, S = scalds (Courtesy Tongway and Hindley, 2004).

organization and process can be derived. These include patch types, patch lengths and widths, patch area index, landscape organization indices, and inter-patch lengths (Tongway and Hindley, 2004). The SSIs are then weighted according to their respective patch contribution to the transect resulting in two SSI measurements: one for each patch type and one for each patch type weighted by the contribution of the patch length to transect length.

Landscapes that have low or high values are not necessarily degraded or highly functional, but by comparing to suitable reference sites or to previous measurements at the same site, conclusions can be drawn as to whether a site is losing or gaining resources and therefore experiencing degradation or improved functionality through more tightly retained resources (Tongway and Hindley, 2004). A reference site would be one that illustrates what the landscape would look like under optimum conditions when disturbances are at a minimum, and in a rehabilitation context, would illustrate the target outcome for rehabilitation. However such sites can be difficult to locate for many reasons. Also the nature of the disturbance, gold tailings storage facilities for example where the tailings provide quite unnatural edaphic conditions, may render a site highly unlikely to proceed to a state comparable to undisturbed local landscapes. In such conditions, time series may be more appropriate to define a trajectory for a landscape. If time series are not available, sites may be ranked against each other to provide the range of functionality within the landscape. However it needs to be noted that this merely provides a range under the current conditions and does not necessarily reflect the potential functionality a site may achieve. As the purpose of this study was to test the ability to predict LFA and ecological measurements from remote sensing data, and not to evaluate the degree of functionality and disturbance in the study landscape, no reference site or time series were needed.

Most methods developed in South Africa for measuring ecological degradation in rangelands are based on veld condition with the underlying aim of maximizing grazing potential (Tainton, 1999) and therefore may not be suitable for assessment of mine-impacted environments. LFA was developed to monitor changes in landscape heterogeneity and biogeochemical processes in disturbed landscapes in Australia (Tongway and Hindley, 2004). It has since been adapted for monitoring rehabilitation processes in Australian mining environments (Tongway and Hindley, 2004) and has been accepted by Australian mining regulators (Lacy *et al.*, 2008). A limitation is that LFA is a subjective method dependant on

the practitioner's skill and consistency in correctly evaluating the environmental features concerned. It was therefore deemed worthwhile to include more empirical methods of measuring ecosystem processes within the experimental design. This would provide a broader base to the ecological data, and potentially, some independent verification of the LFA indices. The ecosystem properties that LFA identifies are patch size, soil stability, infiltration and nutrient cycling (Tongway and Hindley, 2004). Patch size and type are descriptors of landscape organization and an aid to organizing and categorizing landscape heterogeneity, whereas soil stability, infiltration and nutrient cycling are closely interrelated characteristics of a soil and local environment.

2.1.2 Landscape Patches

From an LFA perspective, patches form the first level in defining landscape sampling units (Tongway and Hindley, 2004). Patch size is a relatively simple concept to measure, but together with patch type, is far more difficult to define (Gustafson, 1998). Patch boundaries as defined in the LFA concept are delineated by an observed change in biogeochemical processes at the soil surface (Tongway and Hindley, 2004). These include changes in vegetation cover, litter, soil erosion and/or deposition. Of the different types of patch that could be defined in a landscape, the degree of bare, exposed ground is most important as this is generally a strong indicator of either disturbance (Bartley *et al.*, 2006) or intense competition for limiting resource (D'odorico *et al.*, 2007, Noy-Meir, 1973). Towards the arid side of a climate gradient, vegetation tends to be distributed through the landscape in clusters of "high vegetation cover" interspersed through a matrix of "low vegetation cover" or bare ground (Aguar and Sala, 1999). The shapes of "high vegetation cover" patches seem to depend on the dominance of wind or water processes (Aguar and Sala, 1999). As aridity decreases, the inter-patch areas decrease in size and abundance and the "high vegetation cover" patches become more complex in physiognomic structure and begin to dominate the landscape. At mesic levels of a climate gradient "high vegetation cover" is the norm (Greig-Smith, 1979), with bare ground being a sign of disturbance (Pyke *et al.*, 2002). In these climates, patches become a mosaic of different and intergrading physiognomic and community structures (Greig-Smith, 1979).

384 2.1.3 Soil Stability

385 The next level of sampling under the LFA concept is the quantification of the processes that
386 occur in and around these patches. Amongst many processes, LFA focuses on the three
387 processes of soil stability, infiltration, and nutrient cycling. Soil stability is a complex
388 concept. Dexter and Czyż (2007) defined soil physical stability as the ability of a soil's
389 microstructure to survive immersion in water. These authors regarded the microstructure
390 and colloidal properties as critical to the healthy functioning of a soil, and argued that clays
391 were the predominant colloidal material in soils they studied, and therefore an important
392 factor in a soil's stability (Dexter and Czyż, 2007). Chaney and Swift (1984) found that
393 organic content of a soil had the strongest correlation with soil aggregate stability.
394 Furthermore they found no correlation between soil aggregate stability and the soil physical
395 constituents of sand, silt and clay in the British soils they studied. In hydraulic threshold
396 studies in the UK, Carling (1997) found that soil organic matter (SOM) was closely correlated
397 with increased resistance to erosion. Tongway and Hindley (2004), in their LFA manual
398 define soil stability "as the ability of a soil to withstand erosive forces, and to reform after
399 disturbance." Erosion is an important indicator of landscape and soil health (Fernández *et al.*,
400 *et al.*, 2009, Guerrero-Campo and Montserrat-Martí, 2004, Lal, 2001, 2000).

401 In LFA, the soil stability index is calculated as the sum of values (maximum 100) scored on
402 specific ranking scales for soil cover, litter cover, biological soil crust cover, soil crust
403 brokenness, erosion type and severity, deposited materials, surface resistance to
404 disturbance, and a slake test (Tongway and Hindley, 2004). From this list it is clear that a
405 landscape's susceptibility to erosion is a complex phenomenon to describe and measure.
406 Multiple factors influence erosion although vegetation cover (Hartanto *et al.*, 2003) and soil
407 properties (Greene and Hairsine, 2004, Carling *et al.*, 1997) play critical roles. Sheet erosion
408 and rain splash, possibly combined with deflation, seem to be the most common types of
409 soil erosion at the two sites where this research was conducted (personal observation, D.
410 Furniss). Sheet erosion is the uniform removal of soil from sloping ground by sheet wash or
411 overland flow (Egboka and Orajaka, 1987), and rain splash is the detachment and movement
412 of soil particles from the soil surface by the impact of rain drops (Biielders and Grymonprez,
413 2010, Legout *et al.*, 2005). Whereas sheet erosion and rain splash are water mediated,
414 deflation is the removal of fine soil particles through wind action (Cornelis and Gabriels,

2004). Soil erosion is a three stage process: (1) detachment, (2) transport and (3) deposition (Lal, 2001). The fractional area of exposed soil however could be used as a first order proxy to susceptibility to erosion (Bartley *et al.*, 2006) notwithstanding the effects of other factors such as slope angle, run-on area, slaking and climate to mention a few. Second order proxies for soil stability would include soil organic content (SOM), clay content, physical constituents, and position in the landscape.

2.1.4 **Soil Infiltration**

Soil infiltration is an important and complex property of a landscape which is closely related to soil stability. Almost all water available to the biological processes within a landscape is in the form of soil water (Daly and Porporato, 2005, Barbour *et al.*, 1987) which comes mainly from precipitation that infiltrates the soil. Furthermore, the precipitation that does not infiltrate a soil or evaporate back to the atmosphere runs off the landscape, and is the principle agent of erosion. Infiltration is affected by many factors such as the amount and type of precipitation reaching the soil surface (Liu *et al.*, 2011), soil structure and composition (Connolly, 1998, Dixon, 1995, Cresswell *et al.*, 1992), biological crusts (Belnap *et al.*, 2005, Eldridge *et al.*, 2000), physical crusts (Casenave and Valentin, 1992), vegetation cover and rooting systems (Bogner *et al.*, 2010, Bhark and Small, 2003), litter (Hartanto *et al.*, 2003), microorganisms and soil fauna (Morales *et al.*, 2010), and the moisture content of the soil (Ben-Hur and Lado, 2008, Zhou *et al.*, 2008). In LFA, a proxy measure of infiltration is calculated as the sum of values (again to a maximum of 100) scored on specific ranking scales for perennial basal grass and shrub and tree cover, litter cover and its origin and degree of decomposition, surface roughness, surface resistance to disturbance, a slake test, and soil texture.

2.1.5 **Nutrient Cycling**

Nutrient cycling is the movement of chemical elements that are essential to organisms (Barbour *et al.*, 1987). At a landscape scale it is a complex interplay of biological, physical and historical factors. All nutrient elements (K, Ca, Mg, P, S, Cl, Fe, Mn, B, Zn) except C and N are derived from mineral soils. As minerals become available to biological processes through weathering processes, they are retained by and cycled through these biological processes (Heal and Harrison, 1990). Carbon and nitrogen generally enter biological processes from

445 the atmosphere via photosynthesis and nitrogen fixation respectively. Nutrients that are
446 limited in supply in an undisturbed ecosystem are tightly cycled, with inputs and losses
447 generally being small compared to the pool of nutrients within the ecosystem (Heal and
448 Harrison, 1990). Nitrogen is often the element most in demand due to its central
449 importance in the composition of proteins and nucleic acids (Barbour *et al.*, 1987), and
450 often the most limited in supply (Barbour *et al.*, 1987) due to the complexities of the
451 nitrogen cycle described below.

452 The reactive and diverse chemistry of nitrogen, which can take seven oxidative states and
453 has numerous mechanisms to convert between chemical species, makes the nitrogen cycle
454 one of the most complex elemental cycles (Galloway *et al.*, 2004). The largest pool of
455 nitrogen is the atmosphere where dinitrogen (N_2) gas makes up approximately 78% of the
456 composition of air. This atmospheric nitrogen pool is largely unavailable to most biological
457 entities (Galloway *et al.*, 2004). Nitrogen enters the biological system from this atmospheric
458 pool through one of two pathways (Cleveland *et al.*, 1999). It can be abiotically fixed where
459 N_2 gas is converted to various nitrogen oxides such as NO, NO_2 or N_2O by lightning during
460 storms (Liaw *et al.*, 1990, Bhetanabhotla *et al.*, 1985) and then enters biological pathways
461 through wet or dry deposition on to soils and vegetation (Burns, 2004, Morecroft *et al.*,
462 1994). The second path is biological fixation which occurs when atmospheric N_2 gas is fixed
463 into NH_3^+ and then assimilated into organic nitrogen by a limited set of bacteria and Archaea
464 which are able to break the triple bond between atmospheric N_2 molecules (Houlton, 2008,
465 Galloway *et al.*, 2004). It is generally considered that biological nitrogen fixation exceeds
466 natural abiotic fixation by an order of magnitude (Cleveland *et al.*, 1999). Nitrogen fixation
467 either assimilates the nitrogen directly into living soil microbial nitrogen pools, or into plant
468 organic nitrogen pools if the biological nitrogen fixer is in a symbiotic relationship with a
469 leguminous plant (Houlton, 2008). Organic plant and microbial nitrogen returns to the
470 environment either through herbivory and excretion, or mortality and
471 decomposition/mineralization processes (Moore *et al.*, 2005).

472 Decomposition is the process of breaking down complex organic molecules into simpler
473 molecules, and mineralization is the process of converting the elements in the organic
474 molecules into inorganic form (White, 2006). Decomposition and mineralization in and on
475 the soil may follow one of two possible pathways: (1) a rapid bacterial pathway, and (2) a

476 slower fungal pathway (Moore *et al.*, 2005). During these processes, organic matter is either
477 assimilated into microbial tissue as organic nitrogen, or it is broken down to inorganic
478 nitrogen in a process known as ammonification which results in the formation of ammonium
479 (NH_4^+) (Mooshammer *et al.*, 2014, Vitousek *et al.*, 1979). The ammonium may then be
480 converted to nitrite (NO_2^-) through oxidation by certain autotrophic bacteria and Archaea
481 and a variety of heterotrophic bacteria and fungi (Prosser and Nicol, 2008, Hooper *et al.*,
482 1997) and then further oxidation of the nitrite by different bacteria produces nitrate (NO_3^-)
483 (Klotz and Stein, 2008, De Boer and Kowalchuk, 2001, Vitousek *et al.*, 1979). In the
484 predominant paradigm of the last century, plants, through absorption by their roots, could
485 only access the inorganic components (NH_4^+ and NO_3^-) of the nitrogen cycle when they were
486 freely available in the soil (Jalali *et al.*, 2014, Woodmansee *et al.*, 1978, Haines, 1977,
487 Wiltshire, 1973). However, it is now known that plants can absorb and compete against soil
488 microbes for the monomer forms of organic nitrogen such as peptides and amino acids, as
489 well as the inorganic forms such as NH_4^+ and NO_3^- (Schimel and Bennett, 2004).

490 At each step from organic nitrogen to nitrate, nitrogen can be lost to the atmosphere in a
491 gaseous form. Ammonia (NH_3), the first product of the mineralization of organic material,
492 can be lost from the soil to the atmosphere as a gas through volatilization unless it is
493 assimilated by microbes and plants, or kept in solution in the form of ammonium (NH_4^+).
494 Under anaerobic conditions the nitrite can be converted to nitrous oxide (N_2O), instead of
495 nitrate (Poth and Focht, 1985), or N_2 gas (Poth, 1986), both of which are gaseous and will be
496 lost from the nutrient pools in the soil and enter the atmosphere. Some microbes can use
497 nitrate as an oxygen source under anaerobic conditions and release gaseous nitrous oxide
498 and N_2 gas as a waste (Heinen, 2006, Zumft, 1997). These losses of nitrite and nitrate from
499 soil pools to atmospheric pools as nitrous oxides and dinitrogen gas are collectively referred
500 to as denitrification. Nitrate is highly soluble and together with ammonium can be leached
501 to deeper positions in the soil profile where it may not be available for biological functioning
502 (Pannkuk *et al.*, 2011, Riaz *et al.*, 2011, Mary *et al.*, 1999).

503 Apart from being lost from soil pools through denitrification, leaching and volatilization,
504 nitrogen can also be made unavailable over the long or short term through various kinds of
505 immobilization. When nitrogen is assimilated into microbial or plant tissue, it becomes
506 unavailable to other organisms and so is said to be immobilized (Hoeft *et al.*, 2014,

507 Bengtsson *et al.*, 2003, Burger and Jackson, 2003, Jaeger *et al.*, 1999). The ratio of
508 biologically available carbon to nitrogen is important in determining whether nitrogen is
509 mineralized or microbially immobilized. If the C:N ratio is high, N is likely to be incorporated
510 into microbial tissue and therefore immobilized; whereas if the ratio is low, N is likely to be
511 mineralized and become available for plant uptake (Mooshammer *et al.*, 2014, Holland and
512 Detling, 1990). SOM fractions of varying resistance to mineralization can also immobilize
513 nitrogen over the short term (Barrett and Burke, 2000, Melillo *et al.*, 1989). Long-term
514 nitrogen immobilization can occur through clay-mineral (Ladd *et al.*, 1977, Allison, 1973) and
515 humic interactions (Prescott *et al.*, 2000, Kelley and Stevenson, 1996, Melillo *et al.*, 1989).

516 Terrestrial nutrient cycles are complex with a number of components. Plants, which are
517 responsible for a large degree of terrestrial productivity, take up nutrients mostly from soil
518 nutrient pools, and return them to soil nutrient pools in organic form through exudates and
519 senesced plant parts (Barbour *et al.*, 1987). A variety of soil microbes access energy and/or
520 nutrients by mineralizing this organic material provided by plants or other sources and make
521 elements available in soluble inorganic forms for plant uptake (White, 2006). As this litter
522 breaks down it becomes part of the soil organic matter (SOM) component which, to varying
523 degrees, may be resistant to further mineralization processes (White, 2006). Apart from
524 nitrogen fixation, the critical step in the nitrogen cycle is the mineralization step as plants
525 can only take up nitrogen as inorganic ammonium or nitrate ions, or in simple organic form
526 such as amino acids or peptides (Schimel and Bennett, 2004). Mineralization, or more
527 specifically ammonification, the conversion of large organic molecules containing nitrogen
528 to NH_3 , is thought to be the rate-limiting step (De Boer and Kowalchuk, 2001) in the
529 mineralization process and is dependent, amongst other factors, on microbial action.
530 Measurements of carbon in soils have shown that microbial carbon reacts more quickly to
531 changes in conditions than does total soil carbon and therefore the ratio of microbial carbon
532 (C_{mic}) to total organic carbon (C_{org}) is an indicator of microbial activity and mineralization
533 processes (Anderson and Domsch, 1989). Many different indicators related to soil quality
534 (Bastida *et al.*, 2008, Anderson, 2003), nutrient cycling (Sparling, 1997) and nitrogen
535 mineralization (Schomberg *et al.*, 2009) have been proposed as indicators of a soil's ability
536 to provide nutrients amongst other characteristics. In this study the ratio of microbial
537 nitrogen (N_{mic}) to inorganic nitrogen (N_{ino}) was used as an indicator of nutrient cycling (D.

538 Drake, pers. comm., 2012) and will be referred to as the nutrient cycling ratio index. The LFA
539 nutrient cycling index is calculated by the sum of values obtained from ranking scales for
540 specific soil surface assessment indicators (SSAI). These SSAIs include perennial basal grass
541 and shrub and tree cover, litter cover and its origin and degree of decomposition, biological
542 soil crust cover, and surface roughness.

543 From the above discussion, it has been shown that a number of ecological parameters can
544 be measured which provide an indication of a landscape's functioning (Tongway and
545 Hindley, 2004). A landscape can be portioned into patches or units based on observed
546 changes in biogeochemical functioning. Within these patches, critical ecological processes
547 can be quantified. Soil stability identifies a landscapes susceptibility to erosion, and
548 therefore degradation processes. Infiltration identifies the potential of a soil to absorb
549 water. This is important as soil water is critical to nutrient cycling and plant growth; and
550 runoff influences erosion and soil stability. Finally, by measuring the size of microbial pools,
551 it is possible to obtain an indication of nutrient cycling in these landscapes.

552 2.2 Methods and Materials

553 2.2.1 Research Sites, Climate, Vegetation, Geology and Soils

554 Field research was conducted at two AngloGold Ashanti (AGA) Mining Regions: Vaal River
555 (VR) and West Wits (WW) Mining Regions (Figure 2.2, 2.3 and 2.4, respectively). These two
556 regions are 85 km apart, to the southwest of Johannesburg near Carletonville (WW) in
557 Gauteng Province, and Klerksdorp (VR) in Northwest Province, South Africa. Both mining
558 regions produce gold and uranium from deep-level gold mines (to a depth of 4 km below
559 surface). Sites at VR are situated between 1280 and 1330 metres above sea-level (masl)
560 while those at WW are between 1570 and 1630 masl. Both sites receive summer rainfall,
561 with the southern hemisphere winter from about May to September being the dry, cold
562 period. Mean annual precipitation at WW is 661 mm and VR is 565 mm with mean annual
563 potential evaporation for WW at 2209 mm and VR at 2407 mm (Mucina and Rutherford,
564 2006). The mean annual temperature is 15.6°C for WW and 16.8°C for VR with a mean daily
565 minimum of 0 and 2°C and a mean daily maximum of 25 and 27.5°C for WW and VR
566 respectively (Weiersbye *et al.*, 2006). Frost is common in both sites with mean frost days per
567 annum of 33 and 34 days for WW and VR respectively.

568 At VR all transects were in the Vaal Reefs Dolomite Sinkhole Woodland vegetation type (Gh
569 12) which is classified under the Dry Highveld Grassland Bioregion within the grassland
570 biome as classified by Mucina and Rutherford (2006). Acocks (1988) characterised these
571 grasslands as transitional between the western variation of Bankenveld (61a), a southern
572 variation of *Cymbopogon-Themedra* veld (48a), and a northern variation of Dry *Cymbopogon-*
573 *Themeda* veld (50a). Important taxa include small trees such as *Acacia erioloba* E.Mey.,
574 *Celtis Africana* Burm.f. and *Rhus lancea* L.f., and graminoids such as *Aristida congesta* Roem
575 & Schult, *Digitaria eriantha* Steudel, *Eragrostis curvula* (Schrad.) Nees and *Themeda triandra*
576 Forsk. amongst others (Mucina and Rutherford, 2006).

577 Transects at WW were located in the Gauteng Shale Mountain Bushveld (SVcb 10)
578 vegetation type belonging to the Central Bushveld Bioregion of the Savanna Biome (Mucina
579 and Rutherford, 2006). Acocks (1988) defined the vegetation as Bankenveld (61), possibly a
580 transition between the western and central variations, whereas O'Connor and Bredenkamp

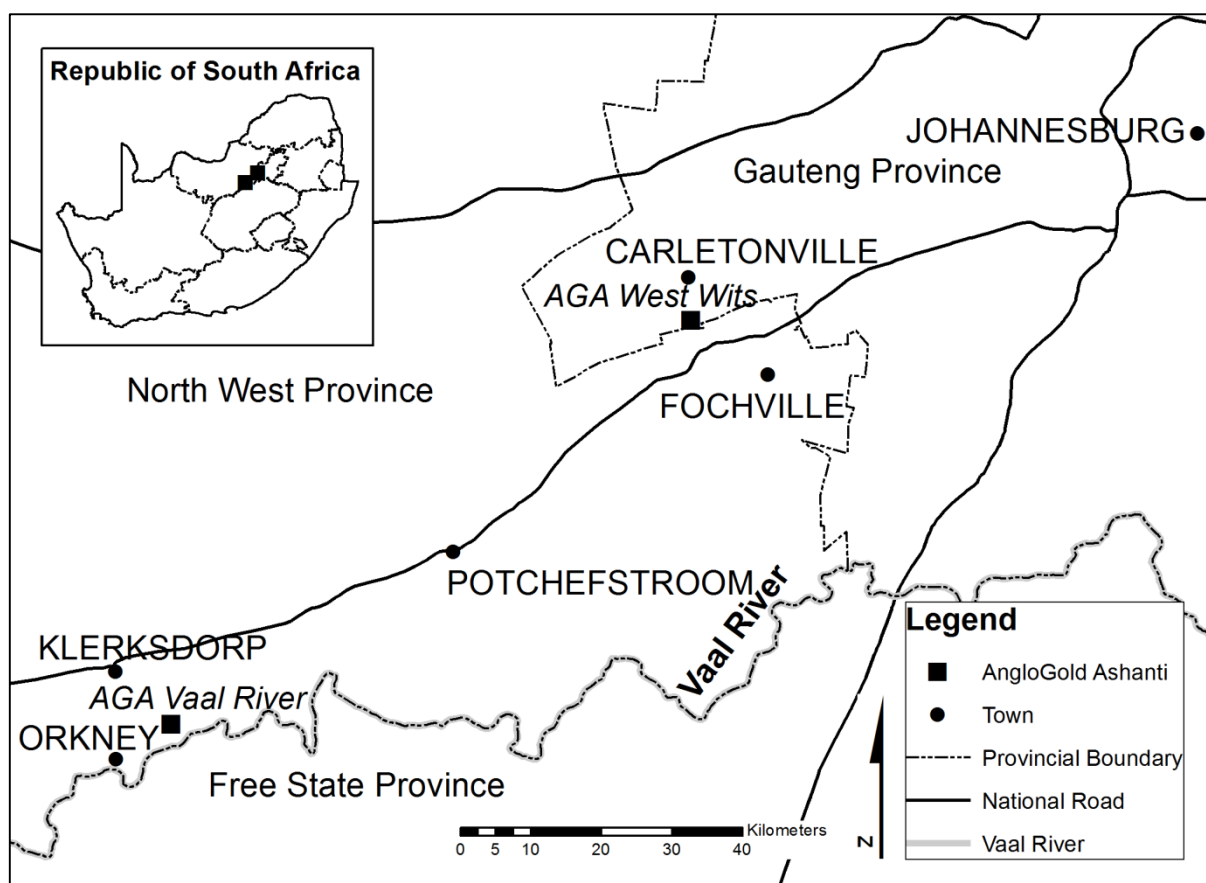


Figure 2.2 Map showing the two mining regions relative to Johannesburg and the provinces of Gauteng, North West and Free State. The insert shows the position of the sites relative to the provinces of South Africa.

(1997) defined the vegetation as Central Plateau Grasslands, being the A2 subdivision, which is dominated by *Panicum coloratum* and *Eragrostis curvula*. All transects at VR were located on dolomites, with subordinate chert, of the Malmani Subgroup (Chuniespoort Group, Transvaal Supergroup) where soils are shallow (50 – 150 mm) and rocky, and dominated by Mispah, Glenrosa and Hutton soil forms. West Wits transects are located mainly on shale and quartzite from the Timeball Shales and Quartzite (Pretoria Group, Transvaal Supergroup) and may stray onto dolomites of the Malmani Subgroup (Chuniespoort Group, Transvaal Supergroup). Soils at WW are mainly shallow Mispah soils (Mucina and Rutherford, 2006). The most common land uses in both areas are cattle grazing (rangelands) and crop agriculture (maize and sunflower). Landscapes around these gold mines have degraded and derelict land from both mining activities and other human activities with varying degrees of disturbed vegetation cover, and swampy lands influenced by seepage from tailings dams and altered drainage patterns (Weiersbye *et al.*, 2006).



Figure 2.3

Vaal River Mining Region showing the distribution of transects across a portion of the mining region. The missing transects (4, 5, 6 and 12) were burnt before they could be sampled, while sampling of transect 15 was abruptly abandoned due to a fire arriving.

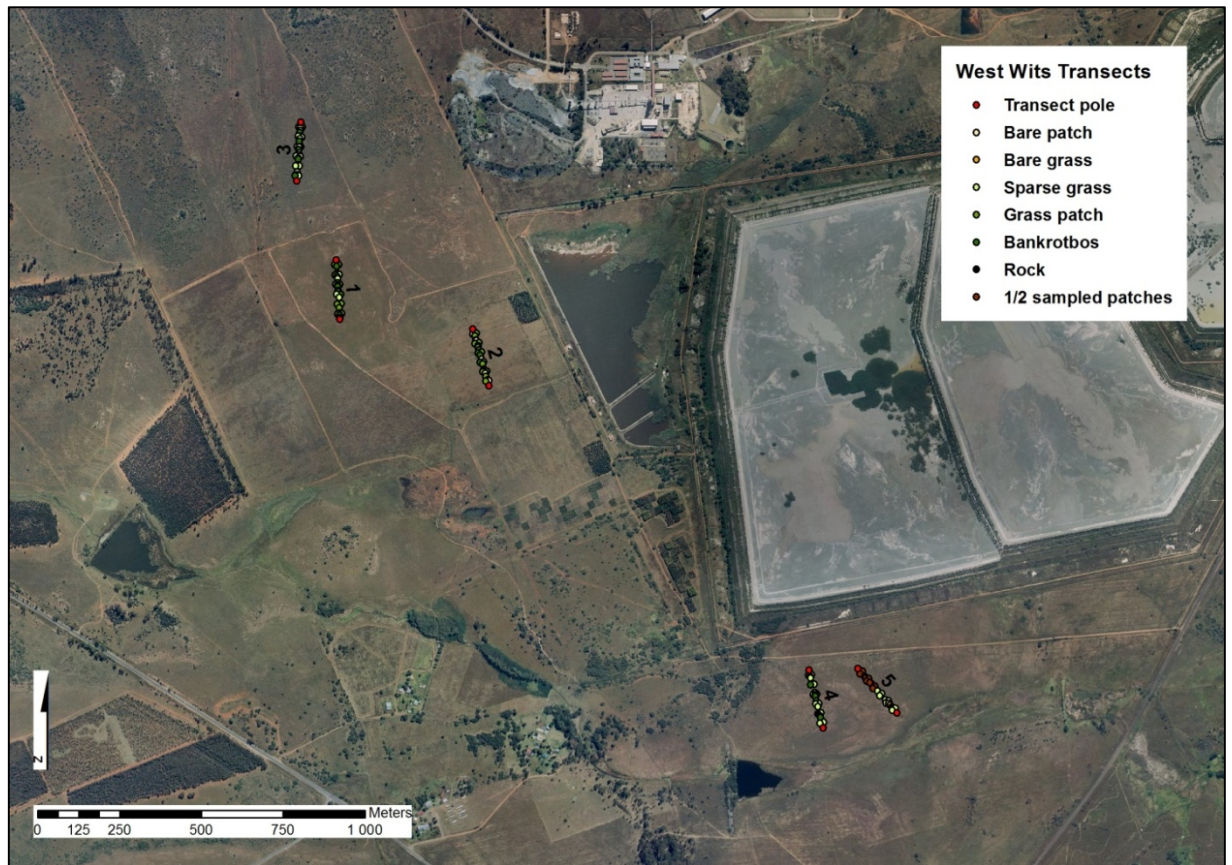


Figure 2.4 The position of transects and patches on the upper Varkenslaagte area in the West Wits Mining Region. The ½ sampled patches in transect 5 were lost overnight as a result of fire.

2.2.2 Landscape Function Analysis Transects

In LFA theory, transects are established following the gradient (sometimes referred to as a gradsect) and the length of each transect is based on the patch size and the need to sample five replicates of each identified patch type (Tongway and Hindley, 2004). As per theory, transects were established following the gradient where this could be clearly discerned. However as this research was designed to model and predict LFA indices from Hyperion hyperspectral imagery, the length of a transect was based on a Hyperion pixel size of 30 m. Furthermore, it was not possible *a priori* to predict the position of a pixel, so to reduce excessive redundant sampling, transects were designed to be 180 m long. Using this length of transect meant five pixels could be accommodated on the ground with only 30 m of transect being redundant in the hyperspectral image regardless of the position of the pixels on the transect.

623 However such a length of transect and associated successive pixels raised a number of other
624 issues. First, this length made it difficult to find areas of grassland that were not bisected by
625 mine infrastructure like roads, pipelines or fences or mining activity like surface scrapping or
626 digging. Thus transects were sited in areas that avoided such infrastructure and mine-
627 related physical disturbed areas. Areas were also avoided that had excessive woodiness that
628 might impinge on the area around a transect and therefore influence either the grassland or
629 the pixel spectra. However, there were no specific guidelines for selecting an area other
630 than it being a natural grassland not influenced by any of the above, and the final location of
631 transects was essentially random in suitable areas. At Vaal River, a largescale guideline was
632 to follow a drainage line that started at the Black Reef ridge and ran between two large
633 tailings storage facilities and onto the Vaal River flood plain so that the influence of AMD
634 might be included in the sampling area. At West Wits, only a small area of grassland was
635 available for sampling and it was difficult to find areas that would fit a 180 m transect.
636 Second, was the influence of spatial autocorrelation. It was deemed that as the purpose of
637 the study was to model and predict LFA indices from hyperspectral data, autocorrelation
638 was not regarded as a major confounding variable in model testing. Third, under the
639 standard LFA protocol the 180 m transect length raised issues around the definition of a
640 patch. Patches in LFA are defined by a change in structure and their boundaries are often
641 indicated by the beginning or end of a deposition zone (Tongway and Hindley, 2004). In the
642 grasslands at VR and WW these were often defined by the size of a grass tuft and would be
643 measured at the decimetre scale which was not feasible over 180 m. As these grass tufts
644 essentially formed a matrix in which other patch types occurred, it was decided to define
645 the grass matrix into areas of visibly homogenous cover with the boundary being defined by
646 a change in physiognomic structure. Thus the dominant patch type was grassland forming a
647 homogenous matrix along the length of each transect. Within this grassland matrix,
648 deviations from this matrix vegetation type indicated subsidiary patches whose type was
649 defined according to physiognomy (Table 2.1).

650 2.2.3 **Measurement and Sampling of Patch Types**

651 During an initial traverse of each transect, patch types were identified (Table 2.1 and Figures
652 2.5, 2.6, 2.7 and 2.8) and their width and length along the entire transect quantified.
653 Thereafter, for each transect, five patches in each patch type were identified for further

654 measurements. These measurements included the eleven Soil Surface Assessment
655 Indicators (SSAI) as defined in the LFA manual (Tongway and Hindley, 2004), a vegetation
656 sample if applicable, and a soil sample. The vegetation sample was collected using a 25 x 25
657 cm quadrat (0.0625 m²). For grasses, all vegetation rooted within the quadrat was clipped at
658 ground level and stored in a packet. For shrubs, all vegetation within a vertical square
659 column of the quadrat was clipped and kept regardless of whether it was rooted in or out
660 the quadrat. A soil core was collected from within the quadrat used to sample above-ground
661 vegetation by driving a stainless steel tube into the soil to a depth of 10 cm if possible. The
662 core with soil was carefully extracted from the soil and stored in a ziplock packet in a cool
663 box in the field and transferred to a fridge/coldroom at the end of the day. Where large
664 roots or stone prevented penetration of the core to 10 cm, the depth at which the core
665 stopped was recorded. At the same time that the core sample was collected, a bulk soil
666 sample was also collected and stored in the same manner.

667 **Table 2.1** Names, codes and definitions of patch types described on transects at West
668 Wits and Vaal River sites.

Patch type	Patch definition
Bare patch (bp)	Exposed soil with no vegetation or biological soil crusts
Biological soil crust patch (bsc)	Soil surface predominantly occupied by biological soil crusts but with no vegetation present
Bare grass patch (bg)	The soil surface with or without biological soil crust predominates and there is a sparse presence of vegetation, generally grasses, which constitutes a minor portion of the cover.
Sparse grass patch (sg)	Grass dominates the cover, but soil surface with or without biological soil crust is visible through the canopy.
Grass patch (gp)	Grass dominates the cover and no soil surface is visible through the canopy
Tall grass patch (tg)	Grasses with extended culms forming extended columnar tufts – most commonly <i>Hyparrhinea hirta</i>
<i>Schoenoplectus</i> patch (sch)	Patches where cover is predominantly a species of <i>Schoenoplectus</i> sedge.
<i>Seriphium plumosum</i> or Bankrotbos patch (bb)	Patches where cover was dominated by the shrub <i>Seriphium plumosum</i>

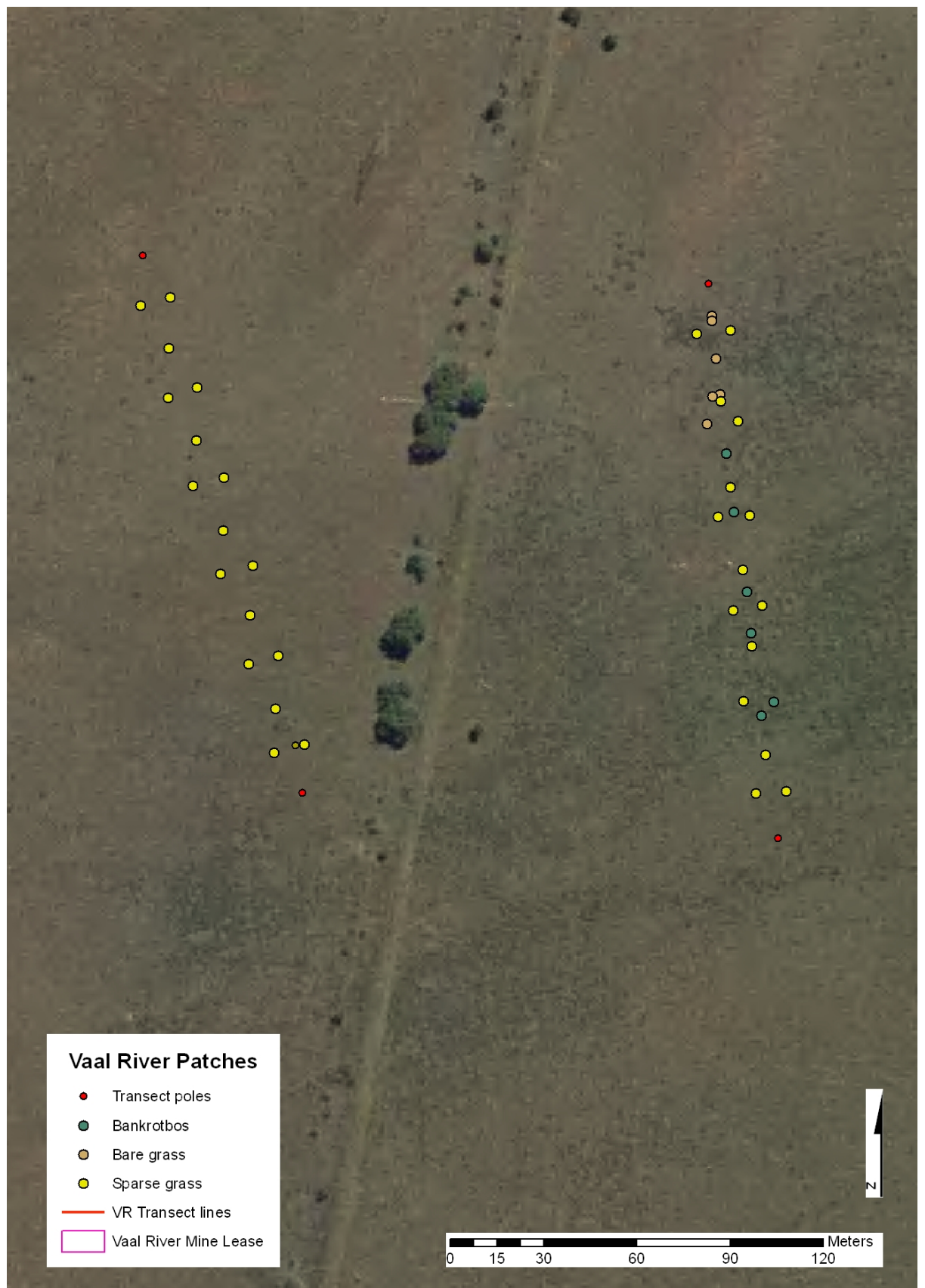
669

670

671 2.2.4 Vegetation and Soil Measurements

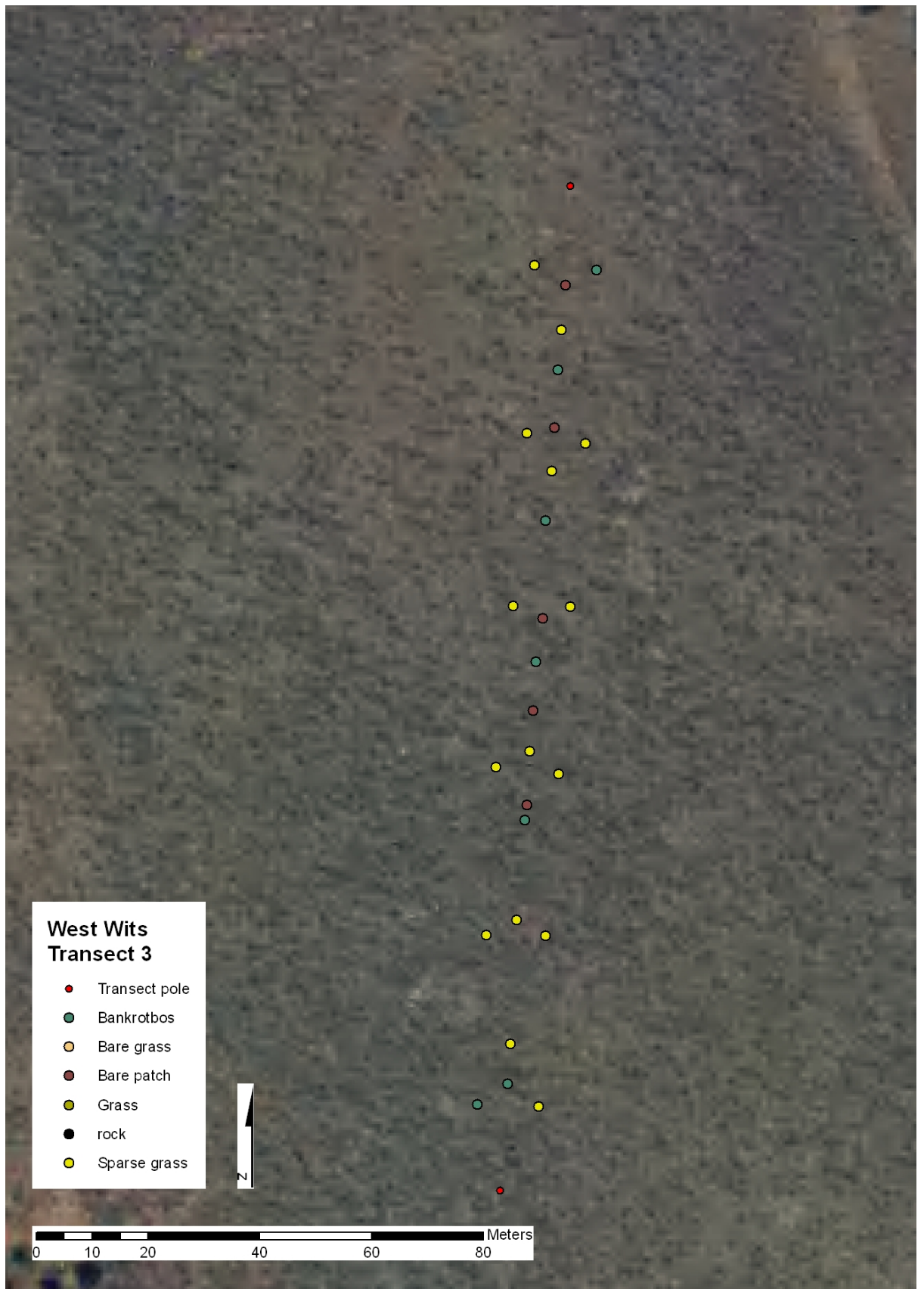
672 Above-ground vegetation samples included all standing biomass within the sample quadrat
673 regardless of whether living or senesced, but did not include litter lying on the soil surface.
674 Litter was rare with the exception of *S. plumosum* patches which often had a fine layer of
675 presumably old seed husks or bracts covering the soil under the canopy. These above-
676 ground vegetation samples were weighed fresh and then dried in an oven at 50°C until
677 there was no change in dry weight and are referred to in this study as above-ground
678 biomass. The cored soil samples were weighed fresh and sieved through a 2 mm sieve.
679 Stones, and organic matter such as roots, were sorted into their separate fractions and
680 weighed fresh and dried. The organic matter large enough to be removed from the soil
681 samples is referred to as roots or below-ground biomass in this study. The sieved core
682 sample was then mixed with the sieved bulk sample for each sample point but there was no
683 mixing between sample points. Electrical conductivity and pH were measured on a 2:1
684 water:soil by mass slurry with hand-held metres (Benton Jones Jr., 2001). Soil moisture was
685 determined by drying a known weight of sieved soil in a crucible at 100°C for 48 hours
686 (Wilke, 2005, Jarrell *et al.*, 1999). Thereafter the same sample was used to determine
687 organic carbon content gravimetrically through loss-on-ignition by heating the sample at
688 550°C for 3 hours (Wilke, 2005, Sollins *et al.*, 1999).

689 An indication of nutrient cycling was obtained for the WW samples by determining the ratio
690 of organic to extractable inorganic nitrogen in each sample (D. Drake, pers. comm., 2012).
691 This index is referred to in this study as the nutrient cycling ratio index (NCR) to separate it
692 from the LFA nutrient cycling index. Extractable inorganic nitrogen was extracted by shaking
693 a known weight of soil (~30 g) in 100 ml of 2 M KCl solution overnight. This was filtered and
694 a 1 ml subsample of the filtrate run through a cadmium column to reduce the nitrate to
695 nitrite before diazotization and nitrite quantification with a spectrophotometer at 540 nm
696 (Mulvaney, 1996). From the same filtrate the ammonium was quantified by treatment of a 2
697 ml sub-sample with the salicylate-hypochlorite colorimetric method and quantification with
698 a spectrophotometer (Mulvaney, 1996).



699

700 **Figure 2.5** Transect 10 and 11 at Vaal River



701

702 **Figure 2.6** Transect 3 at West Wits

703

(a)



(b)



704

(c)



(d)



705

(e)



(f)



706

(g)



(h)



707

Figure 2.7

708

709

Bird's eye view of patch types described at West Wits: (a) grass patch, (b) and (c) sparse grass, (d) and (e) *Seripheum plumosum* patches, (f) bare grass, (g) and (h) bare patch.

710

(a)



(b)



(c)

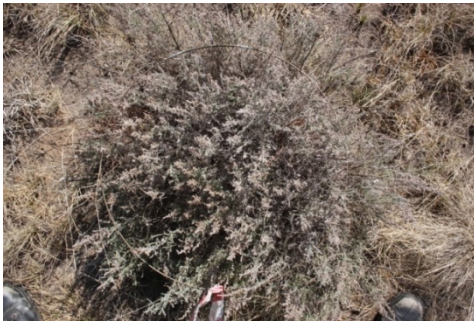


711

(d)



(e)



(f)



712

(g)



(h)



(i)



713

714

Figure 2.8

Bird's eye view of patch types described at Vaal River: (a) and (b) sparse grass, (c) tall grass, (d) *Schoenoplectus* patch, (e) and (f) *Seripheum plumosum* patch, (g) biological soil crust patch, (h) and (i) bare grass patch.

715 For soil organic and microbial nitrogen, a second ~30 g soil sample was incubated in
716 chloroform for 48 hours and then passively aerated for 24 hours (Paul *et al.*, 1999), before
717 following the same extraction as described above for inorganic nitrogen. The filtrate was
718 subjected to a persulphate treatment to oxidize organic nitrogen to nitrate (D'elia *et al.*,
719 1977) before analysis for nitrate by running the sample through a cadmium column to
720 convert nitrate to nitrite followed by colorimetric determination as described above for
721 inorganic nitrate (Mulvaney, 1996).

722 2.2.5 **Statistical Analysis**

723 All statistical analyses were performed using the R statistical package (R Core Team, 2013).
724 Data for most variables were found not to be normally distributed using quantile-quantile
725 plots and the Shapiro-Wilk Normality Test (Royston, 1982). As a result the non-parametric
726 Kruskal-Wallis test that comes with the base R package was used to test if transects differed
727 from each other, and likewise if patch types differed from each other. If significant
728 differences were detected between transects or between patch types with the Kruskal
729 Wallis test, the Multiple Comparisons Test in the pgirmess package (Bouwman *et al.*, 2010)
730 was used to identify which of these transects or patch types differed from the others.
731 Although spatial autocorrelation was not a critical factor in this study, Moran's I
732 autocorrelation Index was calculated using the Moran.I function from the APE package
733 (Kalwij *et al.*, 2010) to better understand the data. To explore how variables such as above-
734 ground biomass, pH, soil moisture or extractable inorganic nitrogen may relate to each
735 other, correlations and the plotting of correlograms was performed with the psych package
736 (Kelly and Walker, 1976). Finally a principle components analysis using the prcomp function
737 that comes with the base stats package of R (R Core Team, 2013) was used to explore the
738 LFA, vegetation and soil data for underlying factors and groupings that may indicate
739 structure in the data.

740 2.3 **Results**

741 2.3.1 **Patch Types**

742 Each transect was dominated by a particular patch type which most commonly was sparse
743 grass, but at West Wits this may have been a grass patch (Table 2.1 and 2.2). This dominant
744 patch type essentially formed a matrix or “sea” in which the other patch types occurred as
745 “islands” of varying density and spatial configuration. Four patch types were described at
746 West Wits while seven were described at Vaal River. The patch types described here were
747 those that a transect line crossed. Not all patch types described were found on each
748 transect. Nor were the patch types described from the sampled transects a complete
749 description of those that may be found in the two study areas.

750 A grass patch (gp) was defined when the canopy of grass tufts was dense enough that the
751 soil surface was not visible (Table 2.1). When the grass canopy was thin enough that a soil
752 surface was visible but the cover was still dominated by grass, the patch was called a sparse
753 grass patch (sg). Bare patches (bp) occurred when an area of soil surface at least 42 cm in
754 diameter (the field-of-view for the ground-based hyperspectral sensor at approximately 1 m
755 above the ground) was exposed with minimal evidence of the presence of biological soil
756 crust and no vegetation cover. Bare patches were not that common as many of these
757 exposed soil surfaces had biological soil crusts dominating the surface cover and were then
758 named biological soil crust patches (bsc). Anecdotal evidence suggested the presence or
759 absence of cattle influenced the balance between biologically crusted surfaces and exposed
760 soil surfaces. Tall grass patches (tg) were also present and were defined by tufted grass with
761 dense extended culms forming columnar canopies approximately a metre or more high. On
762 one transect extending into hydromorphic grassland situated between two tailings storage
763 facilities, a species of sedge, assumed to be *Schoenoplectus corymbosus* (Roth ex Roem. &
764 Schult.) J. Raynal, formed dense, extended patches called *Schoenoplectus* patches (sch).
765 Finally a patch type common on many but not all transects at both sites was called
766 bankrotbos (bb). Bankrotbos is the Afrikaans common name for a shrub recently renamed
767 *Seriphium plumosum* L. and previously known as *Stoebe vulgaris* (Table 2.1).

768 **Table 2.2** The total number of each patch type recorded on transects at each mine site.

Patch type	Number at VR	Number at WW
Bare patch (bp)	5	73
Biological soil crust (bsc)	140	0
Bare grass (bg)	186	3
Sparse grass (sg)	439	85
Grass patch (gp)	21	98
Tall grass (tg)	8	0
Schoenoplectus patch (sch)	3	0
Bankrotbos (bb)	107	106

769

770 2.3.2 Transects and Transect Indices

771 Landscape Function Analysis allows the calculation of a number of indices of landscape
 772 function and organization based on parameters measured along the transect (Table 2.3).
 773 The assumption behind these indices is that a change in resource use should result in a
 774 change in landscape organisation which can be determined by comparing with a reference
 775 site or by comparison to the same site at a previous time. The landscape organization index
 776 (LOI) is defined as the total length of patches divided by the length of transect, while the
 777 Patch Area Index (PAI) is the total patch area divided by the total transect area. To this end a
 778 value of 1 for LOI or PAI suggests that a landscape has maximized the spatial arrangement of
 779 patches and there are no bare patches in the landscape. Thus as a first-order indicator, the
 780 higher the value for these two indices, the healthier the landscape is likely to be. However,
 781 as will be explored further in the discussion, values of 1 are more likely an artefact of the
 782 smoothing of grass tufts into homogenous units and defining bare patches as of a minimum
 783 size corresponding to the field-of-view of the ground-based hyperspectral sensor.

784 The five transects at West Wits show a greater range for LOI between their five transects
 785 (0.83 – 1.00) than did the eleven transects at Vaal River (0.92 – 1.00) suggesting that some
 786 transects at West Wits are functioning at a less efficient level than the Vaal River transects.
 787 At West Wits, Transect 2 recorded the lowest value for both LOI (0.83) and PAI (0.52). Vaal
 788 River showed a much more complex picture with a number of transects (3, 8, 13 and 14) all
 789 having the lowest LOI value (0.92), but a different set of transects (8, 9, 15) having the
 790 lowest value for PAI (0.50, 0.55, 0.56 respectively). For the measure of the number of

791 patches per 10 m, the transects (9, 8 and 15) with the maximum number of patches (6.47,
792 6.05 and 4.60 respectively) corresponded with those for the lowest PAI (Table 2.2).

793 Percentage indices for stability, infiltration and nutrient cycling calculated for each transect,
794 where lower values indicate lower functionality for that specific index, showed a different
795 pattern to those for LAI and PAI above (Table 2.2) in that West Wits tended to show higher
796 or similar values to those at Vaal River. Stability values for all transects were very similar for
797 the Vaal River transects where transect 7 had the lowest value (61.7%) and transect 8
798 (69.5%) the highest, whereas the stability index showed more variability over the fewer
799 West Wits transects where transect 3 had the lowest value (61.9%) and transect 1 had the
800 highest value (73.2%). In contrast, transect 1 at Vaal River had the lowest value (29.0 and
801 19.1%) for infiltration and nutrient cycling respectively, while transect 8 had the highest
802 infiltration value (35.1%) and transect 10 had the highest nutrient cycling index (28.1%). At
803 West Wits, transect 4 had the lowest values (27.2 and 22.0%) for both infiltration and
804 nutrient cycling respectively, and transect 1 had the highest values (38.5 and 35.5%
805 respectively) for these two indices (Table 2.2).

806 Correlations between the various LFA transect indices (Figure 2.9) are generally fairly strong.
807 The exceptions are the LOI having almost no correlation with the three LFA indices of
808 stability, infiltration and nutrient cycling; and inter-patch length having no correlation with
809 PAI, patches per 10 m, and patch width. However PAI had a strong correlation ($r = 0.80$) with
810 patch width and a less strong correlation with LOI ($r = 0.52$), and a negative correlation with
811 the LFA indices of stability ($r = -0.52$), infiltration ($r = -0.72$) and nutrient cycling ($r = -0.57$).
812 The LFA indices of stability, infiltration and nutrient cycling are quite strongly correlated
813 with each other and less well correlated with patches per 10 m ($r = 0.34$, 0.62 , and 0.42
814 respectively), but are negatively correlated with PAI, patch width and inter-patch length
815 (Figure 2.9).

816 **Table 2.3** A summary of indices and measures calculated from the LFA data for each transect. Transect vr15 is missing individual LFA patch
817 values as it was burnt before these could be collected.

Transect	vr1	vr2	vr3	vr7	vr8	vr9	vr10	vr11	vr13	vr14	vr15	ww1	ww2	ww3	ww4	ww5
Landscape Organisation Index	0.96	0.93	0.92	0.98	0.92	0.97	1.00	1.00	0.92	0.92	0.89	1.00	0.83	0.89	0.99	1.00
Patch Area Index	0.89	0.86	0.78	0.77	0.50	0.55	0.75	1.00	0.77	0.67	0.56	0.63	0.52	0.72	0.94	0.94
Number Patches/10m transect	1.91	2.36	2.91	4.21	6.05	6.47	3.33	0.17	3.31	4.17	4.60	5.03	3.91	3.94	1.78	1.67
Mean patch width (cm)	616	715	521	468	284	280	360	1000	562	355	429	364	437	419	565	534
Average Interpatch Length (m)	0.89	0.68	0.56	0.86	0.63	0.33	0.00	0.00	0.57	0.49	0.44	0.00	0.64	0.79	0.95	0.00
Min Interpatch length (m)	0.10	0.10	0.10	0.20	0.20	0.10	0.00	0.40	0.10	0.05	0.05	0.40	0.20	0.10	0.95	0.00
Max Interpatch length (m)	1.60	2.25	1.40	1.40	1.30	1.60	0.00	1.00	1.80	2.00	1.35	0.50	3.60	2.70	0.95	0.00
LFA stability index (%)	64.4	62.1	64.8	61.7	69.5	66.2	68.5	67.4	63.8	66.4		73.2	70.4	61.9	65.8	62.0
LFA infiltration index (%)	29.0	30.1	32.6	32.0	35.1	30.7	34.1	29.8	30.3	31.3		38.5	34.6	31.6	27.2	27.8
LFA nutrient cycling index (%)	19.1	21.9	23.6	21.3	27.7	25.3	28.1	26.2	23.2	24.8		35.5	30.7	23.5	22.0	23.0
Slope (rise/run)	0.02	0.006	0.01	0.006	0.006	0.006	0.01	0.02	0.006	0.006	0.01	0.04	0.02	0.05	0.03	0.03
Aspect	e	se	se	s	e	e	s	s	se	se	se	s	sw	s	s	se
# exposed regolith	8	20	5	12	3	16	2	1	8	10	8		1		1	
# <i>S. plumosum</i> patches				18	43	23	23					32	15	33	15	11
# bare grass patches	5	1	6	9		13	4		18	21	20					3
# bare grass & evaporite			8			13										
# bare patches									5				47	26		
# bsc patches	8	19	10	5		5			21	27	45					
# bsc patches & evaporite			14		23	9										
# grass patches						21						44	54			
# Schoenoplectus spp. patches					3											
# sparse grass patches	24	40	18	44	62	19	31	3	41	52	62	14		38	17	16
# sparse grass & evaporite			20			23										
# tall grass patches	5			3												

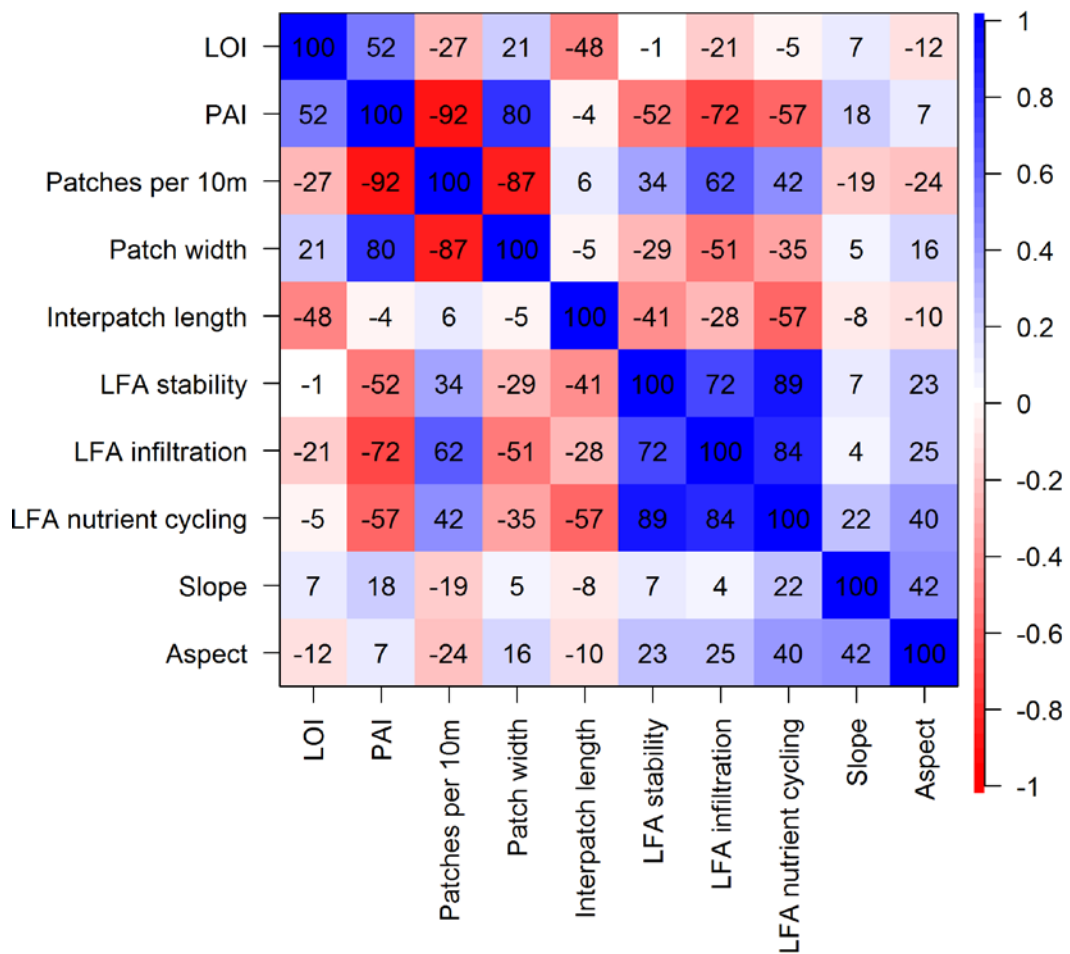


Figure 2.9 A correlogram of summary and index variables for each transect to explore relationships between the various LFA organisation indices and the three indices of function. Vaal River and West Wits transect data were pooled for this analysis.

2.3.3 Patch Measures from Vaal River

Mean values for transects across Vaal River were fairly similar with vr8 having the highest value ($69.6 \pm 0.8\%$) and vr2 having the lowest value ($63.1 \pm 0.7\%$) for stability (Table 2.3). Transect vr8 had the highest value for a number of variables such as infiltration ($35.2 \pm 1.3\%$), nutrient cycling ($27.9 \pm 1.2\%$), fresh and dry above-ground biomass ($1069.9 \pm 164.9 \text{ g.m}^{-2}$ and $837.5 \pm 122.7 \text{ g.m}^{-2}$ respectively), conductivity ($881.5 \pm 178.5 \mu\text{S.cm}^{-1}$), and pH (5.96 ± 0.06) but had the lowest pH (5.96 ± 0.06). The lowest value for these variables were recorded in a mix of transects. Transect vr2 had the lowest stability values ($63.1 \pm 0.7\%$) while transect vr13 had the lowest infiltration ($29.4 \pm 1.0\%$) and transect vr1 had the lowest LFA nutrient cycling index ($21.3 \pm 0.8\%$). Both fresh and dry above-ground biomass was

lowest for transect vr3 (270.6 ± 44.1 and $221.2 \pm 34.5 \text{ g.m}^{-2}$ respectively) although this was only statistically significant when compared with transect vr8 which had the highest biomass (Table 2.3).

Transect vr11 had the lowest conductivity ($23.6 \pm 1.4 \text{ }\mu\text{S.cm}^{-1}$) but had significantly higher values for soil moisture ($5.45 \pm 0.3\%$ by mass) compared to other transects except when compared to vr10 (Table 2.3). However, it should be noted that these were the first transects sampled and that the pattern of soil moisture follows that of sampling into the dry-season winter where the last transects sampled have the driest soils. Transects vr10 and 11 also had no cores shorter than the maximum 10 cm which was unique but not a statistically significant difference. These two transects also had the highest soil bulk density ($1.73 \pm 0.05 \text{ g.cm}^{-3}$ and $1.83 \pm 0.05 \text{ g.cm}^{-3}$ respectively) but this difference was only significant when compared to the two lowest soil bulk densities (1.39 ± 0.04 and $1.41 \pm 0.04 \text{ g.cm}^{-3}$ for transect vr1 and vr8 respectively).

Most transects at VR had essentially neutral soil pH's tending towards the slightly acidic with transect vr1 having the highest pH (6.81 ± 0.06). Soil organic matter or SOM, as determined through loss-on-ignition, was significantly higher in transect vr2 ($5.50 \pm 0.17\%$ by mass) compared to transect vr11 where it was lowest ($2.83 \pm 0.1\%$ by mass). All transects at Vaal River had stone, defined as non-organic particles bigger than 2 mm diameter, present in the soil with transect vr14 having the highest ($2597.9 \pm 566.1 \text{ g.m}^{-2}$) and transect vr11 had the lowest amount of stone in its soil ($216.4 \pm 46.9 \text{ g.m}^{-2}$). Roots in soil, as distinct from SOM through being determined by material left in a 2 mm sieve and separated from the stony material by hand, was highest for transect vr7 ($139.2 \pm 30.7 \text{ g.m}^{-2}$) which was only significantly different from the lowest value ($30.95 \pm 5.4 \text{ g.m}^{-2}$) found for transect vr14 (Table 2.3).

When sorting the data by patch type, generally patches of biological soil crust (bsc) had the lowest or highest mean values depending on the variable (Table 2.4) when compared to other VR patch types. Thus bsc patches had significantly lower values for all three LFA indices ($62.0 \pm 0.4\%$, $24.2 \pm 0.5\%$ and $19.3 \pm 0.4\%$ for stability, infiltration and nutrient cycling respectively) when compared with *S. plumosum* patches which had the highest values for all three indices ($72.7 \pm 0.7\%$, $47.9 \pm 1.5\%$ and $39.9 \pm 1.3\%$, respectively). Shoot

864 biomass was significantly higher in the tall grass patches (2902.2 ± 618.3 and 2506.5 ± 487.9
865 g.m^{-2} for fresh and dry respectively), closely followed by the *Schoenoplectus* patches and *S.*
866 *plumosum* patches when compared to the bare grass patches which had the lowest shoot
867 biomass (175.7 ± 16.0 and $164.7 \pm 13.4 \text{ g.m}^{-2}$, respectively). *Schoenoplectus* patches had the
868 highest conductivity ($511.4 \pm 193.1 \mu\text{S.cm}^{-1}$) and the lowest pH (6.03 ± 0.05), while tall grass
869 patches had the lowest conductivity ($51.0 \pm 5.9 \mu\text{S.cm}^{-1}$) and both bsc and bare grass
870 patches had the highest pH (6.57 ± 0.08 or 0.07 respectively). The lowest mean soil moisture
871 was found in soils from bsc patches ($2.28 \pm 0.19\%$ by mass) while *Schoenoplectus* patches
872 had significantly higher soil moisture ($4.44 \pm 0.32\%$ by mass), as did *S. plumosum* patches
873 ($4.14 \pm 0.28\%$ by mass).

874 No significant differences for SOM values were found between any patches at Vaal River.
875 However when comparing root biomass, *Schoenoplectus* patches had significantly higher
876 mean values (362.4 ± 85.7 and $161.9 \pm 51.6 \text{ g.m}^{-2}$ for fresh and dry biomass respectively), as
877 did *S. plumosum* (200.5 ± 37.1 and $105.2 \pm 21.2 \text{ g.m}^{-2}$, respectively), tall grass (230.6 ± 52.9 and
878 $114.6 \pm 29.1 \text{ g.m}^{-2}$, respectively) and sparse grass patches (185.0 ± 17.2 and $98.4 \pm 11.4 \text{ g.m}^{-2}$
879 respectively) when compared to those of bsc patches, which had the lowest root biomass
880 (26.4 ± 3.2 and $11.6 \pm 1.5 \text{ g.m}^{-2}$, respectively). Both the amount of stone (2585.4 ± 326.3 and
881 $24312.5 \pm 326.3 \text{ g.m}^{-2}$ for fresh and dry stone, respectively) and the soil density (1.75 ± 0.05
882 g.cm^{-3}) was significantly higher in bsc patches as compared to those for patches with the
883 lowest values for these variables. The lowest values for soil stone were found in tall grass
884 patches (696.5 ± 129.8 and $649.3 \pm 120.6 \text{ g.m}^{-2}$ for fresh and dry stone respectively), while
885 the lowest soil density was ($1.2 \pm 0.12 \text{ g.cm}^{-3}$) for *Schoenoplectus* patches.

886 **Table 2.4** Comparisons of transects from Vaal River showing the mean and standard error calculated from measurements collected from
887 individual patches within each transect. Variables with superscripted asterisks are significantly different at the 5% level (*) and
888 1% level (**), values with common superscripted letters are not significantly different, while those with different letters were
889 significantly different at the 5% level using a non-parametric Kruskal-Wallis with Multiple Comparison test.

	vr1	vr2	vr3	vr7	vr8	vr9	vr10	vr11	vr13	vr14
N	32	27	37	37	32	28	27	17	27	27
Stability (%) **	64.3 ± 0.5 ^a	63.1 ± 0.7 ^a	64.9 ± 0.8 ^a	64.0 ± 0.7 ^a	69.6 ± 0.8 ^b	69.5 ± 0.9 ^b	68.3 ± 0.5 ^{ab}	67.8 ± 0.5 ^{ab}	65.0 ± 0.5 ^{ab}	66.5 ± 0.6 ^{ab}
Infiltration (%) *	29.9 ± 0.6 ^a	30.6 ± 0.7 ^a	31.2 ± 0.8 ^a	34.9 ± 2.0 ^a	35.2 ± 1.3 ^a	32.7 ± 1.8 ^a	33.0 ± 1.6 ^a	29.9 ± 1.0 ^a	29.4 ± 1.0 ^a	30.1 ± 1.1 ^a
Nutrient cycling (%) **	21.3 ± 0.8 ^a	22.5 ± 0.7 ^{ab}	23.8 ± 0.7 ^{ab}	25.9 ± 1.9 ^{ab}	27.9 ± 1.2 ^b	27.7 ± 1.5 ^b	27.8 ± 1.6 ^b	26.4 ± 0.8 ^b	22.8 ± 0.9 ^{ab}	24.2 ± 0.9 ^{ab}
Above-ground biomass (g/m²) **	771.1 ± 270.9 ^{ab}	285.2 ± 47.6 ^a	270.6 ± 44.1 ^a	929.7 ± 179.2 ^{ab}	1069.9 ± 164.9 ^b	645.3 ± 133.1 ^{ab}	633.9 ± 114.4 ^{ab}	342.9 ± 31.5 ^{ab}	329.5 ± 70.2 ^a	289.5 ± 45.5 ^a
Above-ground biomass dry (g/m²) **	655.6 ± 220.9 ^{ab}	241.9 ± 39.4 ^a	221.2 ± 34.5 ^a	769.7 ± 149.1 ^{ab}	837.5 ± 122.7 ^b	510.9 ± 91.6 ^{ab}	506.9 ± 79.9 ^{ab}	310.7 ± 27.5 ^{ab}	274.6 ± 55.8 ^a	240.8 ± 35.7 ^a
Conductivity (µS/cm) **	70.7 ± 11.4 ^b	26.5 ± 1.9 ^a	385.5 ± 80.3 ^b	39.1 ± 3.2 ^{ab}	881.5 ± 178.5 ^c	568.9 ± 160.4 ^{bc}	36.2 ± 2.9 ^{ab}	23.6 ± 1.4 ^a	31.9 ± 2.2 ^{ab}	27.1 ± 2.3 ^a
pH **	6.81 ± 0.06 ^a	6.62 ± 0.05 ^a	6.54 ± 0.07 ^a	6.37 ± 0.04 ^{ab}	5.96 ± 0.06 ^b	6.51 ± 0.08 ^a	6.57 ± 0.07 ^a	6.23 ± 0.08 ^{ab}	6.39 ± 0.07 ^{ab}	6.25 ± 0.09 ^{ab}
Soil moisture (%) **	3.49 ± 0.27 ^b	2.87 ± 0.11 ^b	3.41 ± 0.23 ^b	2.46 ± 0.11 ^{ab}	4.50 ± 0.42 ^b	3.08 ± 0.23 ^b	5.25 ± 0.30 ^{cb}	5.45 ± 0.3 ^c	2.02 ± 0.08 ^{ab}	1.77 ± 0.12 ^a
SOM (%) **	4.73 ± 0.20 ^{cb}	5.50 ± 0.17 ^c	5.13 ± 0.16 ^c	4.19 ± 0.13 ^b	4.01 ± 0.17 ^b	3.86 ± 0.07 ^b	3.72 ± 0.36 ^{ab}	2.83 ± 0.1 ^a	3.40 ± 0.11 ^{ab}	3.34 ± 0.12 ^{ab}
Core depth (cm) **	9.38 ± 0.34 ^a	8.62 ± 0.44 ^a	8.85 ± 0.34 ^a	8.80 ± 0.41 ^a	9.72 ± 0.22 ^a	8.89 ± 0.37 ^a	10.0 ± 0.0 ^a	10.0 ± 0.0 ^a	8.43 ± 0.50 ^a	8.81 ± 0.39 ^a
Core density (g/cm³) **	1.39 ± 0.04 ^a	1.52 ± 0.03 ^{ab}	1.61 ± 0.03 ^{ab}	1.47 ± 0.05 ^{ab}	1.41 ± 0.04 ^a	1.63 ± 0.02 ^{ab}	1.73 ± 0.05 ^{ab}	1.83 ± 0.05 ^b	1.65 ± 0.06 ^{ab}	1.72 ± 0.08 ^{ab}
Stone fresh (g/m²) **	762.1 ± 98.9 ^{ab}	1231.4 ± 107.6 ^{ab}	702.2 ± 59.1 ^{ab}	1819.6 ± 275.5 ^{bc}		1713.6 ± 181.1 ^b	561.5 ± 290.9 ^a	228.3 ± 50.6 ^a	1855.8 ± 384.9 ^{bc}	2669.9 ± 576.6 ^c
Stone dry (g/m²) **	723.4 ± 94.3 ^{ab}	1180.8 ± 100.8 ^{ab}	672.1 ± 57.4 ^{ab}	1703.1 ± 266.2 ^{bc}	1266.8 ± 104.9 ^{ab}	1608.8 ± 173.9 ^b	529.3 ± 276.7 ^a	216.4 ± 46.9 ^a	1780.3 ± 376.2 ^{bc}	2597.9 ± 566.1 ^c
Root fresh (g/m²) **	138.1 ± 25.3 ^{ab}	98.4 ± 15.7 ^a	115.4 ± 42.9 ^{ab}	247.4 ± 46.4 ^b	201.3 ± 32.7 ^b	100.9 ± 16.6 ^{ab}	161.7 ± 28.5 ^{ab}	127.2 ± 14.9 ^{ab}	139.1 ± 31.6 ^{ab}	62.4 ± 10.3 ^a
Root dry (g/m²) **	76.6 ± 16.9 ^{ab}	52.6 ± 9.4 ^{ab}	63.6 ± 27.4 ^{ab}	139.2 ± 30.7 ^b	87.9 ± 17.6 ^{ab}	47.2 ± 6.9 ^{ab}	79.2 ± 15.5 ^{ab}	61.3 ± 7.9 ^{ab}	89.1 ± 23.2 ^{ab}	30.9 ± 5.4 ^a

890

891 **Table 2.5** Comparisons between patch types from Vaal River showing the mean and standard error calculated from individual patches.
892 Variables with superscripted asterisks are significantly different at the 5% level (*) and 1% level (**), values with superscripted
893 common letters are not significantly different while those with different letters were significantly different at the 5% level using
894 a non-parametric Kruskal-Wallis with Multiple Comparison test.

	Biological soil crust patch	Bare grass patch	Sparse grass patch	Grass patch	Tall grass patch	<i>Schoenoplectus</i> patch	<i>S. plumosum</i> patch
N	42	56	145	5	11	5	27
Stability (%) **	62.0 ± 0.4 ^a	65.2 ± 0.4 ^{ab}	66.4 ± 0.3 ^{ab}	72.3 ± 0.8 ^c	63.2 ± 1.2 ^{ab}	68.9 ± 1.1 ^{bc}	72.7 ± 0.7 ^c
Infiltration (%) **	24.2 ± 0.5 ^a	29.1 ± 0.7 ^{ab}	31.8 ± 0.3 ^{ab}	35.5 ± 2.2 ^b	33.5 ± 0.8 ^b	36.8 ± 1.5 ^b	47.9 ± 1.5 ^b
Nutrient cycling (%) **	19.3 ± 0.4 ^a	24.7 ± 0.5 ^{ab}	24.0 ± 0.4 ^{ab}	28.4 ± 2.7 ^b	22.0 ± 1.3 ^{ab}	26.6 ± 0.6 ^b	39.9 ± 1.3 ^b
Above-ground fresh biomass (g.m⁻²) **		175.7 ± 16.0 ^a	460.5 ± 19.6 ^{ab}	623.6 ± 153.3 ^{ab}	2902.2 ± 618.3 ^b	2726.9 ± 199.5 ^b	1603.0 ± 119.0 ^b
Above-ground dry biomass (g.m⁻²) **		164.7 ± 13.4 ^a	389.5 ± 16.1 ^{ab}	552.1 ± 128.4 ^b	2506.5 ± 487.9 ^c	2067.2 ± 166.8 ^c	1173.3 ± 80.9 ^{bc}
Conductivity (µS.cm⁻¹) **	341.8 ± 116.5 ^{ab}	176.4 ± 46.2 ^{ab}	171.0 ± 36.7 ^{ab}	60.3 ± 10.2 ^a	51.0 ± 5.9 ^a	511.4 ± 193.1 ^b	459.8 ± 173.6 ^b
pH **	6.57 ± 0.08 ^b	6.57 ± 0.07 ^b	6.38 ± 0.03 ^{ab}	6.55 ± 0.12 ^{ab}	6.53 ± 0.07 ^{ab}	6.03 ± 0.05 ^a	6.22 ± 0.09 ^{ab}
Soil moisture (% mass) **	2.28 ± 0.19 ^a	2.34 ± 0.11 ^a	3.78 ± 0.15 ^b	3.73 ± 0.38 ^{ab}	3.15 ± 0.28 ^{ab}	4.44 ± 0.32 ^b	4.14 ± 0.28 ^b
SOM (% mass)	4.32 ± 0.22 ^a	4.20 ± 0.13 ^a	4.0 ± 0.1 ^a	3.65 ± 0.14 ^a	4.69 ± 0.28 ^a	4.3 ± 0.39 ^a	4.66 ± 0.37 ^a
Core depth (cm) **	7.12 ± 0.43 ^a	7.89 ± 0.35 ^a	9.83 ± 0.08 ^b	10 ± 0 ^b	9.75 ± 0.25 ^{ab}	10 ± 0 ^b	10 ± 0 ^b
Bulk density (g.cm⁻³) **	1.75 ± 0.05 ^b	1.73 ± 0.03 ^b	1.54 ± 0.02 ^a	1.49 ± 0.04 ^{ab}	1.35 ± 0.09 ^a	1.2 ± 0.12 ^a	1.40 ± 0.05 ^a
Stone fresh (g.m⁻²) **	2585.4 ± 392.2 ^b	2038.2 ± 309.9 ^b	836.9 ± 76.8 ^a	1302.5 ± 112.1 ^{ab}	696.5 ± 129.8 ^a		782.6 ± 130.0 ^a
Stone dry (g.m⁻²) **	2431.5 ± 326.3 ^b	1959.2 ± 302.4 ^{ab}	829.5 ± 67.5 ^a	1216.3 ± 105.8 ^{ab}	649.3 ± 120.6 ^a	922.0 ± 238.9 ^{ab}	795.3 ± 95.9 ^a
Root mass, fresh (g.m⁻²) **	26.4 ± 3.2 ^a	57.5 ± 13.2 ^a	185.0 ± 17.2 ^b	142.7 ± 28.9 ^{ab}	230.6 ± 52.9 ^b	362.4 ± 85.7 ^b	200.5 ± 37.1 ^b
Root mass, dry (g.m⁻²) **	11.6 ± 1.5 ^a	33.5 ± 9.2 ^a	98.4 ± 11.4 ^b	68.5 ± 14.5 ^{ab}	114.6 ± 29.1 ^b	161.9 ± 51.6 ^b	105.2 ± 21.2 ^b

895 2.3.4 Patch Measures for West Wits

896 Transect ww1 had the highest level of biogeophysical functioning when comparing the three
897 LFA indices (71.8 ± 0.9 , 37.7 ± 1.4 and $34.2 \pm 1.6\%$ for stability, infiltration and nutrient
898 cycling, respectively) and transect ww3 has the lowest values (61.9 ± 1.5 , 32.8 ± 2.5 and 24.6
899 $\pm 2.4\%$, respectively; Table 2.6). However these differences were only significant for stability
900 and nutrient cycling, and for infiltration there were no statistically significant differences
901 between the West Wits transects. For conductivity measurements, transect ww4 had
902 significantly higher values but this was for a small sample and the differences in conductivity
903 for the rest of the transects were not significantly different at the 5% probability level.
904 Likewise, soil pH was not found to have any statistical differences between transects at
905 West Wits. This lack of statistically significant differences also applied to shoot biomass, soil
906 extractable ammonium and nitrate, inorganic soil nitrogen as calculated by the sum of
907 nitrate and ammonium measures, and the nutrient cycling ratio index.

908 However organic nitrogen was significantly higher in transects ww4 ($37.7 \pm 2.0 \text{ mg.g}^{-1} \text{ soil}$)
909 and ww3 ($34.2 \pm 1.9 \text{ mg.g}^{-1} \text{ soil}$) when compared to transect ww2 which had the lowest
910 organic nitrogen value ($25.8 \pm 1.3 \text{ mg.g}^{-1} \text{ soil}$; Table 2.5). Soil moisture was significantly
911 higher in transect ww1 ($12.7 \pm 0.3\%$ by mass) than the other three transects where transect
912 ww3 had the lowest soil moisture ($7.86 \pm 0.46\%$ by mass). The pattern for SOM also showed
913 significant differences where transect ww4 had the highest value ($8.32 \pm 0.90\%$ by mass)
914 compared to the other transects. However it should be noted that this transect was only
915 represented by nine samples and the other three transects with many more samples (22, 25
916 and 18) in each transect showed no significant differences (Table 2.6).

917 When examining the West Wits data by patch type, bare patches had consistently lower
918 values than did the other patch types (Table 2.7). Bare patches had significantly lower mean
919 values for the LFA indices of stability, infiltration and nutrient cycling ($57.2 \pm 1.8\%$, $22.1 \pm$
920 1.4% and $14.3 \pm 1.8\%$, respectively), while *S. plumosum* patches had the highest values for
921 the same three indices ($72.6 \pm 0.6\%$, $43.4 \pm 0.7\%$ and $38.2 \pm 0.8\%$, respectively). Bare
922 patches had significantly lower pH values (4.88 ± 0.07) than *S. plumosum* which had the
923 highest pH values (5.45 ± 0.07). *Seriphium plumosum* patches also had significantly higher
924 above-ground biomass (1901.5 ± 124.6 and $1517.9 \pm 96.3 \text{ g.m}^{-2}$ for fresh and dry

925 respectively) and SOM values ($5.72 \pm 0.51\%$ by mass) compared to sparse grass patches
926 which had the lowest above-ground biomass (549.5 ± 65.5 and 515.2 ± 60.5 for fresh and
927 dry respectively) and bare patches which had the lowest SOM values ($3.84 \pm 0.23\%$ by
928 mass). Bare patches also had significantly lower soil moisture ($6.59 \pm 0.36\%$ by mass) when
929 compared to grass patches ($11.8 \pm 0.4\%$ by mass) which had the highest soil moisture. Grass
930 patches also had significantly higher extractable soil ammonium ($1.65 \pm 0.13 \text{ mg.g}^{-1}$)
931 compared to bare patches which had the lowest extractable soil ammonium (1.05 ± 0.27
932 mg.g^{-1} , Table 2.7). However no significant differences were observed for soil nitrate,
933 inorganic nitrogen, organic nitrogen and nutrient cycling ratio index.

934 2.3.5 **Spatial Autocorrelation of Patch Measures**

935 In essence spatial autocorrelation is testing for the independence of the samples, and as
936 samples were not chosen in a random spatial fashion, but restricted to transects and areas
937 of specific patch types within a transect following the LFA protocol, the element of spatial
938 independence in samples is already, to some extent, compromised. Testing for spatial
939 autocorrelation using Moran's I and pooling all samples from Vaal River showed that for all
940 variables, observed spatial autocorrelation was significantly different ($P < 0.01$) to the null
941 hypothesis of no spatial autocorrelation with the exception of grams of root per m^2 which
942 was only significant at the 5% level (results not shown). However examining transects and
943 variables individually (10 transects * 11 variables = 110 tests) showed that approximately
944 20% of these tests produced observed Moran's I results that differed significantly from
945 expected values of autocorrelation (22 total with 15 at $P > 0.01$ and 7 at $P > 0.05$). To
946 illustrate this pattern, Table 2.8 shows a selection of the results when testing above-ground
947 biomass from the Vaal River transects. The overall data is highly significant ($P < 0.01$)
948 whereas only 2 transects (vr7 and vr8 with $P < 0.0000$ and $P = 0.0076$, respectively) showed
949 significant autocorrelation between measurements. However, transect vr7 had the highest
950 observed Moran's I value of 0.198 for above-ground biomass (Table 2.8) which is low on the
951 Moran's I scale where a value of 1 equals very high autocorrelation or clustering while a
952 value of -1 represents a very dispersed or negative spatial correlation. A value of zero for
953 Moran's I would indicate no spatial autocorrelation. Therefore, although results for spatial
954 autocorrelation for transects vr7 and vr8 and for all points at Vaal River are highly
955 significant, the observed values for spatial autocorrelation are far closer to 0 than they are

Table 2.6

Comparisons of transects from West Wits showing the mean and standard error calculated from individual patches within each transect. Variables with superscripted asterisks are significantly different at the 5% level (*) and 1% level (**), while values with superscripted common letters were not significantly different and those with different letters were significantly different at a 5% level using a non-parametric Kruskal-Wallis with Multiple Comparison test.

	ww1	ww2	ww3	ww4	ww5
N	22	25	18	9	1
Stability (%) **	71.8 ± 0.9 ^b	70.7 ± 1.2 ^b	61.9 ± 1.5 ^a	69.7 ± 1.5 ^{ab}	62.5 ± 0
infiltration (%)	37.7 ± 1.4 ^a	35.6 ± 1.3 ^a	32.8 ± 2.5 ^a	36.8 ± 3.4 ^a	28.05 ± 0
Nutrient cycling (%) *	34.2 ± 1.6 ^b	31.8 ± 1.7 ^{ab}	24.6 ± 2.4 ^a	29.7 ± 2.8 ^{ab}	23.23 ± 0
Conductivity (µS.cm⁻¹) **	47.7 ± 4.8 ^{ab}	32.5 ± 2.6 ^a	42.3 ± 3.8 ^{ab}	63.6 ± 11.7 ^b	15 ± 0
pH *	5.18 ± 0.06 ^a	5.36 ± 0.08 ^a	5.41 ± 0.08 ^a	5.04 ± 0.08 ^a	5.38 ± 0
Shoot fresh biomass (g.m⁻²)	1371.0 ± 149.7 ^a	1305.9 ± 181.2 ^a	928.7 ± 206.5 ^a	1234.2 ± 228.0 ^a	292 ± 0
Shoot dry biomass (g.m⁻²)	1152.0 ± 113.3 ^a	1123.0 ± 144.9 ^a	814.5 ± 172.3 ^a	945.4 ± 157.9 ^a	295.2 ± 0
Soil moisture (% mass) **	12.7 ± 0.3 ^b	9.72 ± 0.41 ^a	7.86 ± 0.46 ^a	8.01 ± 0.96 ^a	4.92 ± 0
SOM (% mass) **	4.31 ± 0.11 ^a	4.22 ± 0.16 ^a	5.28 ± 0.45 ^{ab}	8.32 ± 0.9 ^b	5.25 ± 0
Ammonium-N (µg.g⁻¹ soil)	1.27 ± 0.12 ^a	1.36 ± 0.11 ^a	1.49 ± 0.32 ^a	1.5 ± 0.17 ^a	0.55 ± 0
Nitrate-N (µg.g⁻¹ soil)	1.77 ± 0.13 ^a	1.45 ± 0.12 ^a	1.76 ± 0.28 ^a	2.24 ± 0.21 ^a	2 ± 0
Organic-N (µg.g⁻¹ soil) **	26.8 ± 1.2 ^a	25.8 ± 1.3 ^a	34.2 ± 1.9 ^b	37.7 ± 2.0 ^b	31.76 ± 0
Inorganic-N (µg.g⁻¹ soil)	3.04 ± 0.2 ^a	2.81 ± 0.16 ^a	3.25 ± 0.52 ^a	3.74 ± 0.26 ^a	2.55 ± 0
Nutrient cycling ratio index	9.33 ± 0.55 ^a	9.74 ± 0.62 ^a	15.5 ± 2.7 ^a	10.2 ± 0.3 ^a	12.46 ± 0

Table 2.7 Comparisons between patch types from Vaal River showing the mean and standard error calculated from individual patches. Variables with superscripted asterisks are significantly different at the 5% level (*) and 1% level (**), while values with superscripted common letters are not significantly different and those with different superscripted letters were significantly different at a 5% level using a non-parametric Kruskal-Wallis with Multiple Comparison test.

	Bare patch	Sparse grass	Grass patch	<i>S. Plumosum</i>
N	9	17	22	27
LFA stability (%) **	57.2 ± 1.8 ^a	63.5 ± 0.9 ^a	72.6 ± 0.3 ^b	72.6 ± 0.6 ^b
LFA infiltration (%) **	22.1 ± 1.4 ^a	29.1 ± 1.0 ^a	36.5 ± 0.7 ^b	43.4 ± 0.7 ^c
LFA nutrient cycling (%) **	14.3 ± 1.8 ^a	23.0 ± 1.0 ^a	33.2 ± 0.9 ^b	38.2 ± 0.8 ^b
Conductivity (µS.cm⁻¹)	35.7 ± 3.1 ^a	37.7 ± 3.2 ^a	41.2 ± 4.3 ^a	49.8 ± 5.7 ^a
pH **	4.88 ± 0.07 ^a	5.12 ± 0.07 ^{ab}	5.36 ± 0.05 ^{bc}	5.45 ± 0.07 ^c
Shoot fresh biomass (g.m⁻²) **	0 ± 0	549.5 ± 65.5 ^a	1374.7 ± 90.2 ^b	1901.5 ± 124.6 ^b
Shoot dry biomass (g.m⁻²) **	0 ± 0	515.2 ± 60.5 ^a	1233.7 ± 73.5 ^b	1517.9 ± 96.3 ^b
Soil moisture (% mass) **	6.59 ± 0.36 ^a	8.77 ± 0.69 ^{ab}	11.8 ± 0.4 ^c	10.1 ± 0.5 ^{bc}
SOM (% mass) *	3.84 ± 0.23 ^a	5.03 ± 0.40 ^{ab}	4.59 ± 0.14 ^{ab}	5.72 ± 0.51 ^b
Ammonium (µg.g⁻¹ soil) **	1.05 ± 0.27 ^a	1.37 ± 0.30 ^{ab}	1.65 ± 0.13 ^b	1.26 ± 0.10 ^{ab}
Nitrate (µg.g⁻¹ soil)	2.21 ± 0.37 ^a	1.48 ± 0.18 ^a	1.65 ± 0.15 ^a	1.77 ± 0.14 ^a
Organic-N (µg.g⁻¹ soil)	25.5 ± 3.32 ^a	31.3 ± 2.0 ^a	28.4 ± 1.3 ^a	30.8 ± 1.5 ^a
Inorganic-N (µg.g⁻¹ soil)	3.26 ± 0.59 ^a	2.85 ± 0.45 ^a	3.30 ± 0.22 ^a	3.02 ± 0.19 ^a
Nutrient cycling ratio index *	8.34 ± 0.61 ^a	14.7 ± 2.5 ^a	9.23 ± 0.62 ^a	11.3 ± 1.0 ^a

to 1 suggesting that spatial autocorrelation for above-ground biomass is low at Vaal River. West Wits results have not been presented but showed similar patterns to those described for Vaal River. Thus, as all across-site spatial autocorrelation analyses showed significant results, one can draw the conclusion that across the two sites, Moran's I is detecting that sampling was clustered at the level of being localized to specific areas where a transect was established. However, as 93% of 184 tests of specific transects or across-site tests produced values between -0.2 and 0.2 for Moran's I, any spatial autocorrelation between measures within a transect or across the two mining regions was low or non-existent. Therefore taking into account the above points it is concluded that spatial autocorrelation can be regarded as a negligible factor in interpreting further results and all points can be regarded as independent samples.

981 **Table 2.8** Spatial autocorrelation using Moran's I Autocorrelation index for shoot
 982 biomass at Vaal River as a whole and for individual transects from Vaal River.
 983

Transect	Observed	Expected	sd	Probability
Vr all	0.158	-0.004	0.021	0.0000**
vr1	-0.064	-0.037	0.041	0.5125
vr2	0.007	-0.042	0.069	0.4862
vr3	0.002	-0.028	0.048	0.5287
vr7	0.198	-0.029	0.046	0.0000**
vr8	0.185	-0.045	0.086	0.0076**
vr9	-0.059	-0.053	0.078	0.9395
vr10	-0.022	-0.042	0.068	0.7725
vr11	-0.014	-0.071	0.069	0.4002
vr13	-0.099	-0.038	0.062	0.3264
vr14	-0.061	-0.038	0.062	0.7179

984

985 2.3.6 Correlations between Patch Variables

986 Some variables from Vaal River sites that are highly correlated such as the LFA infiltration
 987 index and above-ground fresh and dry biomass (Figure 2.10a) show similar patterns of
 988 correlation in the data from West Wits sites (Figure 2.10b), however this pattern does not
 989 hold for all variables common to the two sites. Furthermore, the strong binomial nature of
 990 the West Wits correlations is not displayed in the Vaal River correlations although it should
 991 be noted that the set of variables measured at the two mining sites are slightly different. For
 992 instance, soil nitrogen was not measured for Vaal River sites, whereas gravel and soil below-
 993 ground organic matter (root) content was not measured at West Wits sites. At both Vaal
 994 River and West Wits, the three LFA variables of soil stability, infiltration and nutrient cycling
 995 calculated for each sample site were positively correlated with each other, above-ground
 996 biomass and soil moisture (Figure 2.10a and b, respectively). However, the correlation was
 997 much stronger for West Wits than it was for Vaal River variables. Another difference
 998 between the two sites was that soil pH at West Wits was positively correlated with the
 999 above group of LFA indices and particularly above-ground vegetation, while it had a weak,
 1000 negative correlation with the same group of variables at Vaal River. This difference in

1001 correlation for soil pH between sites may reflect differences in soil type between the two
1002 mining regions.

1003 The strong correlation between fresh and dry above-ground biomass at both mining sites is
1004 important as sampling was conducted moving into winter when most of the vegetation,
1005 such as the graminoid component, is becoming dry and winter-senesced or is already in that
1006 condition, while other vegetation, such as *S. plumosum*, is evergreen and therefore moist.
1007 Thus, the strong correlation between fresh and dry above-ground vegetation means that
1008 only one of these variables needs to be considered.

1009 The strong correlation between the LFA indices of stability, infiltration and nutrient cycling is
1010 to be expected (Figure 2.10a and b) when one considers that there is considerable overlap
1011 in the eleven soil surface assessment indicators (SSAI) employed to calculate each index and
1012 therefore some covariation is expected (Table 2.2). For instance stability is calculated with
1013 eight of the SSAIs and the infiltration index is calculated from six of the SSAIs with 2½ SSAIs
1014 common to both calculations. Likewise, the nutrient cycling index uses four SSAIs to
1015 calculate its index value and shares two of these with the stability index and the other two
1016 with the infiltration index. This argument is supported by the fact that at both Vaal River
1017 and West Wits, the stability index has a stronger correlation with the nutrient cycling index
1018 ($r = 68$ and 91 for Vaal River and West Wits, respectively) than with the infiltration index ($r =$
1019 44 and 79 for Vaal River and West Wits, respectively). The correlation between the
1020 infiltration and nutrient cycling indices is the strongest ($r = 85$ and 95 for Vaal River and
1021 West Wits, respectively) which mirrors the fact that these two indices share the greatest
1022 degree of overlap in SSAIs relative to the total number of SSAIs used to calculate their value.

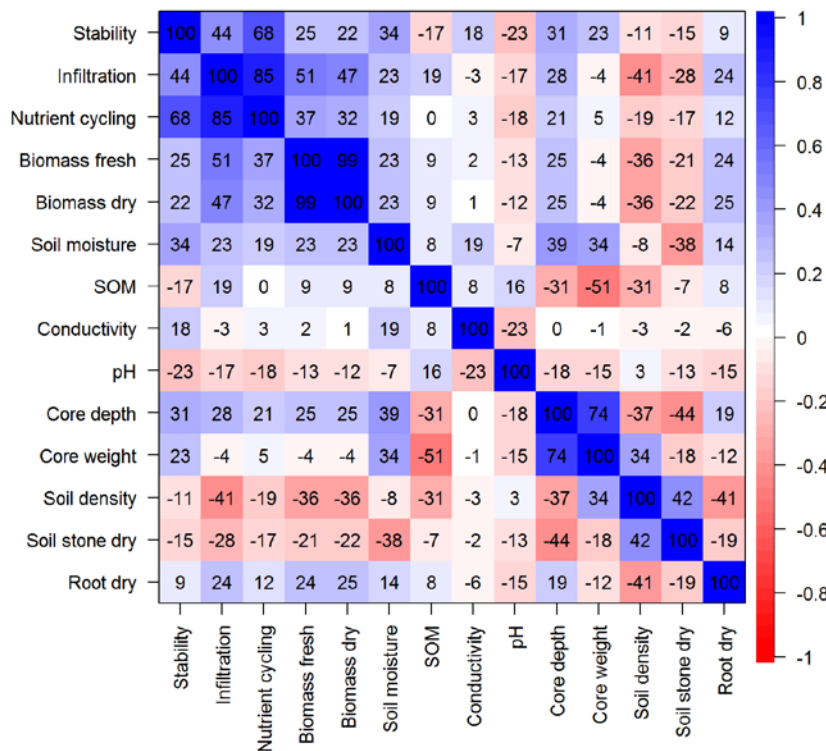
1023 At Vaal River, the amount of stone in the soil (> 2 mm) and bulk soil density had a weakly
1024 negative correlation with the LFA indices, while the core depth had a weakly positive
1025 correlation with the LFA indices and with above-ground biomass. Likewise below-ground
1026 root biomass had a weakly positive correlation with the LFA indices. These patterns are as
1027 expected where an increase in an LFA index should indicate increased biological
1028 functionality, as should increased above- or below-ground biomass; while increased soil
1029 stoniness, soil bulk density and shallower soil cores are likely to indicate lower biological
1030 functionality.

The over-riding pattern observable in the correlation of variables from West Wits is the strong separation of the data into two uncorrelated groups (Figure 2.10b). The three LFA indices of soil stability, soil infiltration and nutrient cycling, together with above-ground biomass, soil moisture and soil solution pH form a group of positively correlated variables. A second group of variables, uncorrelated to the first group, included the inorganic soil nitrogen ions of ammonium and nitrate, organic soil nitrogen, soil electrical conductivity and SOM. All these variables were positively correlated with each other but showed no correlation to the first group consisting of the vegetation and LFA indices.

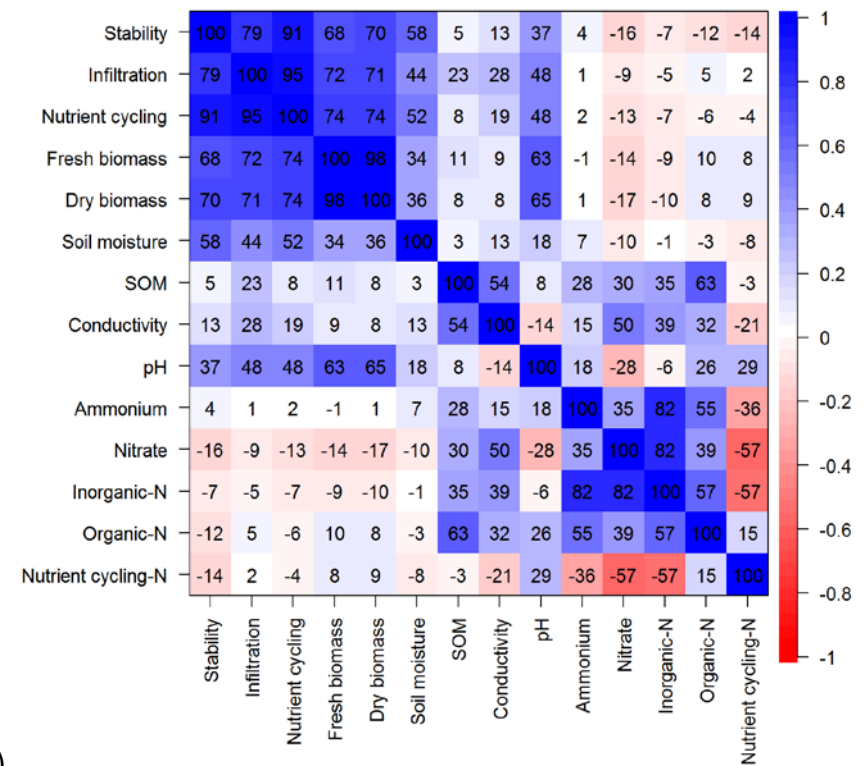
2.3.7 Principal Components Analysis (PCA)

The dominant observation to note in the PCA results is how the variables measured at the West Wits site sort into two clearly uncorrelated groups organized either directly along the axis of the first or second principal component (Figure 2.11b) and these strongly equate to the two groups identified through correlation analysis. For measurements from Vaal River there is no dominant pattern with variables more-or-less evenly distributed around the space defined by the first two principal components (Figure 2.11a). However the set of variables measured for each site and therefore used in the respective PCA analyses are slightly different (Table 2.9 and 2.10) and therefore the PCA results for each site shouldn't be compared. For West Wits, the first two components described 59.4% of the variability and the first three components described 72.3% of the variability (Table 2.10). However, for Vaal River, the first five principal components could only explain 74.1% (Table 2.9) and the first two accounted for 44.6% of the variability. This suggests that for Vaal River, PCA is not providing a robust interpretation of the values and variables measured at this site.

For the Vaal River, the first principal component described 27.5% of the variability and was dominated by no specific variable (Table 2.9). However the three LFA indices of infiltration (0.4055), nutrient cycling (0.3785) and stability (0.3402), together with soil-core depth and to a lesser extent dry shoot biomass showed the highest loadings for the first principal component. The second component, which accounted for 17% of the variability, was dominated by core-fresh weight, SOM and to a lesser extent soil bulk-density and soil-core depth. The third component only accounted for 12.2% of the variability and was dominated



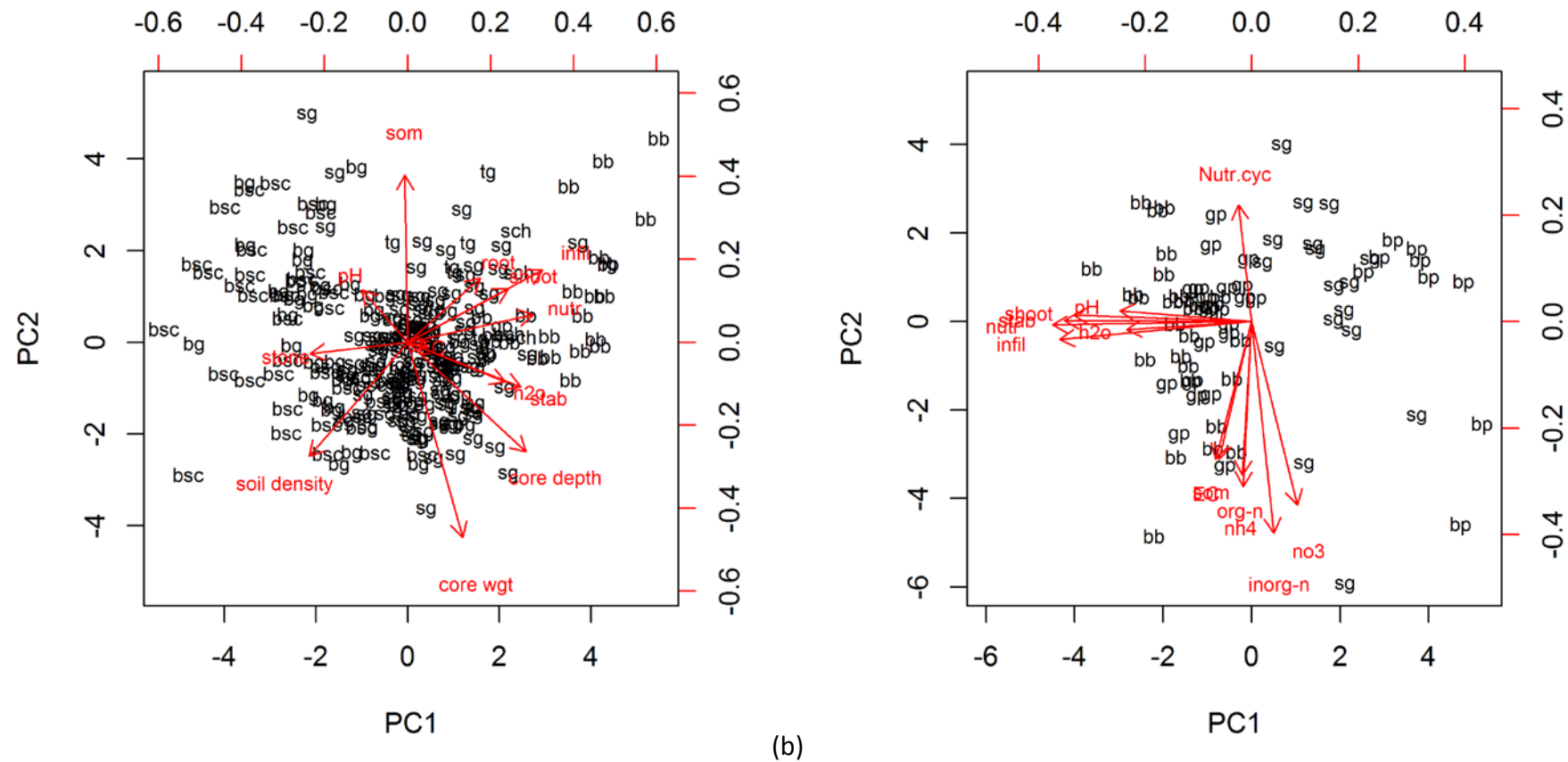
(a)



(b)

Figure 2.10 Correlograms showing correlations between variables for patches from (a) Vaal River and (b) West Wits. The numbers in each block indicate the actual correlation between the two variables.

1064



1065 (a) 1066 (b) 1067 **Figure 2.11** Biplots of the first two components from the PCA for (a) Vaal River and (b) West Wits. Character symbols in each plot refer to 1068 the following patch types: bb = *S. plumosum*, bg = bare grass, bp = bare patch, bsc = biological soil crust, gp = grass patch, 1069 sch = *Schoenoplectus* patch, sg = sparse grass and tg = tall grass.

1070

1071 **Table 2.9** Principal Components Analysis for the Vaal River data showing the first five
1072 principal components

	PC1	PC2	PC3	PC4	PC5
Standard deviation	1.892	1.488	1.261	1.099	1.021
Proportion of Variance (%)	27.5	17.0	12.2	9.3	8.0
Cumulative Proportion (%)	27.5	44.6	56.8	66.1	74.1
Stability	0.3402	-0.1342	0.3515	-0.0598	0.2041
Infiltration	0.4055	0.2171	0.2410	0.2113	0.1393
Nutrient cycling	0.3785	0.0830	0.4244	0.2031	0.2654
Conductivity	0.0592	-0.0118	0.1864	-0.7778	-0.0743
pH	-0.1382	0.1575	-0.3753	0.1981	0.5869
Soil moisture	0.2939	-0.1200	-0.1942	-0.3948	0.2352
SOM	-0.0084	0.5032	-0.0813	-0.2684	0.2179
Soil core depth	0.3559	-0.3291	-0.3082	0.0555	-0.1393
Soil-core fresh weight	0.1655	-0.5874	-0.1407	0.0438	0.0791
Soil density	-0.2960	-0.3434	0.2891	0.0468	0.2291
Shoot biomass, dry	0.3033	0.1601	-0.0073	0.0877	-0.1936
Soil stoniness	-0.2931	-0.0343	0.4426	0.0800	-0.2206
Root biomass	0.2190	0.1932	-0.1550	0.1310	-0.5015

1073

1074 by soil stoniness, LFA nutrient cycling index, soil pH and the LFA soil stability index (Table
1075 2.9).

1076 There does not seem to be any readily apparent trends explaining how these variables
1077 relate to the PC axes for the Vaal River data and this is supported by the biplot (Figure
1078 2.11a) where variables distribute themselves almost evenly around the space defined by the
1079 first two principal components. Sample points, as denoted by character codes for patch
1080 types in Figure 2.11a, mostly congregate around the center of the biplot space defined by
1081 the first two principal components. However there is a trend visible in Figure 2.11a of patch
1082 types with no or very low above-ground (shoot) biomass (bsc and bg codes respectively)
1083 being located in the negative area of the first principal component, and patch types with
1084 high above ground biomass (bb) being located towards the positive zone of the first

1085 **Table 2.10** The first five components from the Principal Components Analysis of the
1086 West Wits Variables

	PC1	PC2	PC3	PC4	PC5
Standard deviation	2.038	1.889	1.293	1.109	0.854
Proportion of Variance (%)	31.9	27.4	12.9	9.5	5.6
Cumulative Proportion (%)	31.9	59.4	72.3	81.7	87.3
Stability	-0.4398	0.0027	-0.2188	0.0222	-0.0406
Infiltration	-0.4507	-0.0420	-0.0364	-0.1381	0.1623
Nutrient cycling	-0.4670	-0.0077	-0.1339	-0.0323	0.1089
Conductivity	-0.0846	-0.3231	-0.0693	-0.5744	0.0373
pH	-0.3091	0.0247	0.3957	0.3419	0.2056
Shoot biomass dry	-0.4177	0.0156	0.0912	0.0949	0.2802
Soil moisture	-0.2937	-0.0190	-0.2030	0.0108	-0.8282
SOM	-0.0762	-0.3234	0.3088	-0.4161	-0.0956
Soil-extractable Ammonium	-0.0205	-0.3874	0.0964	0.4948	-0.1967
Soil-extractable Nitrate	0.1070	-0.4306	-0.2107	-0.0778	0.2422
Soil organic nitrogen	-0.0219	-0.3614	0.5012	0.0007	-0.1364
Soil inorganic nitrogen	0.0522	-0.4968	-0.0685	0.2549	0.0264
Nutrient cycling ratio	-0.0312	0.2739	0.5630	-0.1857	-0.1619

1087

1088 component. The implication is that the patch definitions used in the field to categorize the
1089 transect space were significant in some ecological sense.

1090 Examining the influence of variables on specific principal components for West Wits (Table
1091 2.10), it was clear that no specific variables dominated any of the first three components.
1092 The first PCA component from West Wits had maximum correlation with four variables, as
1093 indicated by the loading values. These four variables were the three LFA indices of nutrient
1094 cycling, infiltration, stability together with dry above-ground biomass, and to a lesser extent
1095 soil solution pH and soil moisture. The second principal component for West Wits accounted
1096 for almost the same amount of variability as the first component. This second component
1097 was dominated by inorganic soil nitrogen and its two components, nitrates and ammonium,
1098 organic soil nitrogen, SOM and soil conductivity (Table 2.9). These are all chemical
1099 characteristics and therefore the second component may represent soil chemistry while

1100 noting that soil solution pH featured much more strongly on the first component and was
1101 negligible on the second component. The first component is possibly related to vegetation.
1102 The third component, which accounted for 12.9% of variance in the West Wits data, was
1103 best correlated with the nutrient cycling ratio and soil organic nitrogen. The fact that
1104 organic and inorganic nitrogen dominate different PC axes may suggest that some of the
1105 factors influencing these two nitrogen pools are different.

1106 The biplot for West Wits using patch types to code for sample points showed that patches
1107 tended to group together in relation to at least one of the principal components. Bare
1108 patches (bp) tended to cluster in the upper right-hand quadrant (Figure 2.11b) meaning
1109 these sample points scored high on the first principal component, and mostly scored fairly
1110 high in the second principal component although a few bare patches scored very low on the
1111 second component. Another trend visible in the biplot (Figure 2.11b) is that patch types
1112 distribute themselves across the first principal component where patches to the right are
1113 bare patches corresponding with very low above-ground biomass and as we move to the
1114 left-hand side of PC1 the patch types have increasing biomass with *S. plumosum* patches on
1115 the left (bb) having the highest biomass.

1116 2.4. Discussion

1117 2.4.1 Patches Types

1118 Landscape Function Analysis was developed as a system to rapidly assess the effectiveness
1119 of a site in structuring its biophysical components and how tightly these biophysical
1120 components cycled resources such as soil, water and nutrients (Tongway and Hindley,
1121 2004). The identification of patch types and their delineation is the first step in an LFA
1122 transect. Wiens (1976) defined patches as areas defined by discontinuities in “ecological
1123 character states from their surroundings” and these discontinuities must have some sort of
1124 biological importance. Levin and Paine (1974) described a patch as an area separated from a
1125 “homogenous reference background” by some form of discontinuity with no limits in size,
1126 persistence, invasiveness or species composition. Watt (1947) defined patches as
1127 aggregations of individuals of different plant species, where different patches have a
1128 different species assemblage, and together these patches form a mosaic that constitutes a
1129 community. In LFA, a patch is an area where resources tend to accumulate (sink) and its
1130 boundary is the line where deposition starts (Tongway and Hindley, 2004).

1131 In its simplest form, a patch is separated from other patches by an inter-patch area from
1132 where resources tend to be removed (source) or transported over from other patches
1133 (Tongway and Hindley, 2004). Worrall (1960) described such patchiness consisting of patch
1134 and bare-ground inter-patches in the Sudan and noted that at about the 300 mm rainfall
1135 isohyet such patches coalesce and bare-ground becomes rare. Worrall (1960) further points
1136 out that below the 300 mm isohyets, wind may be responsible for some of these two-phase
1137 mosaics through feedback between vegetation, deposition and soil moisture retention.
1138 Below the 300 mm isohyet, inter-patch areas increased in size while patch size decreased as
1139 one moved down the rainfall gradient into the desert.

1140 Lauenroth and Coffin (Burke *et al.*, 1998) propose a basic model for natural vegetation that
1141 follows a precipitation gradient and divide this precipitation gradient into three zones, each
1142 zone with the following characteristics: below-ground constraints, indeterminate
1143 dominance and above-ground constraints. Grasslands in this model generally occupy the
1144 portion of the precipitation axis below 1000 mm of annual rainfall which falls into either the
1145 below-ground constraints (≤ 700 mm annual rainfall) or the indeterminate dominance zone

1146 (>700 mm annual rainfall) although the cross-over point between the two will be influenced
1147 by soil texture and mean annual temperature as they influence soil moisture availability
1148 (Burke *et al.*, 1998). At both research sites, where average annual precipitation was 565 mm
1149 and 661 mm for Vaal River and West Wits respectively, the grassland should form a
1150 continuous layer in support of Worrall's (1960) observations in Sudan. However careful
1151 examination at both sites showed that between the grass tufts were runoff areas.
1152 Furthermore, the most common patch type at Vaal River and the second most common
1153 patch type at West Wits was sparse grass (sg) which was defined as tufts with the soil
1154 surface visible which implies an open canopy structure. At West Wits grass patches (gp)
1155 dominated. These grass patches were defined as grass tufts in which the soil surface is not
1156 visible through the canopy. Thus these two sites fit neatly into Lauenroth and Coffin's (1992)
1157 model where Vaal River, being lower on the precipitation gradient shows clear signs of
1158 below-ground limitation whereas West Wits sites, being higher on the precipitation
1159 gradient, seem to be closer to the boundary between the below-ground limited and
1160 indeterminate dominance zone.

1161 In the grass component of these grasslands, the patch size under a purely LFA definition was
1162 generally at the level of the individual tuft or bunch as generally there were narrow runoff
1163 areas between the grass tufts where overland flow would run. However, for reasons
1164 outlined in the methods section, grass tufts and associated runoff zones between tufts were
1165 merged and these grasslands classified into two patch types based on a homogenous
1166 density of tufts. So essentially these grasslands were mostly sparse grass patches where
1167 canopies from tufts did not form a closed canopy, or less commonly as grass patches where
1168 the aerial portion of the grass did form a mostly closed canopy between tufts. Under this
1169 definition, a variable number of grass tufts could be included in a single patch whereas
1170 under a rigorous LFA definition such a patch would actually be a number of patches. This
1171 distinction becomes important in interpreting LFA values as changes in patch size relative to
1172 a reference site or time series of measurements can be important indicators of efficient
1173 resource cycling or loss thereof and concomitant degradational processes in a landscape.
1174 The basal area and density of grass tufts has been shown to be related to a complex
1175 interplay between annual precipitation, edaphic conditions (Fuhlendorf *et al.*, 2001, Worrall,

1960) and above-ground factors such as fire, herbivory and competition for light (Koerner and Collins, 2013, Burke *et al.*, 1998, Wedin, 1995).

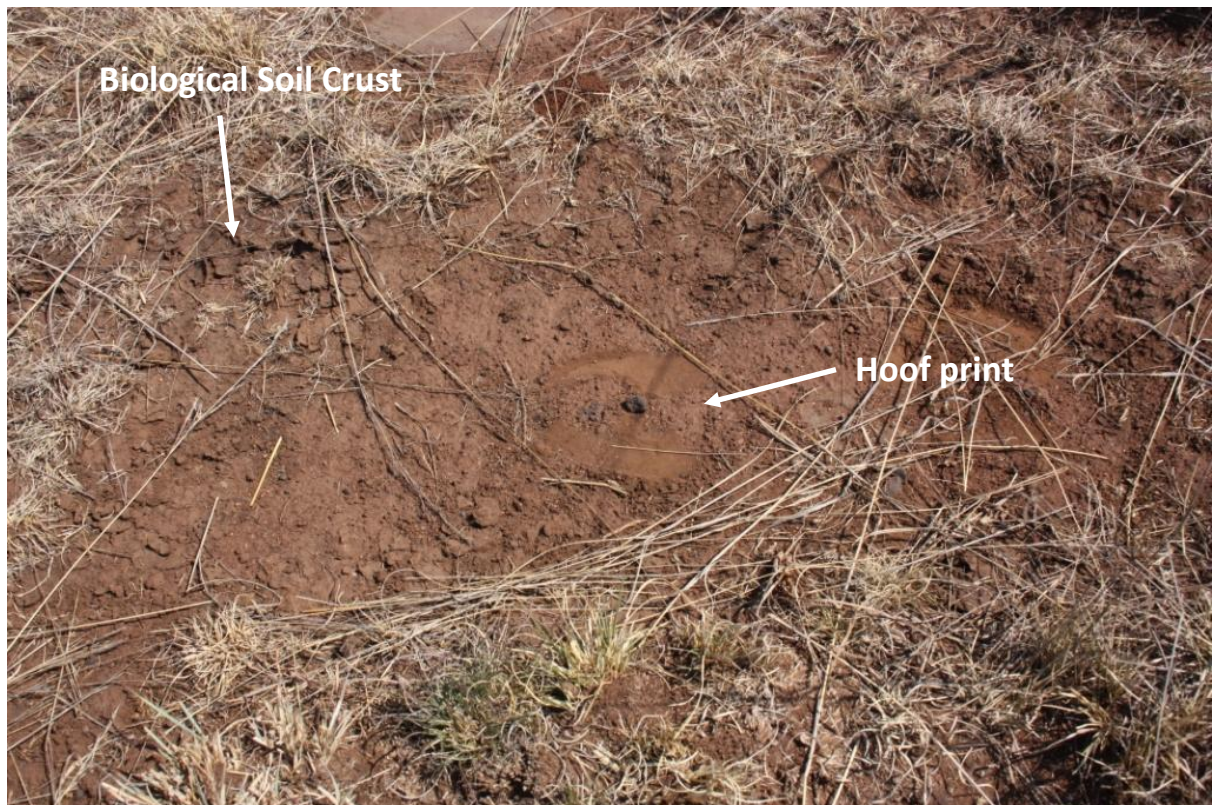
This merging of grass tufts into homogenous patches, together with the fact that these patches (grass and sparse grass patches) far out-numbered other patch types (Table 2.1) and dominated the patch types for each transect (Table 2.2) suggests that these grass patches, as defined in the field, match the “homogenous reference background” of Levin and Paine (1974) and form a matrix in which the other patch types described (Table 2.1) are embedded in a density that is dependent on ecological factors such as edaphic conditions and historical patterns of usage. At West Wits other patch types dispersed in this grassland matrix were bare patches (bp), bare grass (bg), and bankrotbos (bb) or *S. Plumosum* patches while Vaal River included the above patch types together with biological soil crust patches (bsc), tall grass patches (tg) and *Schoenoplectus* patches (sch). Within Watt’s (1947) view of patches is the assumption that a patch type represents a specific stage or “phase” in a successional sequence.

Schoenoplectus is a sedge associated with wetlands (Marnewecke and Kotze, 1999) and was found on only one transect. It therefore represents a response to specific edaphic conditions rather than a response to disturbance or a part of a natural successional sequence. *Seriphium plumosum* L. (syn. *Stoebe vulgaris*) is an indigenous shrub that has been described as an encroacher species in indigenous grasslands where it can seriously impact on the grazing capacity of veld (2012, Snyman, 2010, Jordaan, 2009). Although the drivers for the invasive response of *S. plumosum* are still not clearly defined, Snyman (2012) showed that *S. plumosum* does not like high soil moisture or poor drainage, high soil organic matter and phosphate, soils high in sodium chloride or clay, or alkaline soils. Snyman (2011) further showed that fire stimulates seedling emergence about a month after the fire and was noticeably higher under grass leaves than in open spaces thus dispelling the concept that *S. plumosum* was a pioneer species. Snyman (2010) also showed that *S. plumosum* has allelopathic effects on other species. Thus it is unclear what processes are favouring the presence of *S. plumosum* in these two sites but as it forms the most common patch type at West Wits and the fourth most prevalent patch type at Vaal River (Table 2.1), the soils can be expected to have good drainage and low nutrient status and are probably subjected to regular fires.

Another patch type was tall grass which, from general observation, was probably more common in the general research area than it features in this study as it only occurred on two transects at Vaal River (Table 2.2, 2.3 and 2.5). These tall grass patches (tg) were composed of *Hyparrhenia hirta* (L.) Stapf, an indigenous perennial tussock grass whose culm length is dependent on rainfall (Van Oudtshoorn, 1999), and in various parts of the world the species is regarded as invasive (Chejara *et al.*, 2010). According to van Oudtshoorn (1999), it is a sub-climax or climax species that is resistant to drought, low in palatability, and is regarded as an increaser I species which means that it is abundant in under-utilized rangeland. Fynn *et al.* (2005) describes *H. hirta* as preferring well drained soils and often occurring in abandoned old fields. Fynn *et al.* (2005) further showed that *H. hirta* had strong competitive abilities at low nutrient levels and tended to out-compete shorter grasses such as *Themeda triandra* when grazing was absent. However the data collected in this study could not provide any insight into the presence of these tall-grass patches of *H. hirta* in the study site.

The prevalence of bare patches versus biological soil crust patches seem to be related to land-use factors. At West Wits all open exposed soil surfaces were classified as bare patches whereas at Vaal River only a few such patches were identified and the majority of open exposed soil surfaces were categorised as biological soil crust patches (Table 2.2). However this did not mean that biological soil crust was not observed at West Wits but rather that its presence was low enough that the soil surfaces were characterised as bare patches rather than as biological soil crust patches. A requirement for a biological soil crust to form is a stable soil surface and numerous studies have shown correlations between level of soil surface disturbance and presence of crust (Neff *et al.*, 2005, Belnap and Gillette, 1998, Belnap, 1995). At both Vaal River and West Wits, grazing had officially been terminated (I. Weiersbye, Pers. Comm.). However at West Wits the presence of cattle was observed at all sites and cattle were observed grazing in some sites, whereas at Vaal River only transects 13 and 14 showed some signs of recent grazing. Unfortunately, transect 15 at Vaal River was burnt just as sampling started and this transect was observed to have grazing occurring every afternoon over the sampling period. Exposed soil surfaces on this transect would have been classified as bare patches (Figure 2.12) as biological soil crust was restricted to the periphery of the patch where grass plants provided some protection from trampling and the

1238 rest of the patch shows weakly consolidated soil surface and a fresh hoof print displayed as
1239 freshly powdered soil surface (Belnap, 1995). Thus I believe that the two patch types: bare
1240 patch and soil biological crust patch are a function of grazing pressure, or lack thereof, on
1241 the same basic patch type – exposed soil bare of a vegetation component, and withdrawal
1242 of grazing can quickly promote the reestablishment of biological soil crust across the
1243 majority of the soil surface (Dojani *et al.*, 2011).



1244

1245 **Figure 2.12** A bare patch from Transect 15 just before it was burnt. In the center of the
1246 bare patch is the powdery outline of a fresh hoof print from a cow, and
1247 curled up biological soil crust can be seen around the periphery of the bare
1248 area where it seems to have some protection from the grass bordering the
1249 bare area.

1250 Watt (1947) described a number of examples of patches, or phases as he preferred to call
1251 them, from various different types of ecosystem including two grasslands that formed
1252 continuous covers. Watt (1947) regarded patches as a clustering of individuals of specific
1253 species forming a repeating unit which usually represented a stage or phase in a sequence
1254 or sere. In other words he saw patches as representing a snap-shot of a successional
1255 sequence. Bare patches, biological soil crust patches, bare grass and sparse grass/grass
1256 patches in some cases seemed to follow a continuum that suggests such a secondary

successional pattern. The bare grass patch described patches which had a markedly reduced grass component compared to the sparse grass (sg) or grass patches (gp), but had at least some grass component when compared to bare patches (bp) or biological soil crust patches (bsc), neither of which had any vegetation present in them. Observations in the field suggested that these bare grass patches may have had different causes although neither these differences nor any causes were documented or analysed. Some of these bare patches looked like they were part of a local intermittent drainage line (Figure 2.13a), while other bare patches looked like they might be where some runoff may be trapped and pool for periods of time after a precipitation event (Figure 2.13b). A third bare patch type resembled local bare or biological soil crust patches but with a thin dispersed grass component as if this grass component represented early colonisers (Figure 2.13c) after some form of disturbance had removed the original vegetation cover.

MacDonald (1978) reported a pattern resembling this third type of bare grass patch in Redsoil veld, a grassland near Bulawayo, Zimbabwe. He described a matrix of rangeland in good condition scattered with islands of degraded rangeland. Culm growth of *Cymbopogon plurinodis* was reduced in the degraded patches when compared with culm growth in healthy rangeland, and likewise, soil moisture and infiltration was lower on degraded patches. MacDonald (1978) also found that degraded patches tended to have a different species composition with annual grass species dominating the degraded patches and perennial grass species dominating the healthy rangeland. This pattern corresponds with species differences observed between bare grass (bg) patches and sparse grass (sg) and grass patches (gp) although in this study species were not identified. MacDonald (1978) showed that during a drought year, the degraded patches became bare patches as the annual grasses failed to germinate but stressed that this was a single season response rather than a long-term response. He further showed that during the above-average wet years that followed the drought there was a colonizer/successional sequence of grass species that occurred in these degraded patches with a shift from annual species to medium-lived perennial species (Macdonald, 1978). However during the four years of study, the degraded patches in MacDonald's sites did not achieve the same species composition as the healthy rangeland next to the degraded patches. MacDonald further showed that infiltration was

Intermittent
local
drainage line



(a)

Depression



(b)



(c)

Figure 2.13 The different bare grass patches described in the text. (a) A local intermittent drainage line. (b) A slightly depressed bare grass patch where water may pool for a period of time. (c) A degraded patch where the grass species differ compared to the surrounding matrix and may represent some early successional sequence initiated by some unknown disturbance event.

significantly lower in the degraded sites but could not ascertain whether this was a causal factor or a response to some other factor. He postulated a number of possible causal factors such as (1) overgrazing and trampling, or (2) heavy area-selective grazing, or (3) dead trees which caused bare patches when fire resulted in excessive heat through increased fuel load, or (4) excessive runoff generated by mechanical disturbance such as paths and roads. At Vaal River and at West Wits, there was possible observational evidence in specific cases for all four of these mechanisms but this study could not provide concrete evidence for or against any of them.

2.4.2 Transects Measurements and Transect Indices

Transects in this study were much longer than would normally be the case in an LFA study. Normally, the length of a transect would be based on sampling a representative number of patches, usually five of each patch type, such that a specific landscape is adequately characterised (Tongway and Hindley, 2004). In these grasslands where the basic patch size was generally a grass tussock ($\pm 5 - 20$ cm), this would have resulted in a transect length of at most a few decimetres, and not the 180 metres used to compensate for a 30 m pixel size, and the need to reduce redundancy in sampling (see methods, section 2.2.2). This clumping together of grass tussocks into homogenous swards would have smoothed out the actual patch structure. This is important as LOI, PAI and associated measures (patches per 10 m, mean patch width etc.) are empirical measures indicating how patches are organised in a landscape (Table 2.3), which in turn are proximate indicators of the organising effects of resource availability under a specific set of environmental parameters (Burke *et al.*, 1998, Lauenroth and Coffin, 1992, Noy-Meir, 1973, Worrall, 1960). Therefore measures based on these homogenous swards are at least one level of organisation removed from the direct organisational effects of resource availability and landscape management on patch size and landscape organisation, and may not be an accurate reflection of functionality in these grasslands.

In this context, it is interesting that the three LFA indices evaluating aspects of landscape organisation: LOI, PAI and patches per 10 m show quite different degrees of correlation with the three global LFA functional indices: soil stability, infiltration and nutrient cycling (Figure 2.9). Global LFA indices of function were calculated for the entire transect by calculating an

average of all points measured on the transect weighted by the proportional area for each patch type on the transect. I would have expected, as a generality, that as LOI tends towards 1 (i.e. fewer bare patches and more transect length covered in vegetative patches), or PAI tends towards 1 (i.e. fewer bare patches and more transect area covered in vegetation patches), the three indices of stability, infiltration and nutrient cycling would tend towards higher values (i.e. a positive correlation). The LOI shows no correlation with soil stability and nutrient cycling and a weak negative correlation with infiltration. The patch area index (PAI) has a strong negative correlation with these three indices of function and particularly so with the global LFA index for infiltration. The same pattern holds between global indices of function and measures of patch width albeit a less negative relationship. In short, the more the transect area is covered by patches (PAI tending towards 1), the lower the soil stability, infiltration and nutrient cycling. This suggests that either resource availability (high values of stability, infiltration and nutrient cycling) is not influencing landscape organisation, or the relationship is via some other factor not directly quantified in this study, or that the homogenous swards used as grassland patch definitions are undermining the ability of the LOI and PAI to indicate landscape function.

However, for the patches per 10 m index there is a fairly strong positive correlation with the global infiltration index and less so with the nutrient cycling and stability index. Essentially this is saying that the more patches there are per unit length of transect the higher the infiltration, and slightly less high the soil stability and nutrient cycling. For infiltration this makes sense as the more barriers there are to overland flow, the more time water on the surface has to infiltrate the soil. To fit more patches in per unit length, patches may or may not be getting shorter but the number of bare patches must also be increasing, presumably but not necessarily with a concomitant reduction in the total length and area of bare patch per unit transect. Mayor and Bautista (2012) found that at the finest scales the LFA infiltration index accurately reflected the higher infiltration rate on patches and the lower infiltration rate on bare patches. They also found that at the hillslope scale total run-off was inversely related to the LFA infiltration value of bare patches but not related to the infiltration rate for vegetated patches, or the global infiltration values for hillslopes. However, they did find that global values for LFA infiltration correctly described the total runoff from catchment scale areas. In their study hillslope scales were 16 m² whereas

catchment areas were between 3 and 6 Ha (Mayor and Bautista, 2012) compared to transects in my study which were 180 m long (maximum area 1800 m² or 0.18 Ha). In an earlier study, Bartley *et al.* (2006) also found that bare patch area was positively correlated with runoff from slopes while percent cover had little correlation with runoff. Mayor and Bautista (2012) concluded that scale was an important factor when interpreting and comparing LFA infiltration indices across ecosystems.

Notwithstanding the above discussion, values for LOI, which is the ratio of total patch length to transect length, show much more variability at West Wits than they do at Vaal River. This may be a function of at least two factors. Despite more transects at Vaal River over a wider area, Vaal River transects all fall into a common land use – mostly released from grazing at some unknown time in the recent past (≤ 3 years), and they are all situated in the flattish, undulating plains abutting the Vaal River floodplains where slope angles are small. At West Wits, slope angles were greater (Table 2.3) which increases the erosive effects of runoff, grazing was observed although neither the current nor historical grazing regime has been quantified, and land use was more variable with some sites in rangeland while one transect was an abandoned old field with a well-developed grass cover dominated by *Eragrostis curvula* with some *S. plumosum* shrubs and *H. hirta* tall grasses. Ata Rezaei (2006) found that both severe grazing and increasing slope significantly reduced all LFA indices including LOI. This together with the possible recovery processes through release from grazing pressures and low topographic variability at Vaal River has contributed to the reduced variability of LFA transect measures at Vaal River relative to West Wits.

2.4.3 Patch Indices and other Variables at Patch Scale

The overall trend in ecological measurements for patch types tended to follow the degree of biomass found in a patch (Tables 2.5 and 2.7), although at the high biomass end of the continuum, differences between grass, tall grass, *Schoenoplectus* or *S. plumosum* patches tended to be small or reflect unique conditions. Therefore, *Schoenoplectus* patches at Vaal River, which had significantly higher above-ground biomass together with *S. plumosum* and tall grass patches, when compared with bare grass patches, had the highest mean EC value (Table 2.5). This patch type is defined by the sedge *Schoenoplectus corymbosus* (Roth ex Roem. & Schult.) J. Raynal, which is always associated with wetlands (Marnewecke and

1386 Kotze, 1999), was found on a single transect (vr8, Table 2.4) located close to a tailings
1387 storage facility. This transect recorded a significantly higher mean EC than any other
1388 transect at Vaal River (Table 2.4) with the exception of the transect next to it (vr9). Electrical
1389 conductivity is a measure of the presence of soluble salts in a soil solution (Benton Jones Jr.,
1390 2001) and therefore may be an indicator of contaminated soil. Transect vr8, with the
1391 highest mean EC of $881.5 \pm 178.5 \mu\text{S/cm}$, could be regarded as very slightly to moderately
1392 saline (Benton Jones Jr., 2001) which is considerably lower than values quoted in the Vaal
1393 River Environmental Management Plan (VR EMP) of 1520 to 28782 $\mu\text{S/cm}$ for contaminated
1394 areas (Ellis *et al.*, 2009), some of which were located fairly close (a few 100's of metres) to
1395 transect 8 in the vicinity of the same tailings storage facility (TSF).

1396 *Seriphium plumosum* patches, which also had a high mean EC (Table 2.5), were found on
1397 four contiguous transects (vr 8, 9, 10 and 11, Table 2.4) at Vaal River, two of which had the
1398 highest EC values overall. In total, three transects (vr 3, 8 and 9) had high EC values and
1399 these were also transects that had visible white evaporite crusts on the soil surface. All
1400 three transects were situated on lower slopes and flats and presumably are in close contact
1401 with the water table. Evaluation of soil profiles near vr8 and 9, as documented in the Vaal
1402 River EMP, recorded seepage at 0.6 mbgl (Ellis *et al.*, 2009). Seepage from groundwater and
1403 capillarity together with high evaporation rates at the soil surface and low rainfall resulting
1404 in minimal leaching are believed to cause evaporite crusts or efflorescences (Dultz and
1405 Kühn, 2005). As no chemical analysis was carried out on these evaporite crusts, it is unclear
1406 whether they are the weathered soluble products of dolomite (Ca/MgCO_3) which outcrops
1407 throughout the VR study site, gypsum (CaSO_4) or other sulphate salts, or a product of
1408 contamination from tailings storage facilities and mining activities or a combination of all
1409 three (Meza-Figueroa *et al.*, 2009, Carmona *et al.*, 2008, Dultz and Kühn, 2005, Joeckel *et al.*,
1410 2005, Singer *et al.*, 1999, Keller *et al.*, 1986). Values of soil EC for all sites at West Wits were
1411 very low (Table 2.6), lower than any values for Vaal River recorded in this study, suggesting
1412 that surface soils contaminated through groundwater or atmospheric pathways (Singer *et al.*
1413 *et al.*, 1999) was not an issue at the sites studied at West Wits.

1414 *Seriphium plumosum* patches had the highest values for the LFA indices of soil stability,
1415 infiltration and nutrient cycling at both Vaal River and West Wits (Table 2.5 and 2.7,
1416 respectively), and the highest above-ground biomass, SOM and soil pH at West Wits. This

suggests that *S. plumosum* plays a highly functional role in these grasslands, which according to literature, is contradictory. *S. plumosum* has been considered an indicator of severe overgrazing and rangeland degradation (Snyman, 2009b, Roux, 1969a) or poor management (Jordaan, 2009) as illustrated in one of its common names, bankrupt bush, and its encroachment into Highveld grasslands has been well documented although the mechanism/s behind this encroachment remain unknown (Snyman, 2009a). The shrub is a prolific producer of seeds, some of which seem to have dormancy mechanisms, and even after three years may have high viability (Snyman, 2009a). In contradiction to the conclusions that it invades overgrazed areas and may be a pioneer species, its seedlings only germinate in low-light conditions such as those formed close to unpalatable grasses such as *Cymbopogon* or *Elionurus* species (Snyman, 2010, Lecatsas, 1962, Cohen, 1937). Snyman (2010) provided evidence that *S. plumosum* has allelopathic effects on its own seedlings as well as interspecific competitor seedlings through inhibition of germination and seedling development. This may explain the bare ground that tended to surround *S. plumosum* shrubs on transect ww3 at West Wits (Figure 2.14b). However this was the only site at both Vaal River and West Wits where there was evidence that *S. plumosum* may be exerting a negative effect on surrounding vegetation and no investigation has yet been conducted on allelopathic interactions between *S. plumosum* and adult plants of the same or different species. Furthermore, signs of remnant burnt stubble in these bare patches surrounding *S. plumosum* patches in transect ww3 support Snyman's (2011) conclusion that increased temperatures due to the burning of the *S. plumosum* canopy cause increased local mortality of perennial grass species (Figure 2.14a and b) in the vicinity of these shrubs.

According to Snyman (2012, 2011, 2010), *S. plumosum* is not found on soils that get waterlogged, preferring well-drained soils. Yet forty-three *S. plumosum* patches were recorded on the 180 metres of transect vr8 (Table 2.3). This transect also had the only recorded patches at any site in this study of *Schoenoplectus* species, a sedge that definitely prefers at least seasonally waterlogged or wetland conditions (Marnewecke and Kotze, 1999). Possibly these *S. plumosum* patches are a remnant of drier conditions prior to the establishment of the TSF near the foot of transect vr8. Aerial surveys from 1980 show no TSF but it was present in 1992. Hattingh (1953) estimated that *S. plumosum* had an average lifespan of



a.)



b.)

Figure 2.14 (a) *Seriphium plumosum* burning as a fire swept through transect vr15 showing the increased intensity of the fire as it burns the *S. plumosum* canopy compared to the grass canopy. Notice the grass right up to the *S. Plumosum* suggesting no allelopathic response. (b) Remnant burnt grass stubble and bare patch in the vicinity of an *S. Plumosum* shrub on ww3 transect. The bare patch and dead grass stubble possibly indicative of the increased mortality of grasses close to these shrubs through increased fire intensity when the shrub burns (Snyman, 2011).

about 15 years which is significantly shorter than the time that the TSF has been present (> 20 but < 34 years). However, it is unknown whether or when surface/groundwater conditions might have switched from being possibly well-drained to hydromorphic with the establishment of the TSF. *Seriphium plumosum* shrubs have been associated with low soil nutrient conditions (SOM and % P) and Snyman (2009b) observed that they preferentially invaded abandoned marginal cultivated lands over more fertile sites. Nevertheless, in this study there was no significant difference in SOM between patch types at Vaal River, while at West Wits, SOM values were highest for *S. plumosum* patches and these differences were significant for bare patches but not for other vegetated patch types. Observations while sampling these patches found there was often a well developed fine litter layer of predominantly seeds shed from the shrub. Phosphorus was not examined in this study but no significant differences in soil nitrogen, either organic or inorganic, were found between patch types at West Wits.

Of the seven patch types described from Vaal River, four were some form of graminoid-dominated patch, namely bare grass, sparse grass, grass patch and tall grass, which essentially describes an above-ground biomass gradient (Table 2.5). At West Wits, two of the four described patch types were graminoid-dominated patch types and were similarly named as two of the Vaal River patch types, i.e. sparse grass and grass patch, again based on differences in above-ground cover reflecting a gradient in above-ground biomass (Table 2.1 and 2.7). One factor influencing or influenced by this continuum is soil moisture. However, transect averages for soil moisture tended to follow the date of sampling (not shown) with transects like vr10 and vr11 with the highest average soil moisture being sampled in mid winter whereas transects vr13 and vr14, with the lowest mean soil moisture sampled last towards the end of the dry winter (Table 2.4). But even with the influence of sampling date and winter dry-season effects, above-ground biomass (Figure 2.8) and root biomass (Figure 2.10a) were weakly correlated with soil moisture. Patch type showed significant differences based on soil moisture (Table 2.5 and 2.7) and generally followed the trend of increasing vegetation biomass although there were exceptions. For instance at Vaal River, grass patches had significantly higher above-ground biomass when compared with bare grass and biological soil crust patches, the latter having no measureable above-ground biomass. Yet soil moisture was not significantly different between the three patch types

mentioned, although it was significantly lower than that for *Schoenoplectus* and *S. plumosum* patch types which had the highest mean above-ground biomasses (Table 2.5). As these grass patches belong to transect vr9 and this was sampled towards the end of winter, it is likely that this anomaly in low soil moisture is a result of the length of time the soil has had to dry out over the winter dry period prior to sampling. However, the dominant trend of there being a direct correlation between soil moisture and biomass, and therefore patch type, is one that is well described in the literature (Li *et al.*, 2013, D'odorico *et al.*, 2007, Eldridge *et al.*, 2002, Breshears and Barnes, 1999, Greene, 1992). Bhark and Small (2003) showed that shrub and grassland canopies had higher infiltration and soil moisture than did the inter-canopy spaces and that this effect was stronger for shrubs in their study area than for grass canopies. Pockman and Small (2010) however, found grass patches in the Chihuahuan Desert had higher soil moisture after a rain event than did the shrub *Larrea tridentate* with bare patches the lowest. Thus the relationship between vegetation cover or above-ground biomass and soil moisture is complex.

The primary determinant of soil moisture is climate (Legates *et al.*, 2011, Noy-Meir, 1981). In the United States, Anderson (2006) describes how the central grasslands are separated into three categories which form a continuum from west to east, namely short-grass (0.3 – 0.5 m tall), mixed grass (0.8 – 1.2 m) and tall grass (1.8 – 2.4 m). This continuum essentially reflects an annual precipitation gradient with the arid central continental short-grass prairie in the west receiving 260 – 375 mm annual precipitation, which increases as one moves eastward with the wetter east receiving 625 – 1200 mm (Anderson, 2006), and mixed prairie situated between the two and in the middle of the moisture gradient. It is likely that differences in precipitation underlie differences in biomass and canopy density between the two research sites in this study. At Vaal River 50% of all patches described were sparse grass and this patch type constituted 67% of the graminoid-based patches. At West Wits, 52% of patches were graminoid in nature, split slightly in favour of grass patches (56%) over sparse grass (44%) with no tall grass patches on the transect, but such tall grass patches were observed in off-transect areas of West Wits. Although the composition of the sparse and grass patches was not sampled the only obvious difference was in density of cover as reflected in the above-ground biomass. The fact that this was only significantly higher at West Wits is probably at least partially a function of the higher annual precipitation at West

1519 Wits (Anderson, 2006). Height in these two patch types was not measured but never got
 1520 higher than knee height (0.6 m). However, tall grass patches consisting of *Hyparrhenia hirta*
 1521 tussocks that can reach a height of 1.5 m (Van Wyk and Van Oudtshoorn, 1999), made up
 1522 5% of sampled patches and was limited to particular transects at Vaal River. Although it was
 1523 present at West Wits, tall grass patches were not recorded from any transects sampled
 1524 there. However its presence is not related to a moisture gradient as *H. hirta* is regarded as a
 1525 drought resistant species generally associated with climax or sub-climax grassland
 1526 communities and secondary succession (Van Wyk and Van Oudtshoorn, 1999, Roux, 1969b).

1527 Soil organic matter (SOM) was significantly different between patch types at West Wits
 1528 where bare patches had the lowest mean values and *S. plumosum* patches had the highest
 1529 values (Table 2.7). However at Vaal River there were no significant differences in SOM
 1530 between patch types (Table 2.5). Furthermore, all patches at Vaal River had lower mean
 1531 values for SOM than did any patch type at West Wits although these differences were not
 1532 tested statistically. The reasons for these differences are unknown but the formation and
 1533 breakdown of SOM is a complex process (Jenkinson *et al.*, 1990, Jenkinson and Rayner,
 1534 1977) and a number of factors are known to affect SOM. Temperature (Kätterer *et al.*, 1998,
 1535 Lloyd and Taylor, 1994) and soil moisture (Yang *et al.*, 2007, Lee *et al.*, 2004) have been
 1536 shown to influence SOM where increases in either variable lead to increased soil microbial
 1537 activity and rates of SOM turnover which may result in decreased levels of SOM in soils.
 1538 However, increases in both these factors can also result in increased inputs of carbon to
 1539 soils through increased net primary productivity. Jenkinson and Rayner (1977) showed that
 1540 reduced plant inputs reduced the long-term SOM values of a soil, and sparse grass, grass
 1541 patch and *S. plumosum* patch above-ground mean biomasses were considerably lower at
 1542 Vaal River compared to those for the same patch types respectively at West Wits (Table 2.5
 1543 and 2.7). *Seriphium plumosum* patches at West Wits also had the highest mean above-
 1544 ground biomass for any patch type at West Wits which supports the concept of higher
 1545 carbon in the form of biomass inputs contributing to higher SOM values. Root biomass was
 1546 not quantified at West Wits. At Vaal River, root biomass was found to be only weakly
 1547 correlated with SOM and more strongly correlated with above-ground biomass (Figure
 1548 2.10a). Furthermore, the relationship between root biomass and patch type was not a direct
 1549 relationship between above- and below-ground biomass. This can be seen when comparing

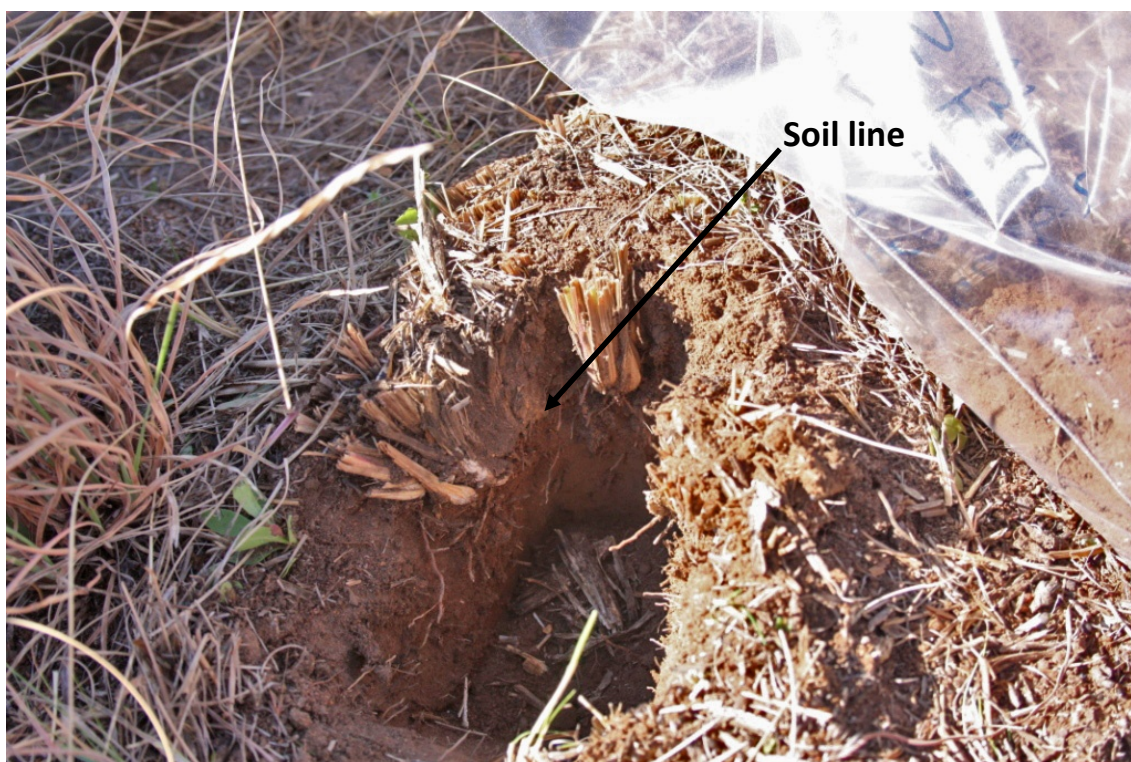
above and below-ground biomasses for sparse grass which had higher mean root biomass and lower above-ground shoot biomass when compared to grass patches at Vaal River. However, the general trend was that increases in above-ground biomass were mirrored in increased root biomass (Table 2.5). The effect of roots on SOM is complex as, on the one hand, they are a source of SOM (Aerts *et al.*, 1992), while Cheng (1990) showed that living roots enhanced the rate of decomposition of SOM. However, amongst other factors, both SOM and the presence of above- and below-ground biomass are important influences on the stability of a soil and its resistance to fluvial based soil erosion (Oades, 1993, Oades, 1984).

The LFA index for soil stability is designed to indicate a site's susceptibility to erosive forces and the ability of a soil to reform after disturbance (Tongway and Hindley, 2004). The index is calculated as a composite of a number of visual or tactile properties of the soil, or features at the soil surface, and uses a relative scale in which high values indicate stable, resistant soils. As the index is a visual system it does not include direct measures of SOM and root presence. Soil organic matter is indirectly estimated through evaluating the degree of incorporation of a leaf litter into soil and the visual presence of fungi in the litter. Root biomass is indirectly estimated through the basal area in grasses, and the canopy cover and density in shrubs and trees (Tongway and Hindley, 2004), on the assumption that the larger the base of a grass tuft is, the greater the root mass that will extend from it into the soil volume. The general trend at both West Wits and Vaal River was that values for LFA soil stability increased from lower values measured for bare soil to highest values for *S. plumosum* patches (Table 2.5 and 2.7). This trend in LFA soil stability was mirrored in most of the environmental factors that promote soil stability and resistance to erosion such as increased above-ground biomass (Chartier *et al.*, 2013, Dlamini *et al.*, 2011, Okin *et al.*, 2006, Abrahams *et al.*, 1995) which was strongly correlated with LFA soil stability at West Wits and less so at Vaal River. Similarly, a weaker but still positive correlation between below-ground biomass and the LFA soil stability index supported the conclusion that this index was providing a viable relative indicator of soil stability and susceptibility to erosion (Burylo *et al.*, 2012, De Baets *et al.*, 2007, Cammeraat *et al.*, 2005).

Other environmental factors measured in this study also provide support for the conclusion that the LFA soil stability index is giving a valid relative indicator of soil stability. Core depth

was taken to 10 cm but in a number of occasions stone prevented the attainment of this depth and there is a statistical correlation between depth of core and patch type. Core depth was significantly shallower in biological soil crust patches and bare grass patches (Table 2.5) compared to all other vegetated patches with the exception of tall grass patches. Also, the amount of stone greater than 2 mm present in the soil in both biological soil crust patches and bare grass patches was significantly higher than in other patch types with the exception of grass, and *Schoenoplectus* patches which were very few and restricted to one transect. There could be a number of explanations for the differences in soil core depths and soil stone content which are unrelated to soil stability and erosion. One such explanation could be that the sub-surface rock is closer to the surface in these patches than it is in neighbouring or nearby patch types and this leads to conditions that are not conducive to the establishment of vegetation and thus a bare patch, biological soil crust patch, or a bare grass patch with very thin annual-like vegetation. However, an alternative explanation would be that deflation is occurring, whereby aeolian erosion in particular, has systematically removed the fines (Cornelis and Gabriels, 2004, Iversen and White, 1982, Bagnold, 1941), leaving behind a soil with increased stone content greater than 2 mm, and a reduced soil depth to the sub-surface stone layer, than neighbouring patch types protected by vegetation. The evidence in this study is insufficient to suggest a mechanism behind these significant differences in core depth and stone content. None the less, these differences do support the conclusion that the LFA stability index is providing a relative and fair indicator of soil stability and susceptibility to erosion.

Other evidence for transport of material comes from observations while collecting above-ground biomass and soil cores. Many grass tufts have a pedestalled nature wherein the soil surface within the tuft is a few millimetres higher than the surrounding soil surface and may be of an observably different constitution and/or colour (Figure 2.15). Again there could be a number of explanations for this soil material within the tuft. It could be the remnants of an old surface protected by the grass tuft when the surrounding soil has been slowly removed. It could also represent material deposited through the grass tufts either trapping saltating material being transported by wind during aeolian erosion (Ravi *et al.*, 2007, Okin *et al.*, 2006, Shachak and Lovett, 1998, Gibbens *et al.*, 1983), or through an accumulation of material in the grass tuft that has been dislodged from the surrounding soil through the



1612 a)



1613 b)

1614

1615 **Figure 2.15** The pedestalled nature of many grass tufts at Vaal River. a) A clipped grass
 1616 tuft showing the definite line between soil and soil material in base of grass
 1617 tuft. b) A pedestalled grass tuft with small stony lag material lying on the soil
 1618 surface to the bottom right of the grass tuft.

action of rain splash from large raindrops (Legout *et al.*, 2005) during the Highveld thunderstorms that commonly occur in the area. The LFA infiltration index was lowest in patch types with the least vegetation such as bare patches and biological soil crust patches at West Wits and Vaal River respectively (Table 2.5 and 2.7) and the general trend was that values increased as the biomass of the different patch types increased. At West Wits above-ground biomass was strongly correlated with the LFA infiltration index (Figure 2.10b) whereas at Vaal River they were less strongly correlated. Also less strongly correlated but still positively correlated with above-ground biomass were core depth, below-ground biomass and SOM from patches sampled at Vaal River. Wilcox *et al.* (1988) found that amongst various edaphic variables in their sites, soil depth was most positively correlated with infiltration where shallow soils had reduced infiltration due to reduced soil water holding capacity. Bhark and Small (2003) showed that infiltration was highest under shrubs in the Chihuahuan Desert of New Mexico, followed by grass patches, and lowest in bare open inter-patch areas between the shrub and grass patches. However, Abrahams *et al.* (1995) found that infiltration was similar or higher in grass patches compared to shrub patches indicating that infiltration is affected by many other factors including edaphic conditions. Wilcox *et al.* (1988) found that vegetation cover and biomass were important attributes in the positive influence of vegetation on infiltration, but they also found that the importance of grass, shrub and litter patches varied between years and canopy cover was more important than basal cover. As mentioned previously, LFA indices use basal cover as a surrogate indicator of below-ground root biomass and it has been shown that root biomass, which was positively correlated with the LFA infiltration index, increases infiltration of water into soils (Williamson *et al.*, 2004, Hino *et al.*, 1987). Finally, Wilcox *et al.* (1988) found a positive correlation between SOM and infiltration in one year but no correlation in the other year of their study. Soil density was most negatively correlated with LFA infiltration (Figure 10a) as high soil density means smaller pores and spaces between soil grains and aggregates and therefore lower infiltration and movement of water through the soil profile. So it can be concluded that the environmental factors measured in this study support the LFA infiltration index as relative indicator of infiltration in these different patch types.

Whereas infiltration is a single physical phenomenon, nutrient cycling is a complex phenomenon involving the movement of many inorganic elements that are essential to the

survival of the organism (Barbour *et al.*, 1987). The pattern for the LFA nutrient cycling index was similar at both West Wits and Vaal River, with mean LFA nutrient cycling values for patch types with no vegetation being significantly lower than patch types with the most vegetation such as *Schoenoplectus* patches, *S. plumosum* and grass patches (Table 2.5 and 2.7). The exception to this general pattern was the few tall grass patches from Vaal River where the mean LFA nutrient cycling value was the second lowest for Vaal River patches and the biomass was the highest. This low LFA nutrient cycling value for tall grass patches was a result of low litter levels with no signs of litter decomposition and incorporation into soil, and little evidence of cryptogams and biological soil crusts around the base of the tall grass tufts. However, the general trend for the LFA nutrient cycling index was a strong correlation with above-ground biomass at West Wits but less so at Vaal River (Figure 2.10b and a, respectively). Likewise, core depth and below-ground root biomass were weakly correlated with the LFA nutrient cycling index which is supported by the findings of Aerts *et al.* (1992) who showed that below-ground root biomass turnover was closely correlated with C, N and P inputs into soil nutrient pools, although in their study, large differences between species existed. Soil bulk density and soil stone content > 2 mm were negatively correlated with the LFA nutrient cycling index, and increased soil density has been shown to be correlated with decreased nutrient availability and uptake (Oliveira *et al.*, 2014, Da Silva and Rosolem, 2003). All these environmental factors provide support for the LFA nutrient cycling index as a relative and general indicator of nutrient cycling.

However, available nitrate-nitrogen and microbial organic-nitrogen extracted from the soil samples collected from West Wits showed no significant differences between patch types and nor did they show any relationship to the LFA nutrient cycling index, or any of the other LFA indices, or above-ground biomass (Table 2.7 and Figure 2.10b). Available ammonium-nitrogen extracted from the soil was significantly higher in grass and *S. plumosum* patches when compared to bare patches but not to sparse grass patches (Table 2.7). This pattern for ammonium was reflected by soil moisture which was significantly higher for grass patches and lowest for bare patches and may reflect the stimulatory effect soil moisture has on microbial activity and therefore ammonification (Gutiérrez *et al.*, 2012). However, it is interesting that nitrate and organic microbial nitrogen did not seem to reflect this moisture effect. Although organic microbial nitrogen showed no significant differences on a patch

basis, there was a significant difference on a transect basis which essentially reflects different locations in the landscape. Transects 3 and 4 had significantly higher organic nitrogen than did transects 1 and 2, and this was mirrored in a similar pattern in the SOM values for these transects. It is not clear why these transects had higher organic microbial nitrogen and SOM values than the other two transects considering that above-ground biomass and the three LFA indicator values showed a reverse pattern. Nevertheless, SOM and organic nitrogen often do show a strong relationship (Mooshammer *et al.*, 2014) as both mineralisation and plant growth, and therefore organic input to soils, depend on a balance between energy in the form of fixed carbon and the ability to transform that energy in the form of enzymes and proteins which have nitrogen as central to their composition.

Values measured in this study for organic nitrogen (overall mean = $30.35 \pm 0.09 \mu\text{g.g}^{-1}$ soil), which is used as a proxy for microbial biomass, compare favourably with other studies. Gallardo and Schlesinger (1992) recorded a mean value of $15.3 \mu\text{g.g}^{-1}$ soil for microbial nitrogen working in the Chuahuahuan Desert which is substantially more arid than the West Wits sites. However, they also noted values for various forests and a savanna, all higher than the mean value obtained at West Wits. Likewise, in another study, values for ammonium and nitrate-nitrogen ($1.2 - 1.7 \mu\text{g.g}^{-1}$ and $0.5 - 2.0 \mu\text{g.g}^{-1}$ soil respectively) compared with values ($1.1 - 1.7 \mu\text{g.g}^{-1}$ and $1.5 - 2.2 \mu\text{g.g}^{-1}$, respectively, Table 2.7) obtained from West Wits sites (Han *et al.*, 2008). The presence of ammonium in the soil would be dependent on the microbial mineralisation of organic nitrogen from mainly either microbial or vegetative sources (Mooshammer *et al.*, 2014). The presence of nitrate would be dependent on nitrifying bacteria converting ammonium first to nitrite, with a second microbial group converting the nitrite to nitrate (De Boer and Kowalchuk, 2001). The mineralized forms of nitrogen are in high demand by both soil microbes and plants (Schimel and Bennett, 2004) and therefore the ratio of organic microbial nitrogen to inorganic nitrogen ($N_{\text{mic}}:N_{\text{ino}}$) can be used as an indicator of microbial activity and mineralisation (D. Drake, pers. comm., 2012). As nitrogen is often the most limiting nutrient in an ecosystem, this ratio can be taken as an indicator of potential for ammonification and nitrification to occur. No significant differences were found between transects or between patch types for the nitrogen cycling ratio index (Table 2.6 and 2.7) which is consistent with the general lack of significant differences for other measures of nitrogen at West Wits.

One possible explanation for the lack of any significant differences in nitrogen values between patch types may be related to where the grasslands of West Wits fit on the precipitation gradient that defines whether the grassland are dominated by below- or above-ground factors (Burke *et al.*, 1998, Lauenroth and Coffin, 1992). It was previously discussed how West Wits falls on the border between below-ground factors and an indeterminate zone where neither below- nor above-ground factors dominant. The below-ground factors are the supply of water and nutrients while the above-ground factor is the availability of light (Lauenroth and Coffin, 1992). Schlesinger and Pilmanis (1998) argued that the scale of patchiness of nutrients in undisturbed desert and semi-desert grasslands is much smaller than in similar sites with invaded shrubs. In essence, they found that available nitrogen was essentially randomly distributed in these grasslands but in shrub-invaded areas nutrient patches formed on a scale of 20 to 260 cm from the shrub base (Schlesinger and Pilmanis, 1998, Schlesinger *et al.*, 1996). Furthermore, Schlesinger *et al.* (1996) found that the area of autocorrelation in soil nutrients was less than or equal to the diameter of the canopies associated with the island. However, West Wits probably receives more precipitation than that received in Schlesinger's sites, which are described as desert sites (Schlesinger and Pilmanis, 1998, Schlesinger *et al.*, 1996), and this would allow individual perennial grass plants to be more closely spaced and thus the island effects at the scale described by Schlesinger *et al.* (1998, 1996) could become a continuum. Standard LFA transect measurements would normally provide both a measure of tuft size and the spacing between adjacent grass tufts, and could have provided some indication for these possible mechanisms. Unfortunately, the modification made to the LFA transect measurements to accommodate 30 m pixels meant the patch size of individual grass tufts was coalesced into homogenous units which are at a larger scale than would be expected under the assumption of nutrient islands related to canopy size.

However, the overriding observation when examining the West Wits ecological data is the separation of the variables into two unrelated groups as can be most clearly seen in the principal component plot (Figure 2.11b) and the correlogram (Figure 2.10b). All the nitrogen-related measurements, together with the SOM measurement, feature strongly on the second principal component. This first principal component was dominated by the three LFA indices of soil stability, infiltration and nutrient cycling, together with above-ground

biomass, and to a lesser extent, soil-solution pH. These two principal components are orthogonal to each other and suggest that there is no relationship between the LFA nutrient cycling index and any measurement of biochemical nutrient cycling such as SOM, ammonium, nitrate or microbial nitrogen. The purpose of taking measurements of inorganic and microbial nitrogen was to provide an empirical test of whether the visually-based LFA nutrient cycling index was reflecting quantifiable parameters. Thus, the fact that there was no relationship between the strongly correlated LFA nutrient cycling index or above-ground vegetation and any measure of nitrogen raises a number of questions around an LFA indicator for nutrient cycling.

Landscape function analysis was developed to provide a simple, rapid system for quantifying, at that moment in time, in a relative and comparable way, the biophysical organisation of a site and the efficiency of ecological processes retaining resources at that site (Tongway and Hindley, 2004). Thus LFA relies on visual and tactile indicators of such ecological phenomena as the distribution of vegetative components relative to each other and the slope, the movement of organic material around the landscape, the nature and state of decomposition of this organic material, the state of the soil surface in terms of both biological and physical crusts, and the composition of soil as determined from texture tests and slaking potential. From these kinds of observations, values are derived as indicative of a site's relative ability to cycle nutrients, the stability of a site's soil from the perspective of its susceptibility to experience erosion, and the degree to which rainfall or water is likely to infiltrate or run-off the soil surface. The LFA nutrient cycling index is defined as "how efficiently organic matter is cycled back into the soil" and is calculated using four environmental variables measured on relative scales adjudicated by the observer (Tongway and Hindley, 2004). These indirect indicators are (i) perennial grass basal cover or tree and shrub canopy cover as an indicator of below ground vegetative biomass; (ii) litter cover, the litter source as in local or transported, and the degree of decomposition in terms of presence or absence of fungi or sign of other microbial activity, and its degree of incorporation into the soil profile; (iii) the degree to which biological soil crusts cover the exposed soil surface; and (iv) the soil surface roughness in terms of its ability to trap and hold or shed mobile material (Tongway and Hindley, 2004). The final index score is

1773 calculated by adding the scores for the various variables and calculating a percentage value
1774 for the site.

1775 It is clear that the characteristics measured in the LFA nutrient cycling index are not the
1776 characteristics of soil nutrient cycling pools, but rather those affecting or defining above-
1777 ground nutrient pools such as the amount of leaf litter, or the propensity of the soil surface
1778 to trap and hold mobile organic material. There is an attempt to measure below-ground
1779 plant contributions through indirect estimation of root pools by evaluating basal stem cover
1780 in grasses and canopy cover in shrubs and trees (Tongway and Hindley, 2004). So in essence,
1781 the LFA nutrient cycling index is combining a relative estimation of two pools – living-
1782 vegetation nutrient pools, and other above-ground organic pools. This other above-ground
1783 organic nutrient pool includes both local and imported leaf litter, animal products and
1784 biological soil crusts. However, these two nutrient pools are not the same as the soil
1785 nutrient pool although all three pools are interconnected. The term soil nutrient pool is
1786 itself a simplification of a highly complex set of pools and interactions that occur in relation
1787 to nutrients within soil volumes. SOM has at least three basic pools: (1) litter, (2) microbial
1788 biomass, and (3) humus (Porporato *et al.*, 2003) and each category can be divided into more
1789 complex pools depending on the nature of the question. There are some critical
1790 components though that governs specific aspects of nutrient cycling. Microbial and other
1791 microfaunal components are critical for the breakdown steps from complex organic
1792 molecules to simpler inorganic molecules, otherwise known as mineralisation
1793 (Mooshammer *et al.*, 2014). Soil moisture is also critical both for the above process and for
1794 uptake by the plant where inorganic minerals must be dissolved in solution (Porporato *et*
1795 *al.*, 2003, Marschner, 1995).

1796 Nutrient cycling is also more than the supply of nitrogen, even though nitrogen is the
1797 nutrient that most often limits productivity in many ecosystems (Lebauer and Treseder,
1798 2008, Galloway *et al.*, 2004) and plays such a critical role as proteins and nucleic acids in
1799 physiological processes. For instance, Craine *et al* (2008) showed that South African
1800 grasslands tended to be co-limited by nitrogen and phosphorus which means that increasing
1801 the supply of nitrogen will not increase the productivity of these grasslands. Redfield (1958)
1802 first described stoichiometric ratios in the elemental composition of plankton in the ocean
1803 on the basis that synthesis of organic matter is highly selective and results in products that

are specific in composition. Although marine environments are quite different to terrestrial environments that lack the mixing processes prevalent in marine ecosystems, an averaged global molar ratio of C:N:P for forest foliage of 1212:28:1 and 3007:45:1 for forest litter has been calculated (Mcgroddy *et al.*, 2004). However, it should be noted that this is a global average, and there is a great deal of difference between regional forests with clear latitudinal gradients (Mcgroddy *et al.*, 2004, Reich *et al.*, 2004). Stoichiometric ratios have also been shown for soil microbial biomass (Cleveland and Liptzin, 2007) and changes in microbial pools in semi-arid grasslands under conditions of elevated CO₂ (Dijkstra *et al.*, 2012). These stoichiometric ratios are produced through constraints imposed by physiology, soil biogeochemistry and community composition (Reich *et al.*, 2004) and illustrate that an index of nutrient cycling needs to consider a combination of elements.

Marschner (1995) discusses 14 nutrients required by plants for healthy functioning and any of these supplied in deficient amounts can retard plant productivity or cause mortality or failure to complete a life cycle. With the exception of carbon, which is fixed from the atmosphere in the leaves, plants obtain all of these nutrients mostly or exclusively in inorganic form from the soil solution via plant roots. With the exception of nitrogen whose ultimate source is the atmosphere, all the others are initially supplied to nutrient cycles through the decomposition of geological material. Six nutrients: N, S, P, Mg, Ca and K are regarded as macronutrients on the basis that plants require them in large amounts, while the rest are regarded as micronutrients as plants usually only require very small or trace amounts (Marschner, 1995). As plants are responsible for the majority of primary productivity in terrestrial ecosystems, i.e. the majority of energy supplied to ecosystems in the form of CO₂ fixed into a variety of organic compounds (Woodwell and Whittaker, 1968), plants are central to the cycling of nutrients in terrestrial ecosystems .

Over and above supplying a large proportion of the energy available in soils to non-autotrophic organisms and microbes through litterfall, exudates of various sorts, and senescence of plant parts, plants also have a major influence over the distribution of nutrients in soils (Jobbágy and Jackson, 2001, Schlesinger *et al.*, 1996). Nutrients in soils may be distributed in horizontal and vertical patterns and these patterns reflect the processes predominating in that local environment. The vertical distribution of nutrients in soils is influenced by weathering, atmospheric deposition, leaching and biological cycling (Jobbágy

and Jackson, 2001), whereas the horizontal distribution of nutrients in soils is predominantly influenced by plants (Burke *et al.*, 1998, Schlesinger *et al.*, 1996, Jackson and Caldwell, 1993, Noy-Meir, 1973) where climate regulates the distribution of plants. The depth at which inputs occur into the profile is regulated through weathering processes and atmospheric deposition. The vertical transport of nutrients is affected by leaching and biological processes which operate in opposite directions with leaching moving nutrients to lower levels while plants in particular move nutrients towards the soil surface.

Jobbágy and Jackson (2001), in support of their hypothesis that nutrients most limiting to plants would be found most consistently in the topsoil, found that total N, extractable P and exchangeable K^+ occurred at higher concentrations in the topsoil than lower down in the soil profile. In contrast nutrients such as Ca^{2+} had an intermediate soil profile distribution, and ions such as exchangeable Mg^{2+} , Na^+ , Cl^- and SO_4^{4-} increased with soil depth (Jobbágy and Jackson, 2001). Plants exert control by concentrating nutrients like N, P and K in plant tissue that were previously dispersed through the soil profile and then regulating the return of those nutrients to nutrient cycling processes through litter fall, exudates and senescence of plant parts (Jobbágy and Jackson, 2001, Jobbágy and Jackson, 2000). This redistribution of nutrients in the soil profile is influenced by both root distribution and plant allocation (Jobbágy and Jackson, 2001). The greater the allocation by plants to shoots, the faster the redistribution of these nutrients to topsoil (Jobbágy and Jackson, 2001). The depth at which roots occur will determine the nutrient pool available for redistribution by plants (Jobbágy and Jackson, 2001).

Under certain circumstances plants may become disconnected from soil and microbial nutrient pools. Abbadie (1992) showed that in the humid savannas of Lamto, Ivory Coast, grasses were obtaining most of their nitrogen from the decay of root material. In these areas, shoot nitrogen was mobilised from the aerial parts prior to seasonal senescence occurring and translocated and stored in the roots of these grasses. Litter and above-ground biomass was then generally lost to annual fires which directly converted the residual organic nitrogen to various inorganic gaseous types and therefore these above-ground plant parts contributed very little to nitrogen-based nutrient cycling. Other sources of nitrogen such as nitrogen fixation, inputs through precipitation and SOM made only slight contributions to plant nitrogen (Abbadie *et al.*, 1992). The end result is that grasses in these savannas are

1866 decoupled from SOM sources and tightly reabsorb and cycle nitrogen mineralized during
1867 root senescence (Abbadie *et al.*, 1992).

1868 To sum up this discussion on nutrient cycling, Abbadie's (1992) observations would explain
1869 the decoupling of plant nutrient pools from below-ground soil microbial nitrogen pools,
1870 available inorganic nitrogen and SOM in this study although the data were not sufficient to
1871 test this conclusion. Craine's (2008) observation that South African grasslands are both
1872 nitrogen and phosphorus limited indicates that the relationship between soil microbial pools
1873 and vegetation nutrient pools observed in my study may be more complicated than is
1874 illustrated by soil nitrogen and above-ground biomass alone. At the same time, the
1875 influence that plants can exert in regulating the distribution of nutrients both vertically and
1876 horizontally is important (Jobbágy and Jackson, 2001, Jobbágy and Jackson, 2000,
1877 Schlesinger and Pilmanis, 1998, Schlesinger *et al.*, 1996). The strong relationship observed in
1878 the principal components analysis between plant biomass and the LFA nutrient cycling index
1879 allows the conclusion that the LFA nutrient cycling index is adequately characterising plant
1880 nutrient pools in a relative manner. However, as the same analysis shows, the LFA nutrient
1881 cycling index is unlikely to be adequately characterising nutrient cycling as a whole in these
1882 patches at West Wits.

1883 2.5 Conclusion

1884 The purpose of this chapter was to test two objectives related to the LFA and other
1885 ecological data collected in the field. These two aspects are: (i) do the LFA techniques with
1886 modified patch definitions adequately characterise the natural landscape organisation, and
1887 (ii) do the three LFA indicators for soil stability, infiltration and nutrient cycling accurately
1888 indicate these processes in a relative manner?

1889 The patch definitions of bare patch, biological soil crust patch, bare grass, sparse grass and
1890 grass patch were shown to possibly represent a successional sequence related to
1891 disturbance of some sort, possibly historical grazing pressure, although the data could not
1892 provide definite proof of these conclusions. Other patch types seemed to be occurring due
1893 to ecological conditions specific to where these patches occurred. *Schoenoplectus* patches
1894 were related to the hydromorphic conditions that periodically occurred along a specific
1895 transect. No evidence in the data could elucidate the factors driving the presence of either
1896 tall grass patches or *S. plumosum* patches although literature does suggest some
1897 relationship to disturbance regimes. In the case of *S. plumosum* early literature implied
1898 disturbance regimes but later literature questions this relationship but the factors driving its
1899 invasion into grasslands are yet to be elucidated.

1900 Transect length was modified from the standard LFA protocol of a length defined by
1901 adequately sampling the patch heterogeneity, usually more or less a decametre, to one
1902 fitting in 5, thirty-metre Hyperion hyperspectral pixels and allowing an extra 30 metres to
1903 ensure that all five pixels fell on the transect. Two patch types, which essentially formed the
1904 matrix of these grasslands in which the other patch types were dispersed, were modified in
1905 their spatial dimensions from tuft size (centimetre dimensions) to homogenous zones
1906 (metre dimensions) due to the unrealistic time demand of measuring tufts across 180 m of
1907 transect. It was argued that this homogenizing did influence the conclusions that could be
1908 drawn about the organisation of the landscape as these matrix measurements now
1909 recorded dimensions at a level in the hierarchy of the landscape which was at least one level
1910 above the actual organisation of the vegetation components which made up the landscape.
1911 Thus the effect was to smooth out the differences between transects and essentially
1912 obscure how landscape organisation, which is based on tuft size and nearest neighbour

relationships, was affected by processes active in the landscape. As the ground-based hyperspectral data is collected from a sampling footprint of 42 cm at ground level, it is not envisioned that this will affect the ground-based modelling of LFA values, but it may affect Hyperion-based hyperspectral modelling where spectral pixel size is approximately 30 m of ground surface. This is because the spectra are expected to record the subtle changes in tuft organisation and spacing through differences in the total area of exposed soil and canopy cover which are lost by classifying these zones into homogeneous units.

The three LFA indices of soil stability, infiltration and nutrient cycling are dependent entirely on the skill and application of the practitioner in accurately judging the processes they describe. It was concluded that for soil stability and infiltration, other data collected during this study together with literature supported the relative accuracy of these LFA indicators. For nutrient cycling, measures of microbial nitrogen and available inorganic nitrogen did not support the LFA index for nutrient cycling. However, above-ground biomass strongly supported the LFA nutrient cycling index and it was concluded that this index was giving a good indication of above-ground nutrient pools but was not giving a good indication of below-ground microbial-related nutrient pools.

1929 3. Predicting LFA from Ground-based Hyperspectral Remote Sensing

1930 3.1.1 Hyperspectral Remote Sensing

1931 Hyperspectral remote sensing, or imaging spectroscopy, is the use of reflected and emitted
1932 light to determine aspects of the environment or feature under study. When light strikes an
1933 object, it is absorbed, transmitted, or reflected off the object. Reflected light is defined as
1934 the ratio of the intensity of light reflected from an object to the intensity of light incident on
1935 the object (Van Der Meer, 2010). When light interacts with a mineral a number of factors
1936 determine whether it is absorbed or reflected. In minerals with transition metal elements,
1937 such as Fe, Ti and Cr amongst others, electron transitions and charge transfer processes
1938 determine the position of absorption features in the VNIR (visible and near-infrared) region,
1939 while vibrational processes in H_2O , OH^- , CO_3^{2-} and SO_4^{2-} produce fundamental overtone
1940 absorption features (Clarke, 1999, Van Der Meer, 2010). The chemical and crystal structure
1941 of the mineral in which the absorbing species is contained influences the position, shape,
1942 depth, width and asymmetry of these absorption features (Clarke, 1999, Van Der Meer,
1943 2010). Some minerals show absorption features due to colour centers caused by defects in
1944 the crystal lattice (Clarke, 1999, Van Der Meer, 2010).

1945 In comparison to minerals, plant tissues have complex interactions with light. The amount of
1946 light reflected, absorbed or transmitted depends on the angle of incidence, wavelength,
1947 surface roughness of the plant part, presence of wax or hairs on the surface, plant
1948 phenology, and differences in optical properties and biochemical contents of plant parts
1949 (Kumar *et al.*, 2010). The presence of plant pigments (chlorophylls, carotenoids and
1950 anthocyanins) play a central role in the spectrum of plants in the visible portion of the
1951 spectrum. Chlorophyll pigments, which are the dominant pigment in active photosynthetic
1952 tissue, absorb visible light for photosynthesis in the violet-blue and red wavelengths of the
1953 spectrum, and reflect green wavelengths around 420 to 490 nm (Kumar *et al.*, 2010), hence
1954 the green colour of active photosynthetic tissue. Light entering a leaf is scattered mainly by
1955 refraction through differences in refractive index between hydrated plant cells and
1956 intercellular airspaces, although some scattering is due to Rayleigh and Mie scattering
1957 (Kumar *et al.*, 2010). Electron transitions are responsible for most light absorbed by plant
1958 pigments. The photon nature of light results in pigments absorbing light at specific

wavelengths and not at all at other wavelengths (Kumar *et al.*, 2010). During senescence chlorophyll degrades more rapidly than other plant pigments and the carotenoids, xanthophylls and anthocyanins begin to dominate the light absorption processes (Sanger, 1971). Thus red wavelength reflectance increases and this together with the green reflectance results in leaves looking orange and yellow. As senescence progresses and dicotyledonous leaves die, tannins predominate and leaves take on a brown colour with reflectance decreasing across the entire visible spectrum (Kumar *et al.*, 2002, Boyer *et al.*, 1988). In monocotyledons such as grasses, it is not clear if they follow the same process. However, dehydration-induced senescence in *Eragrostis curvula* (Schrad.) Nees showed that both chlorophyll and carotenoids decreased simultaneously while anthocyanins increased as dehydration and senescence progressed (Van Der Willigen *et al.*, 2001). A dominant characteristic of light absorption by chlorophyll is the red-edge (Curran *et al.*, 1991, Miller *et al.*, 1990). The red-edge is defined as the point of maximum inflection (Filella and Penuelas, 1994) between the high reflectance in the NIR region caused by leaf internal structure and water content (Kumar *et al.*, 2010), and the strong absorption by chlorophyll at 670 nm (Gitelson and Merzlyak, 1996, Filella and Penuelas, 1994, Gitelson and Merzlyak, 1994).

In the NIR region from 700 to 1300 nm, reflectance from leaves is generally high (Kumar *et al.*, 2010). This is primarily a result of leaf structural components and cellular arrangement, with tightly packed cells such as the palisade mesophyll having higher transmission and lower scattering, and spongy mesophyll with many intercellular airspaces having lower transmission and higher scattering (Kumar *et al.*, 2010). This spectral zone has two water absorption bands at 760 and 970 nm caused by the OH bonds in the water molecule (Peñuelas *et al.*, 1993). The SWIR region from 1300 to 2500 nm, is dominated by water-absorption regions and minor biochemical features in plant reflectance spectra (Kumar *et al.*, 2010). High absorption of light by water in the plant body occurs at 1200, 1450, 1940 and 2500 nm through overtones from transitions in the vibrational and rotational states of water molecules. Foliar biochemical absorption features are caused by lignin, cellulose, starch, proteins and nitrogen but are often masked by plant water content and so are more apparent in dry leaves (Kumar *et al.*, 2010). Lignin is important in an ecosystem context because it is resistant to decomposition processes and therefore contributes to SOC and soil stability (Austin and Ballaré, 2010, Melillo *et al.*, 1982) and is detectable in the spectrum of

senescent vegetation (Serrano *et al.*, 2002, Ustin *et al.*, 1994). As the study site is semi-arid and the sampling period is during the winter senescent period, lignin and cellulose features may be useful to separate senescent standing grass litter from soil spectra (Ustin *et al.*, 1994).

Soil is a highly variable, complex mixture of primary and secondary minerals, water, organic matter, living organisms and air (Ben-Dor *et al.*, 2009). To the eye, the most obvious difference between soils is their colour, whereas in spectral terms the most obvious difference is their brightness (Huete and Escadafal, 1991). Iron oxides are responsible for the yellow, orange and red colours of soils (Anderson and Robbins, 1998). Many mineral constituents of soils such as iron oxides (Viscarra Rossel *et al.*, 2010, Borisova *et al.*, 2007, Sherman and Waite, 1985), carbonates (Lagacherie *et al.*, 2008) and clays (Chabrilat *et al.*, 2002) are detectable from spectra of soil. Iron oxides have absorption centers at 450, 600 and 900 nm (Preissler and Loercher, 1995). In the SWIR region, clays have an absorption feature at 2200 nm (Chabrilat *et al.*, 2002) and carbonates at 2300 nm (Lagacherie *et al.*, 2008). Soil organic matter influences the colour of a soil and causes a general decrease in reflection (brightness) across the entire spectrum from 400 to 2500 nm (Preissler and Loercher, 1995). Increasing soil water content also reduces soil reflectance (Demattê *et al.*, 2006). Other soil properties such as texture, pH and cation exchange capacity (CEC) have been predicted for soils from soil spectra with varying degrees of success (D'acqui *et al.*, 2010, Goldshleger *et al.*, 2010, Terhoeven-Urselmans *et al.*, 2010, Ben-Dor and Banin, 1995, Ben-Dor and Banin, 1994).

3.1.2 Aims and Objectives

The aim of this chapter is to predict four aspects of LFA, namely patch type, and the LFA indices of stability, infiltration and nutrient cycling from ground-based hyperspectral measurements of solar reflectance collected in the field using partial least squares regression (PLSR). As these regressands or response variables are dependent on the practitioner's skill in identifying patch types and allocating values for soil-surface properties, the same spectral data was explored with PLSR modelling to predict other, more empirical variables of these patch types and soils. The objectives of this chapter are therefore:

- | | | |
|------|----|---|
| 2019 | 1. | to test the applicability of HRS and PLSR to predict patch type from ground-based |
| 2020 | | hyperspectral data in each of two sites, West Wits and Vaal River. |
| 2021 | 2. | to test the applicability of HRS and PLSR to predict the three LFA indices of soil |
| 2022 | | stability, infiltration and nutrient cycling with the same hyperspectral data. |
| 2023 | 3. | to explore any relationships between the same hyperspectral reflectance and |
| 2024 | | empirical variables such as above-ground biomass, below-ground biomass, soil |
| 2025 | | moisture, soil organic matter, inorganic and organic soil nitrogen content, soil core |
| 2026 | | depth and soil properties such as EC, pH, soil bulk density, and gravel and stone |
| 2027 | | content. |

2028 3.2 **Methods and Materials**

2029 3.2.1 **Introduction**

2030 The research site, layout of transects, selection of sampling points, and variables sampled
2031 are described in Chapter 2, Section 2. Spectral reflectance was collected at every point
2032 where LFA soil surface indices were determined, and above-ground biomass and soil cores
2033 collected.

2034 3.2.2 **Spectral Sampling in the Field**

2035 Spectral samples were collected in the field with a portable Fieldspec Pro FR
2036 spectroradiometer made by Analytical Spectral Devices (ASD) and carried in a backpack with
2037 a laptop mounted on a belly board. The ASD was fitted with a fibre optic cable and a pistol
2038 grip held at hand height approximately 1 m from, and perpendicular to, the soil surface. The
2039 ASD collects spectral data from 350 nm to 2500 nm and a single saved spectrum of a target
2040 was the average of sixty spectra integrated to values at 1 nm intervals using the High
2041 Contrast RS³ software supplied with the instrument. Spectral sampling occurred between
2042 12:00 and 13:00 pm when the sun was at zenith and no clouds were visible. Sampling was
2043 restricted to one transect at a time which was sampled in a sequential manner with spectral
2044 measurements collected after the LFA transects were laid out and patches identified, but
2045 prior to collection of soil surface values and destructive sampling for above-ground biomass
2046 and soil cores. Three spectra were collected in succession from each sample point and each
2047 spectrum was only saved once the reflectance sampled by the ASD had stabilized. The
2048 reflectance was standardized using a 99% Spectralon Labsphere reference plate before
2049 sampling each point.

2050 3.2.3 **Spectral Preparation**

2051 Spectra were processed with ASD's Viewspec Pro software and exported as text files
2052 containing reflectance. Each target's three spectra were averaged after individual inspection
2053 to produce single spectra for each sample point. After inspection and averaging of spectra,
2054 some wavelengths were deleted from all spectra as many spectra had spectral noise at
2055 specific wavelengths. These deletions included wavelengths at the beginning and end of

2056 **Table 3.1** Wavelengths removed from the spectra prior to PLSR modelling
 2057

Wavelengths (nm)	Possible source of noise
350 – 399	ASD radiospectrometer sensitivity
911 – 969	Atmospheric H ₂ O
1111 – 1169	Atmospheric H ₂ O
1331 – 1459	Atmospheric H ₂ O
1781 – 1949	Atmospheric H ₂ O
2401 – 2500	ASD radiospectrometer sensitivity

2058

2059 each spectrum due to SNR at the edge of the detector, and noise presumably caused by
 2060 atmospheric gases, aerosols and water (Table 3.1). The spectral data were then imported
 2061 into R (R Core Team, 2013) and combined with the ecological or patch data into a data
 2062 frame prior to regression analysis using the pls package (Mevik and Wehrens, 2007) in R.

2063 3.2.4 The Ecological Data

2064 The ecological data were discussed in detail in chapter 2. They were collected as point
 2065 samples distributed along transects. Each point was classified according to the patch type in
 2066 which it was situated, and all patch types on a transect were sampled. At each point, spectra
 2067 were collected as described in Section 3.2.2, fresh and dry above-ground biomass
 2068 measured, below-ground biomass and stone and gravel were collected while sieving the soil
 2069 samples through a 2 mm sieve and weighed fresh and dry. Soil moisture content, soil pH,
 2070 soil EC, SOM and in the case of the West Wits samples, soil microbial organic nitrogen and
 2071 KCl-extractable soil ammonium and nitrate measured. The depth of penetration of the soil
 2072 core to a maximum of 10 cm was measured and soil dry bulk density calculated. All these
 2073 ecological measurements could thus be paired with a spectral sample and were collected for
 2074 the whole transect immediately after the spectral data for the transect had been acquired.

2075 3.2.5 Partial Least Squares Regression Modelling

2076 Partial Least-Squares Regression (PLSR) is a data compression method where the spectral
 2077 data is the predictor matrix and the ecological data is the response matrix. The compression
 2078 produces a sequence of models with increasing number of components. These models allow
 2079 the prediction of a response variable from spectral data (Frank and Friedman, 1993). The
 2080 method was developed by Wold in the sixties for econometrical problems and has become a
 2081 common technique for chemometrical and spectroscopy modelling (Wold *et al.*, 2001,

Geladi and Kowalski, 1986a). PLSR is particularly useful for spectral data which is generally highly collinear and often noisy (Wold *et al.*, 2001). The method is unusual in that it models both the X and Y-variables through maximizing the explained X/Y covariance (Martens, 2001), and has the flexibility of being able to incorporate a number of response or Y-variables into the modelling process (Martens and Næs, 1989). This means that the LFA variables of soil stability, infiltration and nutrient cycling can be modelled individually against the spectral data, or in the case of patch types, as a matrix of dummy variables in which each column represents a particular patch type.

In essence PLSR uses matrix algebra and a special case of singular value decomposition (SVD) which incorporates both the spectral matrix and the response matrix (Mark and Workman, 2007a). SVD results in three matrices, termed the left singular values matrix, the singular values matrix and the right singular values matrix. The right singular values matrix, also known as the loadings matrix, is used to calculate the scores matrix (Mark and Workman, 2007a). From all these matrices, regression coefficients and predicted values of the response variable can be calculated. However, the process is iterative up to the number of objects in the dataset. Each iteration adds a component or latent vector to the model predicting the response variable. Thus we can use a model constructed from only the first component, or a model constructed from the number of components equal to the number of objects in the dataset, or any number of components in between. However, the first few components generally account for the features in the environment present in the spectra, but thereafter, adding components to the model begins to include more of the noise or error inherent in the data (Mark and Workman, 2007b). At this point it is important to note that the number of components in a PLSR model is not related to the number of variables in the data. So if a PLSR model is derived from spectral data with 200 wavelengths and infiltration as the only response variable, the best model for predicting infiltration might be a single-component model or a three- or 10-component model but regardless of the number of components, all 200 wavelengths would be included in the model.

Selecting the optimal model is not a simple exercise and Mevik and Wold (2007) state that there are no tests for model selection. They further suggest an indication of the best model is when the RMSEP or cross-validation error reaches a minimum value. To avoid over-fitting of the model and keeping the number of components to a minimum, Kooistra *et al.* (2004)

2113 advocated a parsimonious of only adding a component to the model if its addition results in
2114 a reduction in the RMSEP of greater than or equal to 2%. This is a strong and simple rule to
2115 apply when there is a single response. However, for models with a matrix of predictors as in
2116 the patch models, the addition of a component may have a marked reduction in RMSEP for
2117 one response variable but actually increase the RMSEP for another variable.

2118 Wilson *et al.* (2004) used van der Voet's (1994) test to select the model with the least
2119 number of components whose predicted residual sum of squares (PRESS) is insignificantly
2120 larger than the PRESS for the model with the minimum PRESS. PRESS is related to RMSEP via
2121 the MSEP or mean sum of squares which is the mean of the PRESS, and the RMSEP is the
2122 square root of the MSEP. The advantage of the RMSEP is that the units are the same as the
2123 response variable. Many other authors have used the RPD or the ratio of performance to
2124 deviation which was first proposed by Williams (1987, in Minasny and Mcbratney, 2013) to
2125 select the best model (Nawar *et al.*, 2014, Ong and Cudahy, 2014, Wetterlind *et al.*, 2010).
2126 However, it has also been criticized as nothing more than the coefficient of determination
2127 (r^2) transformed to a scale where values above 2 are regarded as being excellent models,
2128 values between 1.4 and 2 as fair models, and values less than 1.4 are regarded as being non-
2129 reliable models (Bellon-Maurel *et al.*, 2010, Chang *et al.*, 2001) although other authors have
2130 used other ranges (Bellon-Maurel *et al.*, 2010). Both Bellon-Maurel *et al.* (2010) and
2131 Minasny and Mcbratney (2013) criticized the RPD on the basis that it depends on samples
2132 being normally distributed (Bellon-Maurel *et al.*, 2010) and that the RPD is the same as r^2
2133 (Minasny and Mcbratney, 2013), which has a more intuitive range of values. Bellon-Maurel
2134 *et al.* (2010) recommend using the RPIQ or ratio of performance to interquartile range, which
2135 better represents the spread of the population, rather than the RPD. Neither RPD nor RPIQ
2136 were used to select the best model in this study due to the above critique of the RPD and
2137 the fact that no one else has used the RQIP.

2138 To summarize the above discussion on model selection, choosing the best model needs to
2139 find a balance between the variance in the data that the model can account for, the amount
2140 of noise being included in the model by adding more components to the model, and the
2141 lowest statistical prediction error in the models. Therefore, in this study with ground-based
2142 spectra as the predictor matrix, optimal models were chosen based on a balance between
2143 local minima for the RMSEP, the number of components in the model, and the validation

2144 coefficient of determination or r^2 value as an indicator of amount of variability explained by
2145 the model for both the spectral data and the response variable (Mevik and Wehrens, 2007).

2146 3.3 **Results**

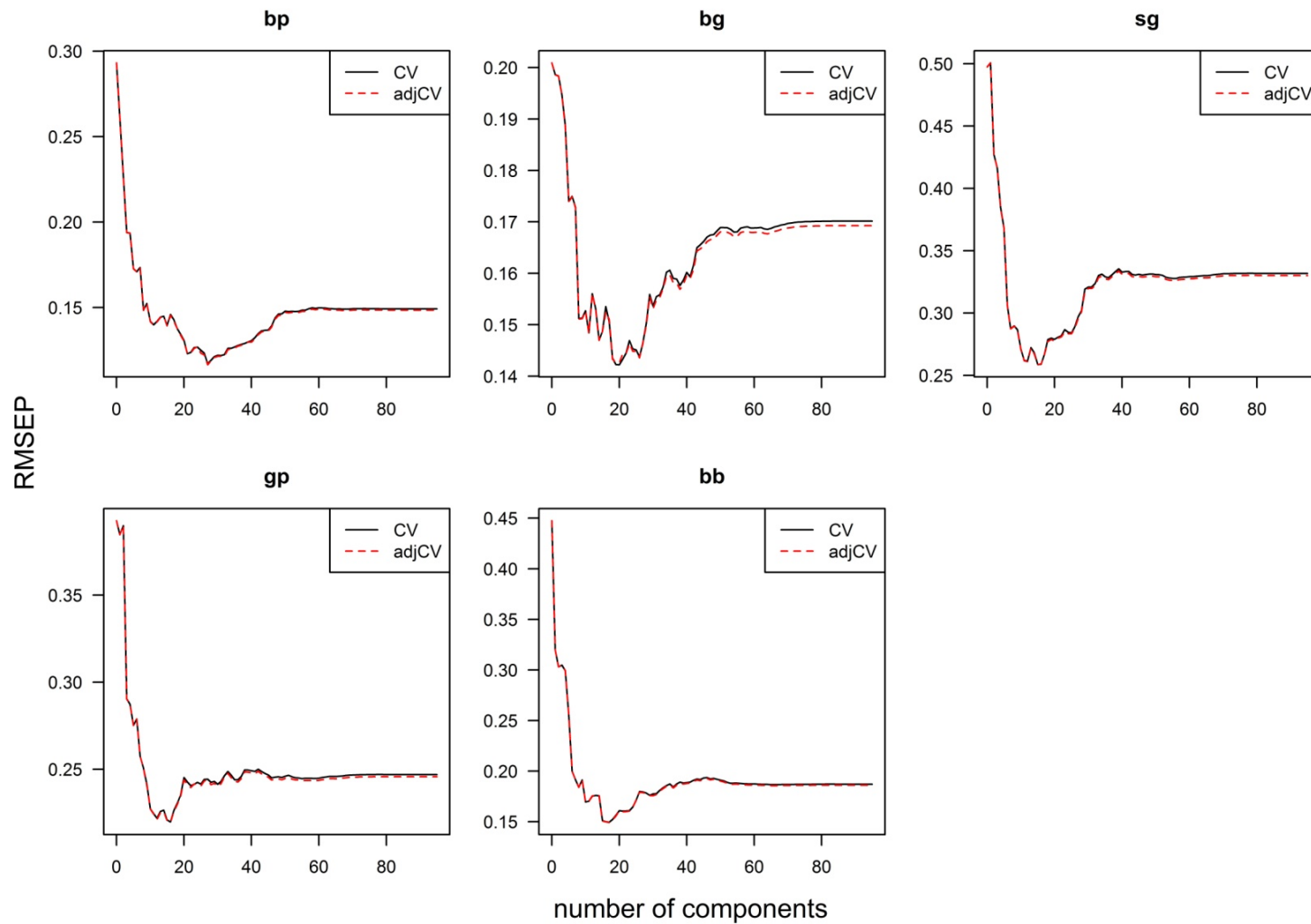
2147 3.3.1 **Patch Models for West Wits**

2148 West Wits patches provide a good introduction to the PLSR models as the site had fewer
2149 patch types than Vaal River, and other variables such as the LFA indices, biomass and soil
2150 properties are often related to the patch type (Aguilar and Sala, 1999, Burke *et al.*, 1998).
2151 Furthermore, having a matrix as the dependent Y variable, in this case a matrix of five patch
2152 types, provides a more complex test of the PLSR method than would be the case when
2153 predicting a single variable such as biomass. Table 3.2 presents the initial results for the first
2154 20 components of a PLSR model derived from calibration data consisting of 97 sample
2155 points in five patch types. The best model should be able to separate the various patch
2156 types from the rest by indicating a value where any values above that would indicate the
2157 specific patch type and values below would indicate the rest of the patch types. There are
2158 no definitive tests for selecting the optimum number of components for a model which
2159 makes this decision to some extent a subjective one (Mevik and Wehrens, 2007). The lowest
2160 Root Mean Square Error of Prediction (RMSEP) derived from calibration data may indicate a
2161 good model. In practice, to avoid over-fitting of the model, one often chooses the number
2162 of components that correspond with the first local RMSEP minimum (Mevik and Wehrens,
2163 2007) rather than the absolute minimum.

2164 From the RMSEP calculated using Leave-One-Out (LOO) cross-validation, it is clear that no
2165 one model is optimal for all five patch types (Table 3.2 and Figure 3.1). The first RMSEP local
2166 minimum is for a 2-component model predicting *S. plumosum* patches (Table 3.2). However,
2167 coefficients of determination for predictions with calibration data for the other four patch
2168 types with a 2-component model are all lower than 50% and that for *S. plumosum* is only
2169 55.8%. The first model that would include a local RMSEP minimum for all five patch types is
2170 a 7-component model but this gives a calibration prediction r^2 of 40.45 for bare grass
2171 patches which suggests we would have very poor separation of this patch type from other
2172 patch types. The lowest RMSEP was obtained for bare patches (Figure 3.1), but this model
2173 had the highest number of components (27) for an absolute RMSEP minimum for any patch

2174 **Table 3.2** Cross-validated (CV) RMSEP using leave-one-out (LOO) estimates for calibration models predicting patch types at West Wits from ground-
2175 based hyperspectral measurements. The table also presents the coefficient of determination (r^2) representing the explained variance for
2176 each variable as one adds components to the model. Values in bold underlined italics reflect models with local RMSEP minima for that
2177 patch type, while the bold underlined values reflect the variation explained for that variable by three selected models using calibration
2178 data. Patch acronyms are: bp = bare patch, bg = bare grass, sg = sparse grass, gp = grass patch, and bb = *S. plumosum* patches.

Components	Cross-validated RMSEP using LOO					Coefficients of determination for calibration predictions					
	bp	bg	sg	gp	bb	Spectral	bp	bg	sg	gp	bb
1	0.261	0.1986	0.5007	0.3848	0.3199	90.75	24.43	4.87	0.48	5.56	49.96
2	0.2268	0.1984	0.4272	0.39	<u>0.3031</u>	96.23	45.69	6.18	28.88	5.57	55.81
3	0.1939	0.1947	0.4164	0.2904	0.3047	98.55	59.99	15.46	33.42	48.75	56.01
4	0.1935	0.1887	0.3848	0.2873	0.2994	99.17	63.50	31.09	47.90	51.29	61.23
5	0.1727	<u>0.174</u>	0.3683	<u>0.2752</u>	0.2558	99.80	69.98	36.98	50.53	55.59	70.50
6	<u>0.1709</u>	0.175	0.3062	0.2789	0.2000	99.85	71.24	37.03	68.51	55.67	83.61
7	0.1734	0.1729	<u>0.2875</u>	0.2575	0.1916	99.91	71.38	40.45	73.15	64.96	85.45
8	<u>0.1484</u>	<u>0.1512</u>	0.2897	0.2506	<u>0.1841</u>	<u>99.93</u>	<u>80.61</u>	<u>55.84</u>	<u>73.42</u>	<u>68.37</u>	<u>87.07</u>
9	0.1523	0.1513	0.2866	0.2414	0.1913	99.94	80.77	56.87	78.47	76.22	87.13
10	0.1419	0.1527	0.2711	0.2273	<u>0.1695</u>	99.95	84.91	57.98	80.73	77.44	90.29
11	<u>0.1398</u>	<u>0.1484</u>	0.262	0.2243	0.1703	<u>99.96</u>	<u>86.42</u>	<u>60.54</u>	<u>81.79</u>	<u>78.27</u>	<u>90.46</u>
12	0.1418	0.156	<u>0.2615</u>	<u>0.2219</u>	0.1752	99.97	86.67	61.51	83.10	80.95	90.54
13	0.1444	0.1531	0.2723	0.2257	0.1761	99.97	86.91	62.41	86.06	83.75	90.83
14	0.1449	<u>0.147</u>	0.2681	0.2264	0.1757	99.98	87.23	63.96	87.99	85.19	91.02
15	<u>0.1395</u>	0.1487	<u>0.2587</u>	0.2208	0.1511	99.98	88.43	64.02	88.07	86.01	94.01
16	0.1459	0.1535	0.2589	<u>0.2197</u>	0.1501	99.98	89.19	67.90	89.48	88.48	94.07
17	0.1429	0.1507	0.267	0.2261	<u>0.1495</u>	99.99	90.35	67.93	90.85	90.00	94.39
18	0.1376	0.1435	0.2787	0.2305	0.1521	<u>99.99</u>	<u>91.48</u>	<u>74.97</u>	<u>92.63</u>	<u>90.57</u>	<u>94.50</u>
19	0.1343	0.1423	0.28	0.2352	0.1563	99.99	92.90	74.99	94.54	90.61	95.50
20	0.1308	<u>0.1422</u>	0.2789	0.2452	0.1609	99.99	93.55	75.78	95.18	93.73	95.78



2179

2180 **Figure 3.1** Cross-validated RMSEP using LOO for all components of a PLSR calibration model with five patch types and ground-based
 2181 hyperspectral measurements collected from West Wits. Patch acronyms are: bp = bare patch, bg = bare grass, sg = sparse grass,
 2182 gp = grass patch, and bb = *S. plumosum* patches.

type. It would be preferred if the chosen model was as parsimonious as possible with the number of components in the model. The patch type with the least number of components at its RMSEP minimum was sparse grass patches with a RMSEP of 0.2587 for a 15-component model (Table 3.2 and Figure 3.1). However, this was also the highest RMSEP value for any patch type at its absolute minimum RMSEP, with both bare grass and *S. plumosum* patches having nearly half the RMSEP value (0.147 and 0.1495, respectively) but slightly more components (14 and 17, respectively) in the model than did sparse grass patches (Table 3.2 and Figure 3.1). Thus it can be clearly seen that no one model is able to predict the five patch types with similar error.

Following the rule of including a component in the model only if it contributes a greater than 2% reduction in RMSEP (Cho *et al.*, 2007, Kooistra *et al.*, 2004), did not simplify model selection, as the addition of components had highly variable impacts on the decrease in RMSEP for different patch types (results not shown). For instance, the first three components reduced the RMSEP for bare patches by 11%, 13% and 14.5% respectively. Whereas for bare grass patches, the reduction in RMSEP for these three components was 1.1, 0.1 and 1.9% respectively, and that for grass patches were 2, -1.4 and 25.5% respectively. An examination of the coefficients of determination (r^2), which indicate the degree to which that number of components in the calibration model account for the variance in a specific variable in the model, show that the first component accounts for 90.75% of the spectral variation (Table 3.2) while the same component accounted for less than 1% of the variability in the sparse grass patches and a maximum of 49.96% for *S. plumosum* patches, with the other three patch types falling between these two extremes. Of the five patch types in the model for West Wits, bare grass patches consistently have the lowest explained variance per added component, with the exception of the first two components (Table 3.2). In fact, the calibration predictions for bare grass patches only pass the 50% explained variance when the 8th component is added to the model. This is substantially more components than the other patch types where *S. plumosum* patches require only two components to achieve at least 50% explained variance, bare grass patches three, grass patches four, and sparse grass patches five components (Table 3.2). Taking into account both the RMSEP and weakness of the model in accounting for the variability in the

2213 bare grass patches, three models were selected for testing against the validation data.
2214 These were models with 8, 11 and 18 components.

2215 The 8-component model was chosen because it was the second local RMSEP minima for
2216 three patch types, namely bare patches, bare grass patches and *S. plumosum* patches (Table
2217 3.2 and Figure 3.1). The other two patch types, sparse grass and grass patches, had one local
2218 RMSEP minima in an 8-component model and this model explained more of the variability in
2219 the bare grass data (55%; Table 3.2 and Figure 3.1) than was left in the residuals for bare
2220 grass patches. The 11-component model was chosen because it matched the third local
2221 RMSEP minima for bare grass and bare patches, and included the third local minima (10th
2222 component) for *S. plumosum* patches (Table 3.2 and Figure 3.1). The 11-component model
2223 was also one component short of the second local RMSEP minima for sparse grass and grass
2224 patches. This twelfth component did not add substantially to the explained variance for any
2225 of the five patch types, leaving the 11-component model as the more parsimonious choice.
2226 The third model chosen to test against calibration data was an 18-component model. This
2227 model included the absolute minima for RMSEP values for three patch types: sparse grass,
2228 grass patches and *S. plumosum* patches with 15, 16 and 17 components respectively (Table
2229 3.2 and Figure 3.1). It also coincided with a decrease in the RMSEP for bare grass patches of
2230 greater than 2% (4.8%) between the 17th and 18th component, although this was not a local
2231 minima in RMSEP. The nearest local minima for bare grass patches was 20 components but
2232 these two extra components did not significantly improve the r^2 for bare grass patches
2233 when compared with the variation explained by an 18-component model (Table 3.2).

2234 Table 3.3 presents the results of predictions made with these three selected models for
2235 both calibration and validation data. First note that bare grass patches have very few
2236 objects in the model: four in the calibration data and one in the validation data. This single
2237 object in the validation data probably accounts to some degree for the fact that for bare
2238 grass patches the validation data predictions had higher coefficients of determination than
2239 for the calibration data with both 8- and 11 component models (Table 3.3). Bare grass
2240 patches also consistently under-predicted values with both calibration and validation data
2241 (Figure 3.2), had the greatest deviation from a perfect 1:1 match between predicted versus
2242 measured values (dotted line = 1:1, solid line = best fit between measured and predicted
2243 values) for any patch type from West Wits (Figure 3.2). Like bare grass patches, sparse grass

Table 3.3 Number of objects in calibration and validation datasets for five patch types from West Wits. The models selected include 8-, 11- and 18-component models and show the coefficient of determination and probability when predicting values from calibration and validation data.

Model	Patch name	Bare patch	Bare grass	Sparse grass	Grass patch	S. Plumosum	Total
	Patch code	bp	bg	sg	gp	bb	
	Calibration objects	9	4	41	17	26	97
	Validation objects	3	1	12	7	9	32
8 component model	Calibration prediction (r^2)	0.81	0.56	0.73	0.68	0.87	
	Probability	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	
	Validation prediction (r^2)	0.76	0.67	0.77	0.68	0.84	
	Probability	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	
11 component model	Calibration prediction (r^2)	0.86	0.61	0.82	0.78	0.90	
	Probability	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	
	Validation prediction (r^2)	0.80	0.66	0.82	0.68	0.89	
	Probability	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	
18 Component model	Calibration prediction (r^2)	0.91	0.75	0.92	0.91	0.95	
	Probability	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	
	Validation prediction (r^2)	0.89	0.75	0.84	0.73	0.91	
	Probability	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	

patches had a higher r^2 value for predictions with validation data ($r^2 = 0.77$, $P < 0.0001$) compared to calibration data ($r^2 = 0.73$, $P < 0.0001$) with the 8-component model, but this converged to the same value for the 11-component model and both data sets ($r^2 = 0.82$, $P < 0.0001$ in both cases), and had fallen below the calibration r^2 value ($r^2 = 0.92$, $P < 0.0001$ and for validation: $r^2 = 0.84$, $P < 0.0001$) in the 18-component model (Table 3.3). Grass patches had the same r^2 value (0.68, $P < 0.0001$) for the 8-component model with both calibration and validation data and this value did not change for the 11-component model and validation data ($r^2 = 0.68$, $P < 0.0001$), but did increase for the 18-component model with

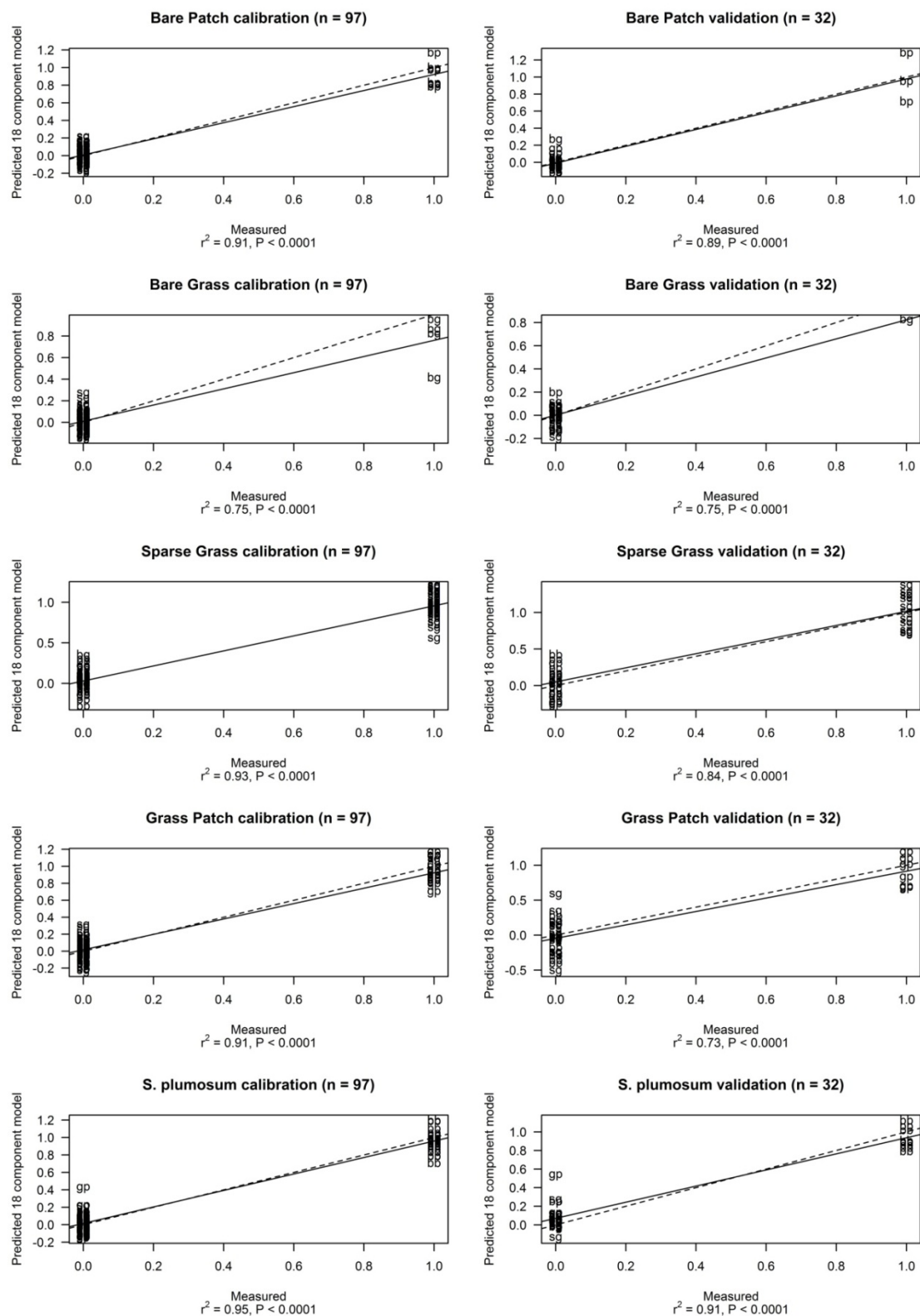


Figure 3.2 Calibration (left column) and validation (right column) predictions from an 18-component model for five patch types at West Wits. The dotted line represents the 1:1 fit between predicted and measured values while the solid line represents the best simple regression fit between predicted and measured values.

both calibration and validation data ($r^2 = 0.91$, $P < 0.0001$ and $r^2 = 0.73$, $P < 0.0001$, Table 3.3). The expected trend is that model predictions would be slightly weaker (lower r^2 values) with validation data compared to calibration data with the same number of components in the model, and that as components are added to the models, predicted values would more closely relate to measured values (higher r^2 values) until an optimum model is selected. This pattern shows clearly in bare patches, and in *S. plumosum* patches the optimum model is probably between 11- and 18-component models (Table 3.3).

Another aspect observable in the predictions from the three models is that bare patches and *S. plumosum* patches had high coefficients of determination ($> 75\%$) for predicted versus measured values from all three models using both calibration and validation data (Table 3.3). This suggests that PLSR is finding it relatively easy to “see” spectra for bare patches (a soil surface) and *S. plumosum* patches (an evergreen shrub) as different (Figure 3.3). This is supported by the fact that the first component explains 50% of the variation in calibration predictions for *S. plumosum* patches, while the first two components explain 46% of the variation in calibration predictions for bare patches (Table 3.2). However, bare grass and grass patches had the lowest r^2 values for all patches when comparing predicted versus measured values with the 8-component model and validation data. Adding another 3 components, while it improved the calibration r^2 values with the 11-component model, actually resulted in a drop in r^2 for bare grass patches and no change in r^2 value for grass patches with validation data (Table 3.3). This suggests that the PLSR model is not separating out the spectra for bare grass and grass patches easily from the spectra for other patches. Sparse grass patches make up more than a third of the objects in both the calibration and validation data and yet only after adding 13 components to the model, do the coefficients of determination for sparse grass reach similar or higher calibration r^2 values than those observed for bare patches for the same number of components in the model (Table 3.2). Also the RMSEP for both sparse grass and grass patches are very similar and considerably higher than values for the other three patch types (Table 3.2).

The implications from the above results are that the model is finding it difficult to spectrally separate bare grass, sparse grass and grass patches where the vegetation is mainly grass species, all of which are in a winter-senesced state. From a physiognomic and species composition perspective, sparse grass and grass patches represent different ends of the

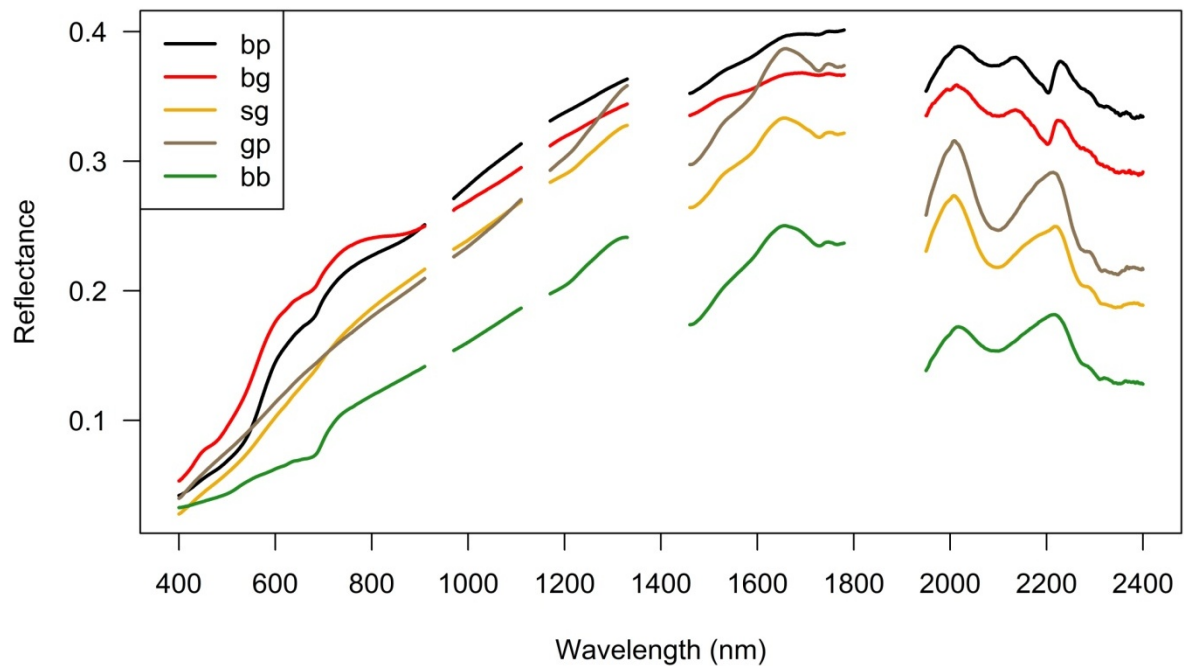
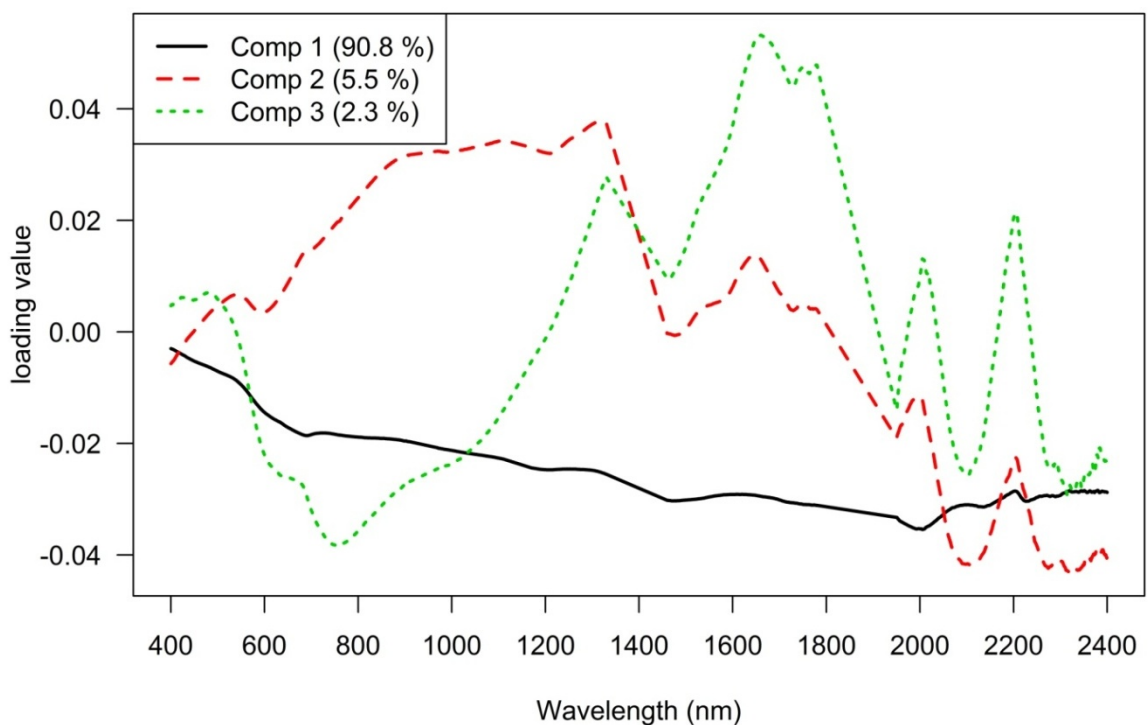


Figure 3.3 The average spectrum for each patch type from West Wits. Patch acronyms in the legend: bp = bare patch, bg = bare grass, sg = sparse grass, gp = grass patch, and bb = *S. plumosum* patches.

same continuum whereas bare grass patches may represent a disturbance, or some other edaphic influence, on this continuum. An examination of the mean spectrum for each patch type from West Wits (Figure 3.3) showed that *S. plumosum* has a distinctly subdued spectrum with a distinct red edge indicating the main chlorophyll absorption feature centered at approximately 680 nm. This absorption feature is not seen in the two spectra representing sparse grass and grass patches, which are in a winter senesced state, but it does seem to be present to a lesser extent in bare patches and bare grass patches. However, in bare patches and bare grass patches it is unknown if this absorption feature represents algae and other photosynthetic components of biological soil crusts (Weber *et al.*, 2008), or some other constituent of the soil surface that is totally or largely exposed in these patches (De Jong *et al.*, 2011). Apart from the lower general reflectance and the chlorophyll feature in the near-infrared (NIR), the *S. plumosum* mean spectrum showed the same features as sparse grass and grass spectra in the shortwave infrared (SWIR). Sparse grass and grass patches had almost identical mean spectra with sparse grass patches having very slightly higher reflectance between 700 and 1100 nm but lower reflectance in the rest of the SWIR region (Figure 3.3). Bare grass and bare patches had almost identical patterns of reflectance across the spectrum with bare grass being brighter in the NIR region up to 900

2316 nm but bare patches being brighter in the rest of the SWIR region (Figure 3.3). An
 2317 absorption feature centered around 1729 nm that may relate to lignin or cellulose
 2318 (Daughtry *et al.*, 2004, Serrano *et al.*, 2002, Fourty *et al.*, 1996) was apparent in all three
 2319 vegetation types but was not present in bare patches and bare grass patches. Both mean
 2320 spectra for bare patch and bare grass patches had a significant absorption feature around
 2321 2202/3 nm which may be related to the ALOH spectral feature from smectite clay in the soil
 2322 (Chabrilat *et al.*, 2002). This feature was not present in the other three predominantly
 2323 vegetative patch types with little or no soil surface exposed. From a spectral perspective,
 2324 the spectra for the five patch types can be organized into three groups: a predominantly soil
 2325 group consisting of bare patches and bare grass patches, a grass group consisting of sparse
 2326 and grass patch types, and a third group consisting of *S. plumosum* patches (Figure 3.3).

2327 The loading values derived during the modelling process can be used for interpretation
 2328 purposes as they reflect features in the original spectra (Mevik and Wehrens, 2007). The
 2329 first three loadings are plotted in Figure 3.4 and together account for 98.6% of the variability
 2330 in the spectral data. Of this variability, the first component accounts for 90.8% and has
 2331 negative values over the entire spectrum. The first component has a local minimum
 2332 centered at 690 nm which probably corresponds with the chlorophyll absorption feature



2333
 2334
 2335

Figure 3.4 Loadings for the first 3 components of a five-patch model from West Wits.

2336 previously noted in the mean spectrum from *S. plumosum*. Other local minima in the first
2337 loading occur at 1203 - 1214 nm, 1992 – 2005 nm, 2126 – 2137 nm and 2226 – 2228 nm
2338 although, apart from the 2126 – 2137 nm feature which may be related to cellulose
2339 (Daughtry *et al.*, 2009, Daughtry *et al.*, 2004, Daughtry *et al.*, 2003) and the 2226 – 2228
2340 feature which may be related to clay (Chabrilat *et al.*, 2002), it is unclear what might be
2341 causing these local minima (Figure 3.4). Beyond this the loadings become too noisy to
2342 separate features from noise. The second loading has a strong local minimum at 595 nm,
2343 weaker local minima at 978 – 1010 nm, which may be related to either water or iron oxide
2344 (De Jong, 1991), and 1210 nm (Figure 3.4). Both the second and third loadings have
2345 common minima in the SWIR at 1729 nm and a deep minimum at 2099 – 2104 nm which
2346 probably corresponds with lignin and cellulose signals in the vegetation spectra (Daughtry *et*
2347 *al.*, 2009, Daughtry *et al.*, 2004, Daughtry *et al.*, 2003, Serrano *et al.*, 2002, Nagler *et al.*,
2348 2000). However, the third loading differs from the second loading in the NIR with a shallow
2349 local minimum at 449 nm and a deep local minimum at 753 – 766 nm (Figure 3.4).

2350 Pearson correlation coefficients between wavelengths for all spectra from West Wits clearly
2351 show the strong collinearity of spectra with almost all wavelengths having correlations
2352 greater than 70% (Figure 3.6). The exception was the visible wavelengths between 400 and
2353 500 nm with wavelengths greater than 600 nm. Of interest is the strong correlation
2354 between wavelengths around 600 nm with wavelengths at 1500 – 1550 nm and 2000 nm
2355 (Figure 3.5). Correlations between spectra from individual sample points at West Wits show
2356 that patch spectra clearly organize themselves into patch types and within these groups
2357 there may be sub-groups (Figure 3.6). Apart from one *S. plumosum* patch in transect 3 at
2358 West Wits, all the other *S. plumosum* patches in this transect organize into one group
2359 together with one *S. plumosum* patch from transect 1 and another from transect 2 (Figure
2360 3.6). The rest of the *S. plumosum* patches in transect 1 and transect 2 form another group,
2361 while a third group lying closer to the transect 3 grouping is made up of *S. plumosum* from
2362 transects 4 and 5 (Figure 3.6). The first *S. plumosum* group seems to have quite strong
2363 correlations with grass patches but less so with sparse grass patches, bare grass patches and
2364 bare patches. Bare patches and bare grass patches are strongly correlated with each other

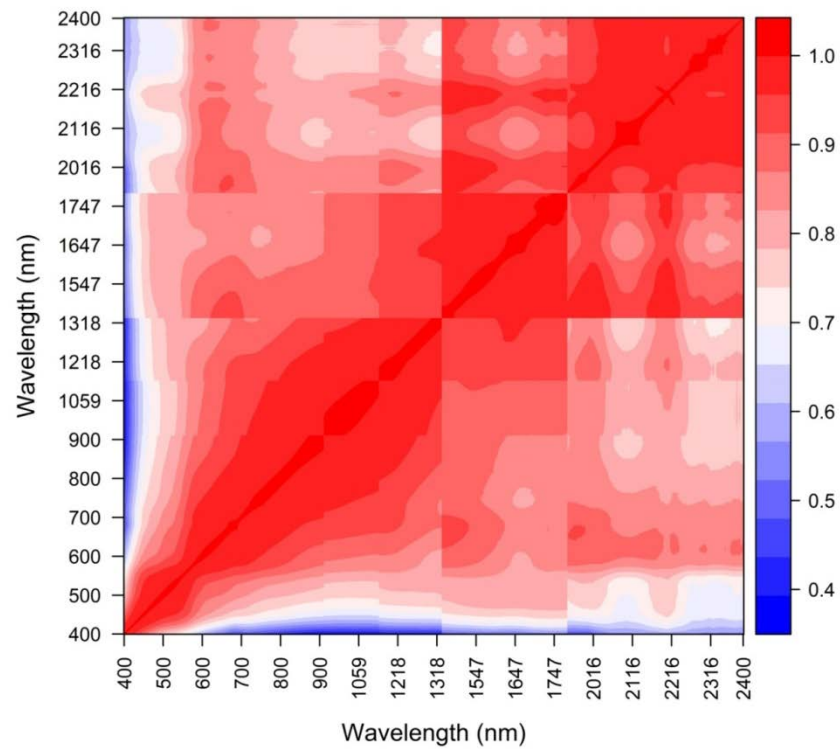


Figure 3.5 Pearson correlation coefficients for wavelengths from spectra from all sample points at West Wits.

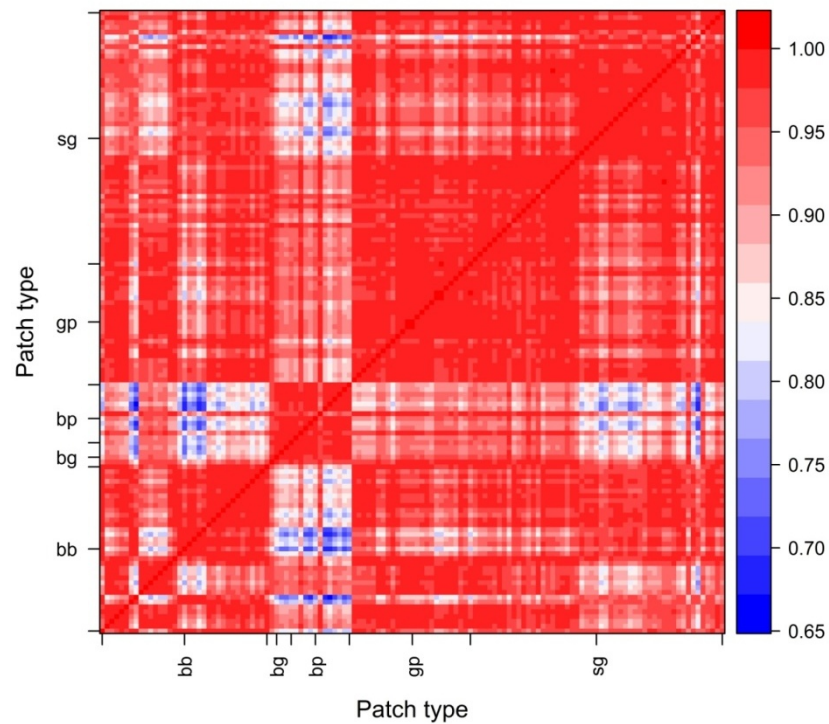
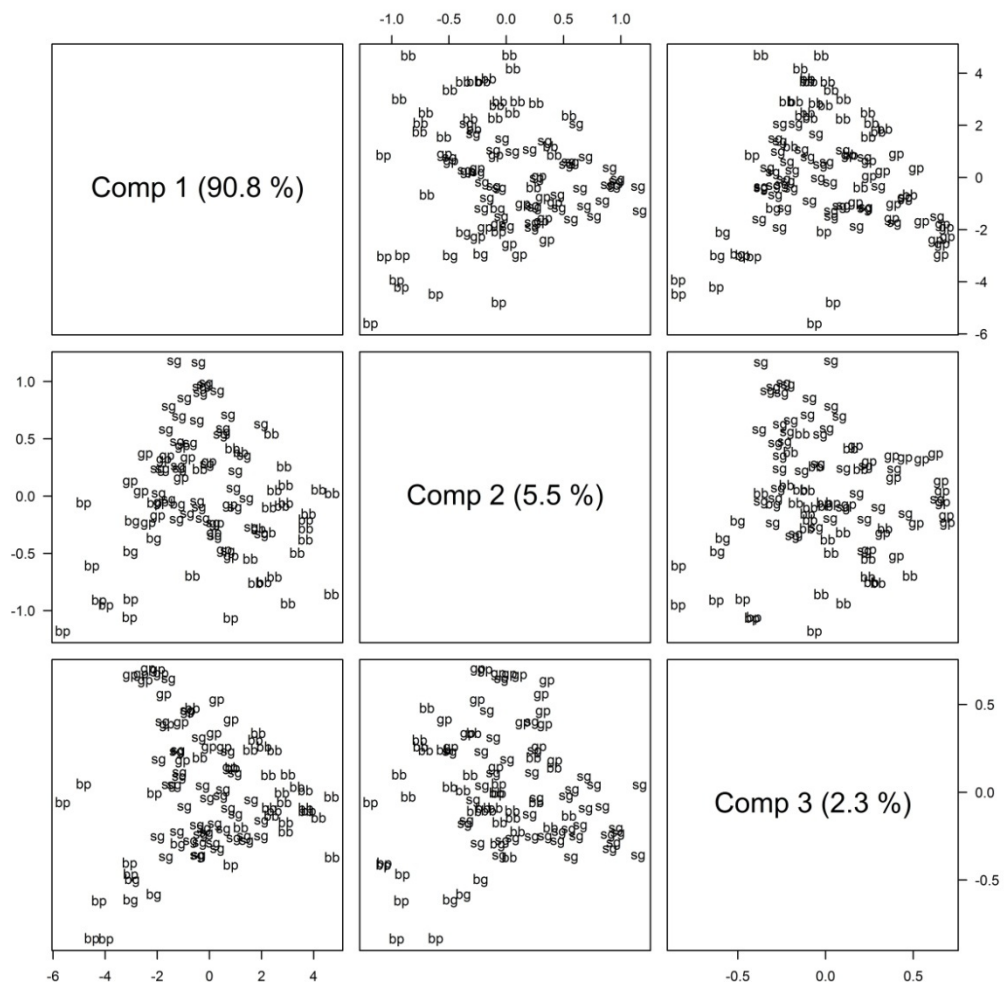


Figure 3.6 Pearson correlation coefficients for spectra from individual sample points from West Wits organized into patch types. Tick marks on the axes without labels indicate boundaries between patch types. Patch acronyms in the legend: bp = bare patch, bg = bare grass, sg = sparse grass, gp = grass patch, and bb = *S. plumosum* patches.

and form a cohesive group that is weakly correlated with *S. plumosum* patches forming group 1 and 3, and sparse grass patches, but slightly better correlated with the second *S. plumosum* grouping and some grass patches (Figure 3.6). Within this bare patch - bare grass cluster are three sample points that show stronger correlations to grass and sparse grass patches than they do to the bare patch-bare grass cluster. Within this bare patch - bare grass cluster are three sample points that show stronger correlations to grass and sparse grass patches than they do to the bare patch-bare grass cluster. Two of these were classified as bare patches and examination of the photos shows that these bare patches were surrounded by large grass tufts with tendrils of dry grass trailing into the edges of the bare patch and thus presumably confounding the spectrum and the classification of the patches. The third patch that forms an exception to the bare-patch-bare grass cluster is a bare-grass patch from transect 5 that had small stunted grass tufts that often defined bare grass patch type. However, there were enough of these individual tufts in the target area to produce a fairly strong senescent grass signal and thus the strong affinity with sparse grass and grass patches for this sample point. Grass patches generally form a strongly correlated group (Figure 3.6). Sparse grass patches on the other hand separate into two distinct groups, one of which is closely correlated with grass patches and the other forms a distinct group of its own (Figure 3.6). This seems to be a landscape effect as the two groups belong to specific transects although the underlying cause is at present unknown. Within these three groups there are individual points in the group that are better correlated with the other grass group rather than their classified grass group. Rather unexpectedly, the bare patch – bare grass group shows weaker correlations with sparse grass than it does with grass patches (Figure 3.6), as the exposure of more soil surface in sparse grass patches than grass patches, at least to the eye, would have suggested closer affinities for bare patches and bare grass patches to sparse grass patches than to grass patches.

Another method of exploring the data for outliers is the scores used in the PLSR modelling process (Mevik and Wehrens, 2007). An examination of the scores for the first three components, which together explained 98.6% of the variability in the spectral data, showed that there were no extreme outliers in the West Wits data (Figure 3.7). Patch types tended to cluster in specific areas of the component space. *S. plumosum* patches are mostly clustered towards the positive end of the first component axis and bare patches towards the

2405 negative end and sparse grass patches spread through the center. Component two clustered
 2406 bare patches and some *S. plumosum* patches towards the negative end and sparse grass
 2407 patches towards the positive end with grass patches around the center. Component three
 2408 followed a similar pattern to component two but switched sparse grass and grass patches
 2409 with sparse grass tending to cluster around the center of the axis and grass patches towards
 2410 the positive end of the third component axis (Figure 3.7). This suggests that the first
 2411 component is principally separating *S. plumosum* and bare patches from the other patches,
 2412 while component two is separating sparse grass and to some extent bare grass, and
 2413 component three is separating grass patches from the bare patch - bare grass complex.



2414
 2415 **Figure 3.7** Score plots for the first three components of a five-patch model from West
 2416 Wits which together account for 98.6% of the spectral variation.

2417 Like the loadings, the regression coefficients can provide interpretable information as they
 2418 relate to spectral patterns in the original spectra (Figure 3.8 and 3.9). The regression
 2419 coefficients for the 8-component model have a common pattern for the five patch types but

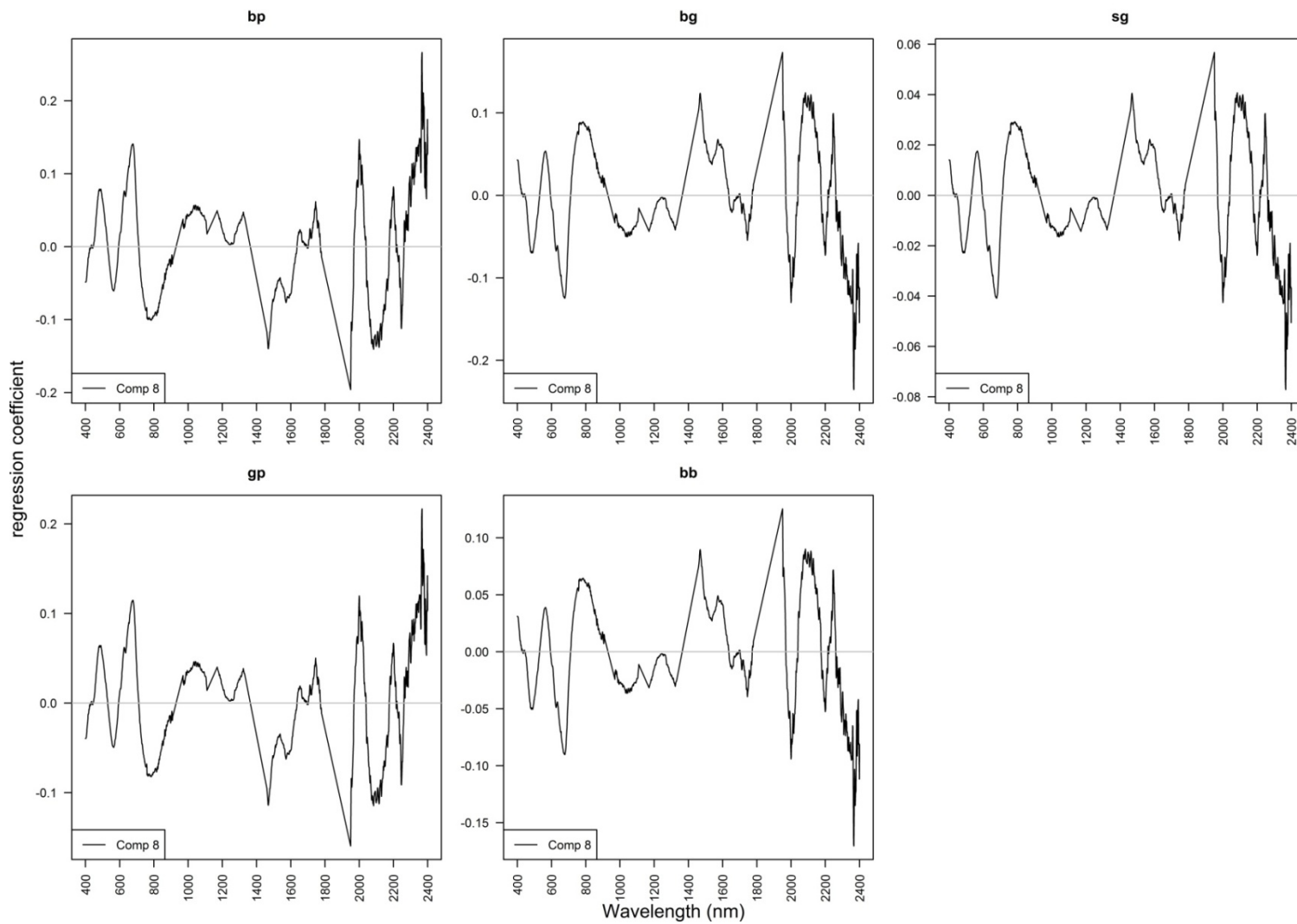


Figure 3.8

Regression coefficients from an 8-component model predicting patch types from ground-based hyperspectral measurements. Patch acronyms are: bp = bare patch, bg = bare grass, sg = sparse grass, gp = grass patch, and bb = *S. plumosum* patches.

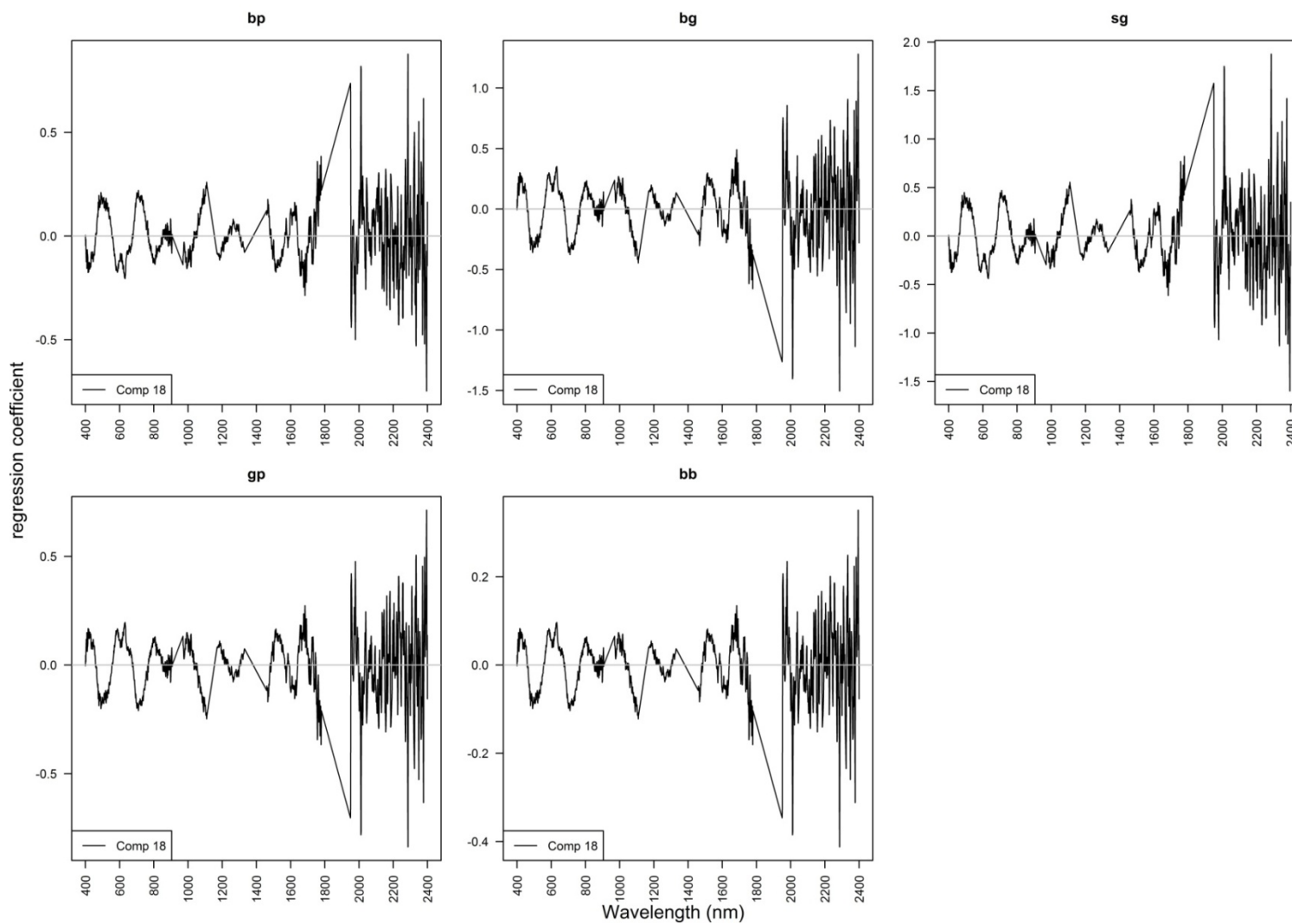


Figure 3.9

Regression coefficients from an 18-component model predicting patch types from ground-based hyperspectral measurements. Patch acronyms are: bp = bare patch, bg = bare grass, sg = sparse grass, gp = grass patch, and bb = *S. plumosum* patches.

may vary in scale and/or be a mirror image depending on the patch type being predicted. So the regression coefficients for bare and grass patches have one shape, while bare grass, sparse grass and *S. plumosum* patches show the mirror image (Figure 3.8). This pattern of mirror image in the regression coefficients for the 8-component model is followed in the 18-component model but with some rearrangement of the patch groupings. Bare and sparse grass patches have one pattern of coefficients and bare grass, grass patches and *S. plumosum* patches have the reverse pattern (Figure 3.9). The scale of the regression coefficients also varies between patch types with bare patches, bare grass patches and grass patches having coefficients between -0.2 and 0.2 while sparse grass patches had coefficients between -0.08 and 0.06 and *S. plumosum* patches had coefficients between -0.15 and 0.15 for the 8-component model (Figure 3.8). The troughs and peaks in the regression coefficients reflect, in a complex manner, the absorption centers and areas of high reflectance in the original spectra. As the peaks and troughs can be mirror images of each other, it is not clear whether they relate to absorption centers or high reflectance features. Furthermore, as the regression coefficients are a combination of all the components in the model selected, they represent a complex combination of the original spectra. Local minima in the regression coefficients for bare patches and grass patches and an 8-component model are found at 565, 635, 816, 1219, 1556, 1582, 1690 – 1721, 2073 – 2131 and 2246 nm (Figure 3.8). For bare grass, sparse grass and *S. plumosum* patches, the regression coefficients had local minima at 482, 678, 1528, 1646, 1992 and 2009 nm (Figure 3.8). It is noticeable that some of these minima are quite noisy, especially when compared with the first three loadings displayed in Figure 3.4, and this increases in the SWIR above 2000 nm for the 8-component model. For the 18-component model it is not feasible to locate the minima and maxima as so much noise is being expressed in the regression coefficients around these points and the entire SWIR at wavelengths longer than 2000 nm seems to be noise. This suggests that models with more than 8 components are not likely to be very robust in terms of reflecting spectral features in the target, but rather improvements in models with many components are improving through modelling some of the noise in the data.

An alternative model was developed by taking into account the spectral similarity, as illustrated in the mean spectra (Figure 3.3), and spectral correlations between sample points

Table 3.4 Cross-validated (CV) RMSEP using leave-one-out (LOO) estimates for calibration models predicting revised patch types at West Wits from ground-based hyperspectral measurements. The table also presents the coefficient of determination (r^2) representing the explained variance for each variable as one adds components to the model. Values in bold underlined italics reflect models with local RMSEP minima for that patch type, while the bold underlined values reflect the variation explained for that variable by three selected models using calibration data. Patch acronyms are: bgp = bare-grass patch combination, sgp = sparse-grass patch combination, and bb = *S. plumosum* patches.

Component	Cross-validated RMSEP			Spectra	Calibration r^2		
	bgp	sgp	bb		bgp	sgp	bb
1	0.2929	0.4868	0.3194	90.74	30.12	6.67	50.12
2	0.209	0.3805	<u>0.3064</u>	96.06	65.35	43.65	54.82
3	<u>0.1719</u>	0.3684	0.3096	98.54	77.29	48.34	54.88
4	0.1734	0.3253	0.26	99.13	77.33	61.5	69.71
5	0.1755	0.3062	0.2315	99.62	77.33	67.72	77.22
6	0.1743	0.2576	0.1862	<u>99.86</u>	<u>78.01</u>	<u>78.44</u>	<u>86.18</u>
7	<u>0.1632</u>	0.2417	0.185	99.9	82.02	82.09	86.51
8	0.1645	0.2268	<u>0.1689</u>	99.92	82.3	86.43	90.1
9	0.1547	0.2159	0.1703	<u>99.94</u>	<u>86.52</u>	<u>87.96</u>	<u>90.14</u>
10	0.1532	0.2138	0.1608	99.95	87.32	88.51	92.44
11	<u>0.1521</u>	<u>0.2095</u>	0.1553	<u>99.96</u>	<u>87.33</u>	<u>89.25</u>	<u>93.46</u>
12	0.1523	0.2143	<u>0.1495</u>	99.97	87.38	89.96	94.68
13	0.1592	0.229	0.1513	99.97	89.41	91.18	94.69
14	0.1628	0.2381	0.1543	99.98	90.4	93.02	95.23
15	<u>0.1581</u>	<u>0.2334</u>	0.1563	99.98	90.76	94.15	95.75
16	0.1623	0.2394	0.1583	99.98	90.95	94.98	96.2
17	0.1627	0.245	0.1587	99.99	92.19	95.58	96.2
18	0.1673	0.2457	0.1592	99.99	94.87	96.45	96.25
19	0.1698	0.2469	0.1673	99.99	95.17	96.87	97.55
20	0.1756	0.2669	0.1751	99.99	95.48	97.69	97.87

(Figure 3.6). Bare patches and bare grass patches were combined into a bare-grass patch (bgp), and sparse grass and grass patches combined into a sparse-grass patch type (sgp). Generally, the results with the calibration data show an immediate improvement with absolute minima being reached in fewer components in the new model compared to the original five-patch model. The bare-grass patch combination took 11 components to reach its absolute minimum RMSEP (0.1521, Table 3.4) compared to either bare patches alone

2475 which took 15 components (RMSEP = 0.1395, Table 3.2), or bare grass alone which took 20
2476 components (RMSEP = 0.1422, Table 3.2). However, it should be noted that the RMSEP for
2477 the individual patches at the absolute minima was slightly lower than that for the
2478 combination patch at its absolute minimum which probably reflects the slightly increased
2479 variability in the spectra for the combined patch type compared to the two original
2480 individual patch types. The first and absolute RMSEP minimum for the sparse-grass patch
2481 combination was at 11 components (Table 3.4) whereas both sparse grass and grass patches
2482 took more components (15 and 17 respectively) to reach their absolute RMSEP minima
2483 (Table 3.2). Furthermore, the absolute RMSEP value for the sparse-grass patch combination
2484 (0.2095, Table 3.4) was lower than that for both sparse grass (0.2587) and grass patches
2485 (0.2197, Table 3.2) alone. Also noticeable in the new 3-patch model was that the
2486 coefficients of determination reached values greater than 50% with four components and
2487 calibration data (Table 3.4) compared to eight components in the five-patch model (Table
2488 3.2).

2489 The selection of the optimum model faced the same conundrums as the five-patch model
2490 did. The first local minimum for *S. plumosum* was at two components and for the bare-grass
2491 patch combination it was at three components. However, for the sparse-grass patch
2492 combination the first local minimum, which was also the absolute minimum, was reached at
2493 11 components. In terms of adding components that decreased the RMSEP by 2%, for the
2494 bare grass patch combination, the first three components had a combined 46% decrease in
2495 RMSEP and the next two components increased the RMSEP by 2% (results not shown).
2496 Similarly, for *S. plumosum* patches the first two components had a 4% decrease in RMSEP
2497 while the third component increased the RMSEP by 1%, but thereafter the RMSEP
2498 decreased by 16%, 11% and 20% for the 4th, 5th and 6th components. For the sparse-grass
2499 patch combination, there was a consistent decrease in RMSEP of greater than 2% for all
2500 components up to and including the 9th component (results not shown).

2501 Taking all this into account, possible models include a 6-component model, a 9-component
2502 model and an 11-component model. The 6-component model is the most parsimonious with
2503 the 6th component contributing a 19% decrease in RMSEP for *S. plumosum* patches even
2504 though the seventh component was a local minimum in RMSEP for the bare-grass patch
2505 combination. Adding the seventh component would have decreased the bare-grass patch

2506 combination RMSEP by 6% but only slightly reduced the RMSEP (0.6%) for *S. plumosum*
2507 patches. The sixth component also had a major decrease in the RMSEP for the sparse-grass
2508 patch combination of 16%. So for both the sparse-grass patch combination and the *S.*
2509 *plumosum* patch type this sixth component made a major contribution to accuracy of the
2510 model in predicting patch types. The sparse-grass patch combination had a continuous
2511 decrease in RMSEP of greater than 2% per component up to and including the 9th
2512 component where the tenth component only decreased the RMSEP by 1%. This 9th
2513 component also decreased the RMSEP for the bare grass patch combination by 6%. This
2514 decrease in RMSEP in both the bare-grass patch combination and the sparse-grass patch
2515 combination was the basis for choosing a 9-component model even though no RMSEP
2516 minima had yet been achieved for the sparse-grass patch combination. The first RMSEP
2517 minimum for the sparse grass-patch combination was recorded at 11 components and
2518 formed the basis for choosing an 11-component model for testing with calibration data.

2519 The 6-component model had strong coefficients of determination for the three patch types
2520 (78%, 78% and 86%, all $P < 0.0001$, for the bare-grass patch combination, sparse-grass patch
2521 combination and *S. plumosum*, respectively) when predicting values with calibration data
2522 (Table 3.4 and Figure 3.10). Apart from *S. plumosum*, the coefficients of determination with
2523 validation data were even stronger with the bare-grass patch combination producing an $r^2 =$
2524 0.91 ($P < 0.0001$) and sparse-grass patch combination an $r^2 = 0.80$ ($P < 0.0001$; Figure 3.10).
2525 For *S. plumosum* the validation coefficients of determination dropped slightly to 84% (Figure
2526 3.10). However, despite these strong coefficients of determination, there was overlap
2527 between the predicted bare-grass patch combination and other patch types with the
2528 calibration data (Figure 3.10). Similarly, the sparse-grass patch combination had poor
2529 separation with the rest of the patch types at the limits of the range of predictions between
2530 the two classes, whereas *S. plumosum* calibration predictions had clear separation for
2531 predicted values and values predicted for the rest of the patch types with calibration data.
2532 With validation data, the bare-grass patch combination had clear separation between
2533 predicted values for the patch type and those for the rest of the patch types and *S.*
2534 *plumosum* also had fairly clear separation for validation data but separation of the sparse-
2535 grass patch combination from the rest of the patch types was small for validation data
2536 (Figure 3.10).

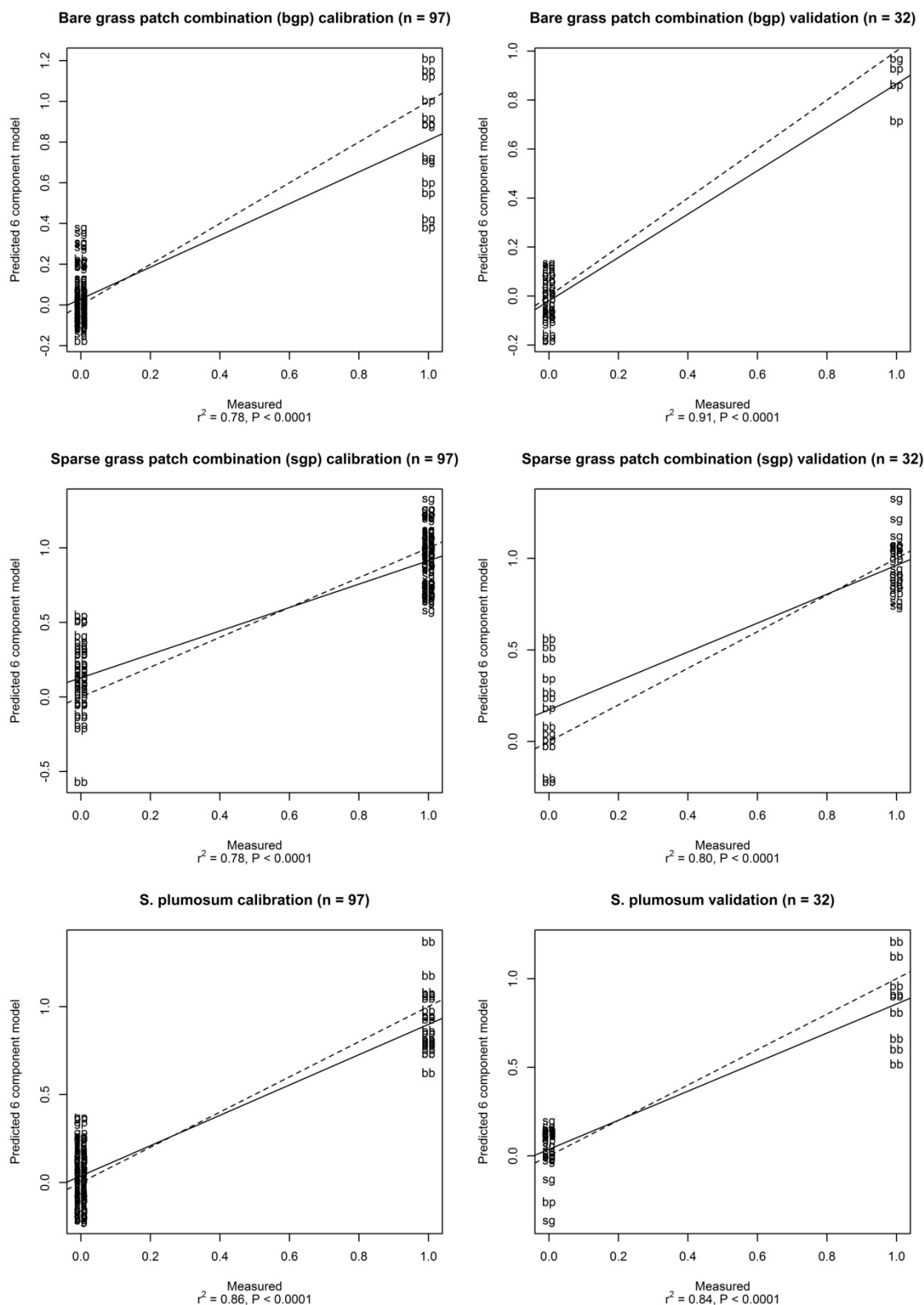


Figure 3.10 Calibration (left column) and validation (right column) predictions from a 6-component model for three patch types at West Wits. The dotted line represents the 1:1 fit between predicted and measured values while the solid line represents the best simple regression fit between predicted and measured values.

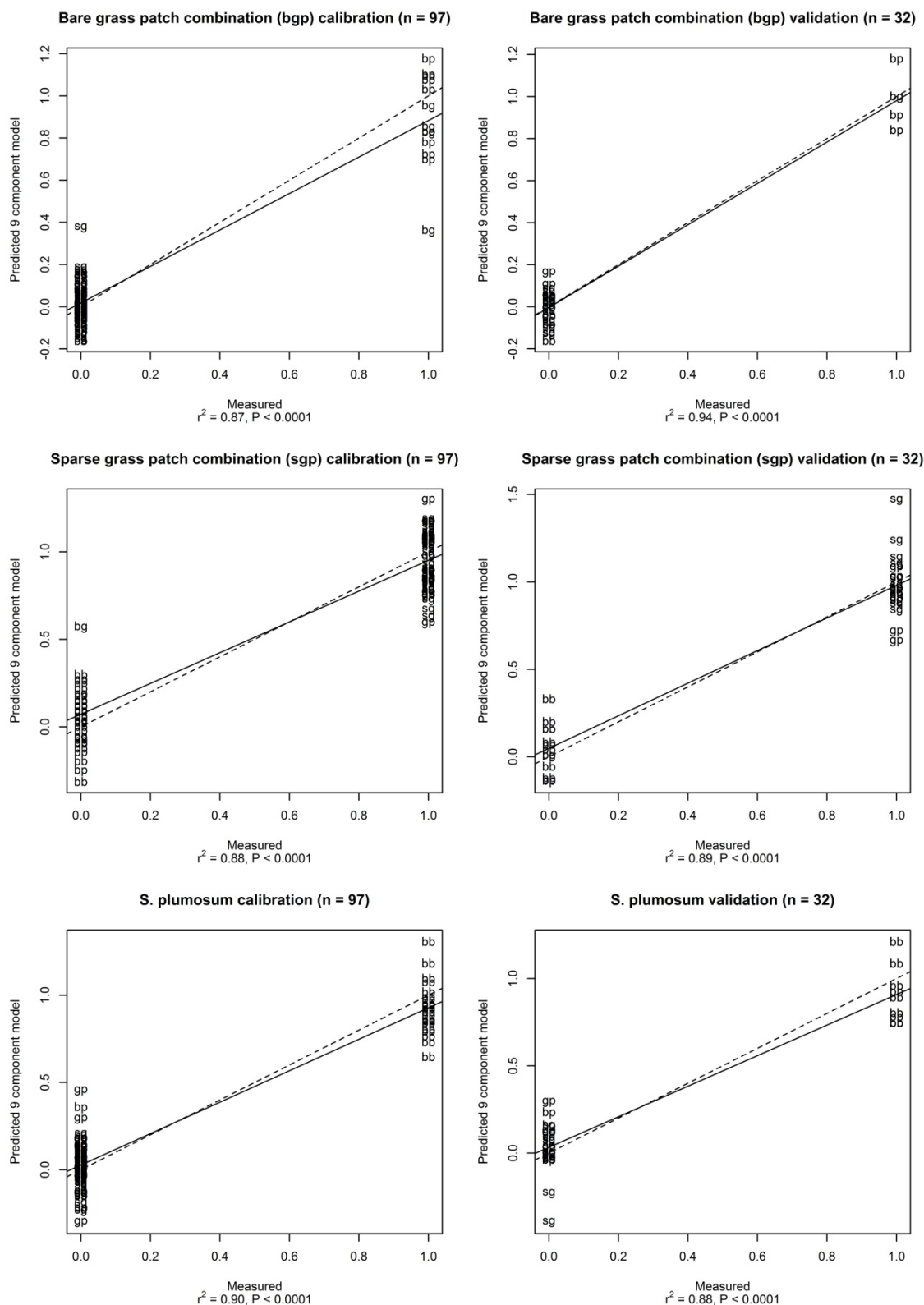


Figure 3.11 Calibration (left column) and validation (right column) predictions from a 9-component model for three patch types at West Wits. The dotted line represents the 1:1 fit between predicted and measured values while the solid line represents the best simple regression fit between predicted and measured values.

On the other hand, the 9-component model achieved good separation for all three predicted patch types with both calibration and validation data with the exception of two points (Figure 3.11). These two points were identified as ww5sgt105e and ww5bg1 which had both been previously identified as possibly miss-categorized based on their spectral correlations. Validation coefficients of determination were 0.94, 0.89, and 0.88 for the bare-grass patch combination, sparse-grass patch combination and *S. plumosum* patches, respectively (Figure 3.11). As the 9-component model achieved such clear separation and in the interests of parsimony in the model, there was no need to test and present the results of the 11-component model. Other patch organizations were tested but none provided better models than this 9-component three-patch model did to predict West Wits patches.

3.3.2 Patch Models for Vaal River

Vaal River had seven patch types spread across all the transects (Table 3.5). These represent a total of 291 points spread unevenly between the patch types but representing in a relative manner the occurrence of the different patch types on the transects, and to some extent in the landscape. A confounding variable was the presence of evaporite (referred to in this study as gypsum but this has not been verified) on some transects and principally affecting some bare grass, sparse grass and biological soil crust patches. Figure 3.12 presents the correlogram of the spectral correlations between the different sample points from Vaal River. In Figure 3.12a it is clear two spectra are significantly different from all others. These two spectra are gypsum-encrusted biological soil crust (vr8bscg2 and 3) with a very high reflectance in the visible region, possibly caused by the globulose, spherical nature of the evaporite in these two patches. They were removed from the subsequent correlogram (Figure 3.12b) which then shows clear differences in the spectra of different patch types. *Seriphium plumosum* (bb) patches show high correlation with each other and less so with the other vegetated patch types but lower correlation generally with biological soil crust and bare grass patches. The other vegetated patches: sparse grass, grass, tall grass and *Schoenoplectus* patches show similar spectral correlation to each other and *S. plumosum* patches.

Bare grass and biological soil crust patches show strong spectral correlation within each group and to each other, although they do differ from each other, and have less correlation

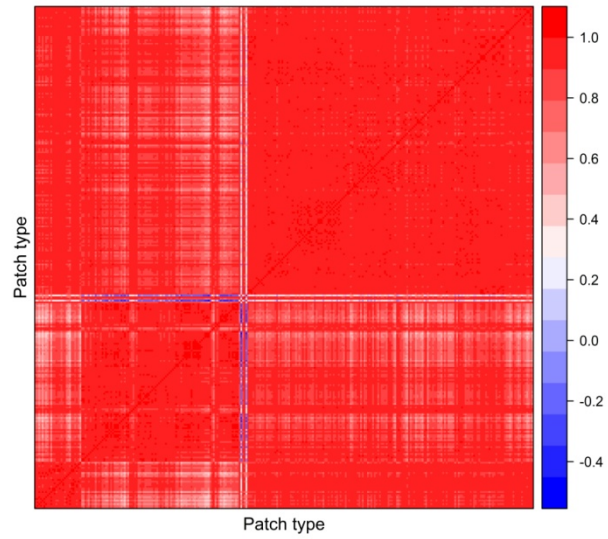
2579 to the vegetated patch types. However, within these two groups are individual patches that
 2580 show stronger correlation with the vegetated patches than they do with their own group
 2581 and this trend is stronger for bare grass patches. Within each patch type are sub-groups that
 2582 show stronger correlation to each other than they do to their specific patch type and these
 2583 sub-groups are often associated with specific transects suggesting edaphic and landscape
 2584 position effects.

2585 Correlations between wavelengths (Figure 3.13) of the spectra from all points sampled at
 2586 Vaal River show less correlation as the distance between wavelengths increases than that
 2587 shown by correlations between wavelengths for West Wits spectra (Figure 3.5). The reason
 2588 for this is not clear but in essence wavelengths in the SWIR between 1950 and 2400 nm
 2589 show weak to no correlation with wavelengths in the visible and NIR spectrum between 400
 2590 and 1330 nm. The general trend was that wavelengths were strongly correlated with nearby
 2591 wavelengths and this correlation dropped off with increasing distance between
 2592 wavelengths. Correlations between wavelengths formed three distinct blocks: the visible
 2593 and NIR region up to 1330 nm, a SWIR block between 1460 and 1780 which was moderately
 2594 correlated with the other two regions, and a second SWIR region between 1950 and 2400
 2595 nm.

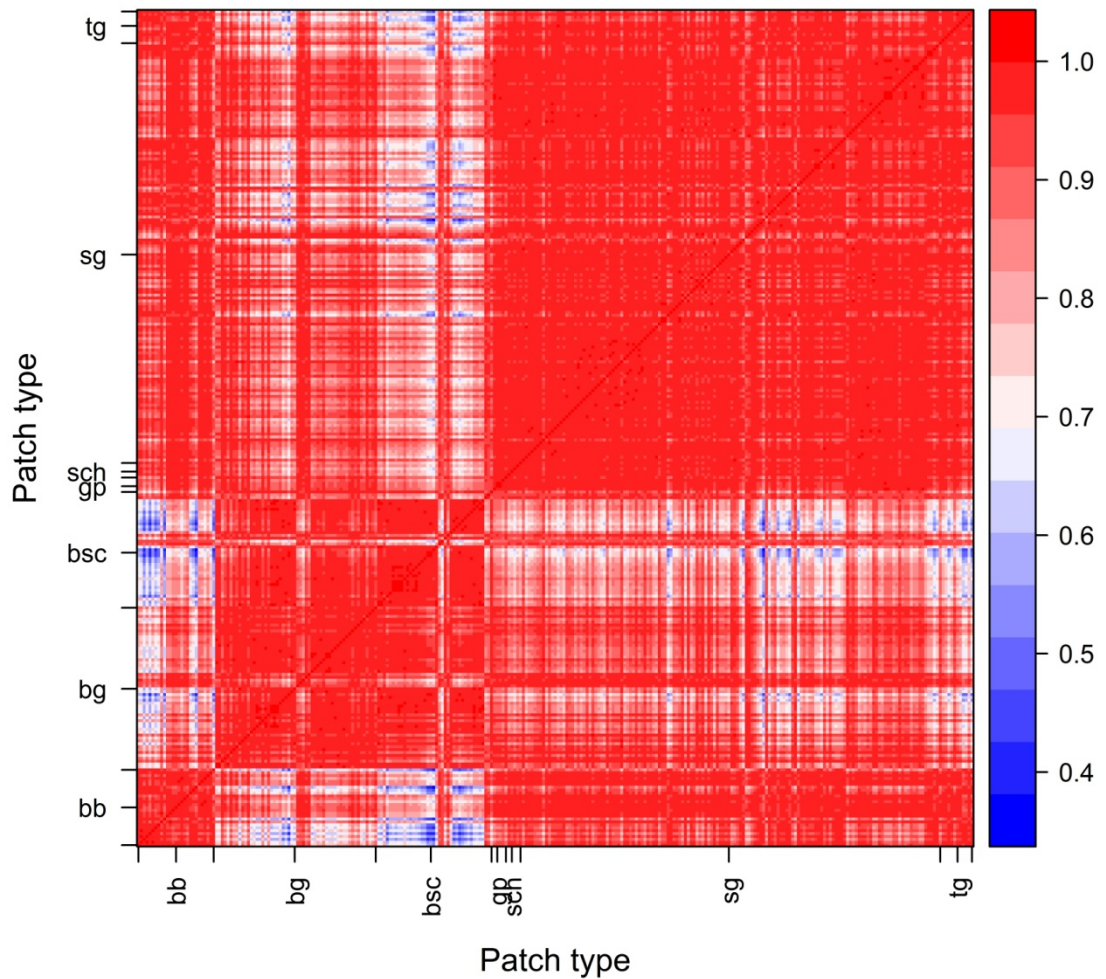
2596 **Table 3.5** Number of samples in each patch type from Vaal River. The patch type
 2597 acronyms are bb = *S. plumosum*, bg = bare grass, bsc = biological soil crust,
 2598 gp = grass patch, sch = *Schoenoplectus* patch, sg = sparse grass, tg = tall grass.

Patch type	Calibration			Validation		
	Total	with gypsum	no gypsum	Total	with gypsum	no gypsum
bb	22	0	22	5	0	5
bg	39	7	32	16	3	13
bsc	31	7	24	11	3	8
gp	4	0	4	1	0	1
sch	4	0	4	1	0	1
sg	111	8	103	35	2	33
tg	8	0	8	3	0	3
Total	219	22	197	72	8	64

2599



2600



2601

2602 **Figure 3.12** (a) Correlogram for all spectra from Vaal River. (b) Correlogram for spectra
 2603 from Vaal River with two gypsum patches removed from the data set. Tic
 2604 marks without labels show the boundary between two patch types. Codes for
 2605 the labels are: bb = *S. plumosum*, bg = bare grass, bsc = biological soil crust,
 2606 gp = grass patch, sch = *Schoenoplectus* patch, sg = sparse grass, tg = tall grass.

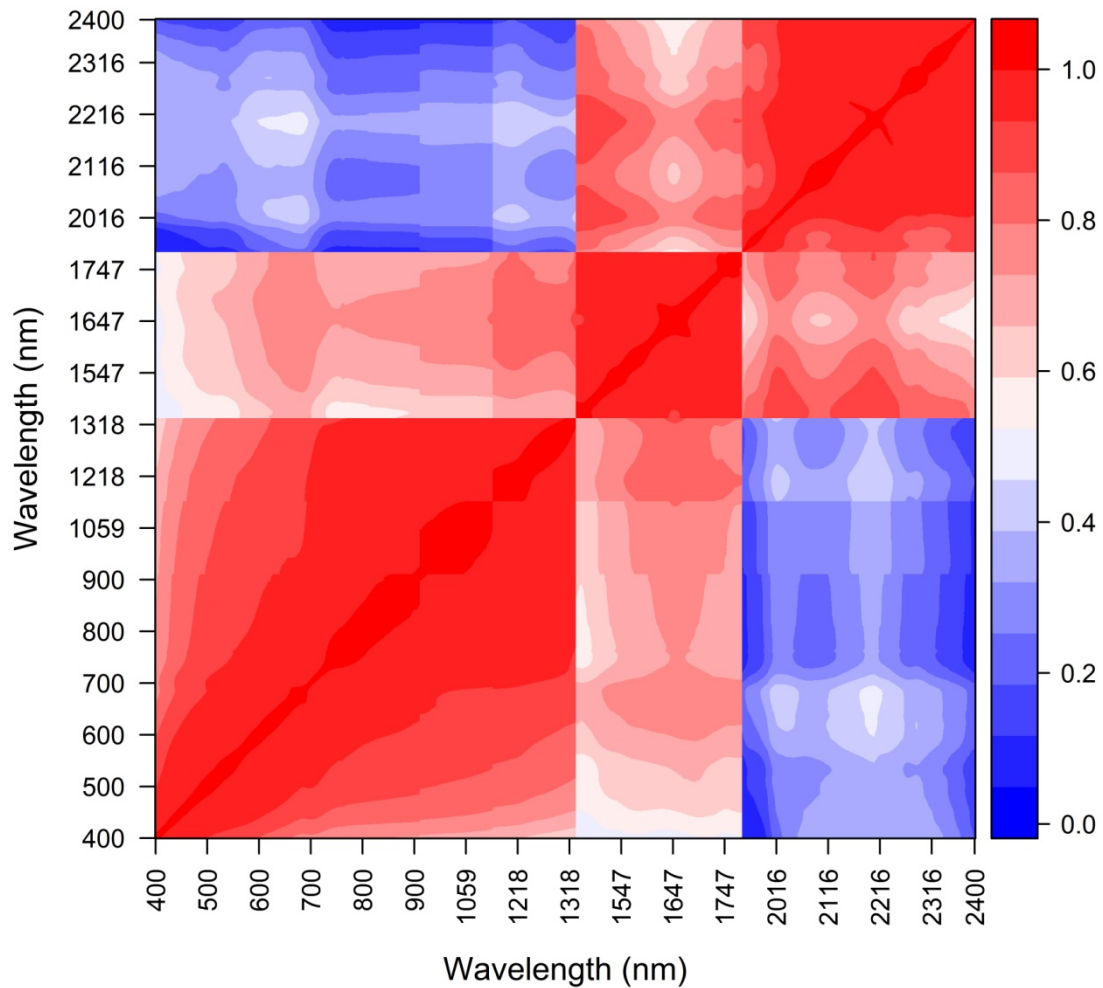


Figure 3.13 Pearson correlation coefficients for wavelengths from spectra from all sample points at Vaal River.

The dominant feature in the mean spectra for each patch type (Figure 3.14) is the red-edge in the *S. plumosum* spectrum due to the chlorophyll absorption center around 670 - 680 nm (Tucker, 1979, Jordan, 1969). This is expected as *S. plumosum* remains evergreen through winter when this study was conducted. The sedge patch dominated by *Schoenoplectus* had a less pronounced red-edge feature and three of the four grass patch types: bare grass, sparse grass and grass patches all show small features around this chlorophyll absorption feature (Figure 3.14). Tall grass only shows a minor inflection at this point but biological soil crusts did show a feature centered on the chlorophyll absorption wavelength which was presumably a result of an algal component in the crust (Figure 3.14). The mean spectra tend to cluster into groups with *S. plumosum* and *Schoenoplectus* patches forming one group with reflectance generally slightly higher for *S. plumosum* patches. Sparse grass and grass

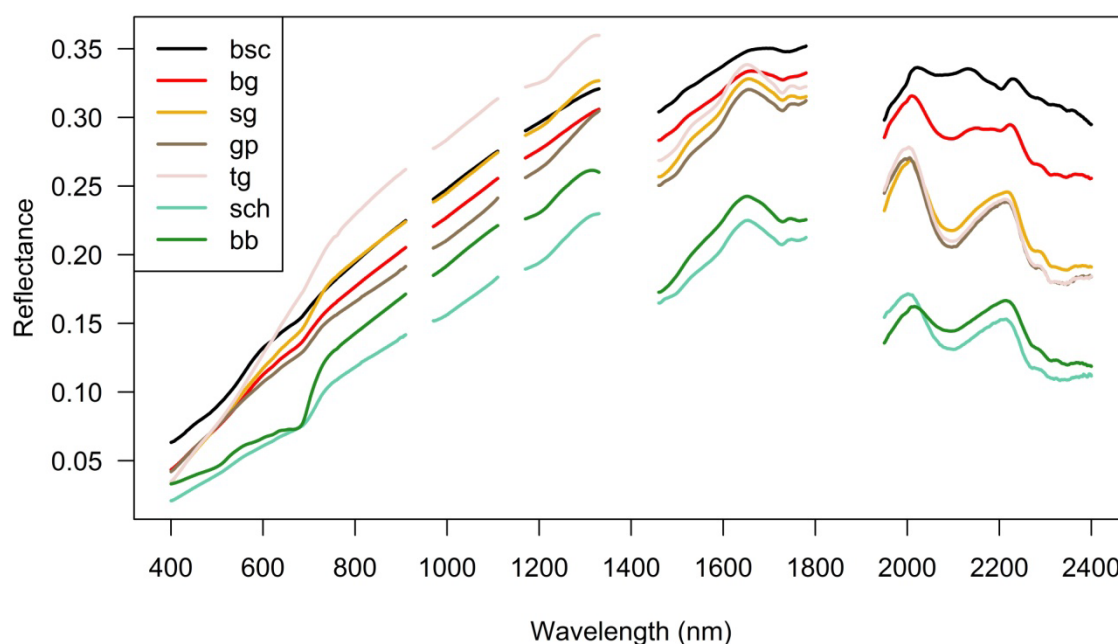


Figure 3.14 Mean spectra for the seven patch types sampled at Vaal River. Codes for patch types are bsc = biological soil crust, bg = bare grass, sg = sparse grass, gp = grass patch, tg = tall grass, sch = *Schoenoplectus* patch, bb = *S. plumosum*.

patches form another group with tall grass patches where the mean spectra are almost on top of each other in the SWIR region (Figure 3.14). However, in the visible and NIR (VNIR) regions from about 600 to 1330 nm there are notable differences in brightness with grass patches the least bright and tall grass patches significantly brighter than the rest of the spectra in this region (Figure 3.14). This brightness of the tall grass patches may reflect the bright yellow culms that make up the bulk of tall grass patches. All five vegetation-based patch types show a small absorption feature around 1726 nm (Figure 3.14) that is possibly a function of lignin (Serrano *et al.*, 2002) and a strong absorption center around 2080 nm related to cellulose (Daughtry *et al.*, 2004, Daughtry *et al.*, 2003).

Biological soil crust and bare grass patches to a lesser extent show an absorption feature around 2203 which is best matches the AlOH clay feature (Figure 3.14). Most likely, the varying amount of vegetation in the bare grass patches is influencing the degree to which the clay signal shows in the mean spectral profile of bare grass patches. Another confounding variable for these two patch types was the presence or absence of surface gypsum crusts, which was also found in some sparse grass patches (Figure 3.15). Biological soil crust

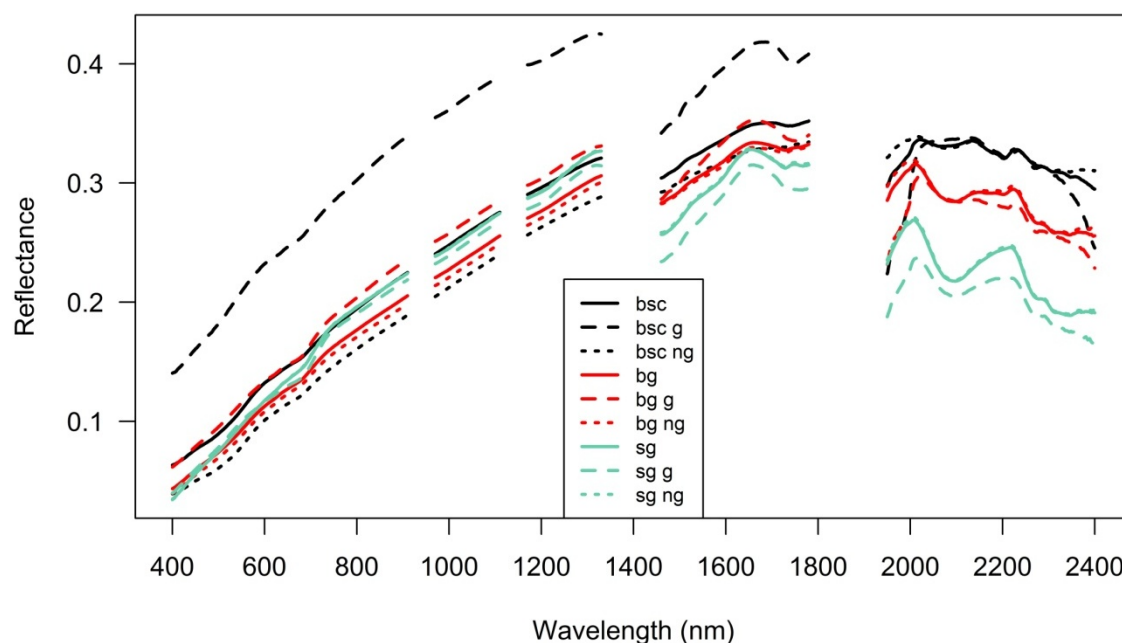


Figure 3.15 Mean spectra for the three patch types that have some samples with efflorescent crusts. Patch labels on their own indicate the mean for all samples of that patch type (bsc = biological soil crust, bg = bare grass, sg = sparse grass). The mean for the sub-sample with efflorescence is indicated by a g and for the subsample without efflorescence is indicated by an ng.

patches with gypsum have a particularly bright spectrum in the VNIR and SWIR wavelengths up to 1780 nm but in the rest of the SWIR (> 1950 nm) there is little difference between the mean spectra with or without gypsum (Figure 3.15). Likewise for bare grass patches, the gypsum dominated patches tend to have a brighter mean spectrum in the VNIR and SWIR region up to 1780 nm but overall this seems to have a much smaller affect on the mean spectra for bare grass patches than did the gypsum-affected patches in biological soil crust patches (Figure 3.15). Sparse grass patches were little influenced by the presence of gypsum which tended to result in a slightly lower brightness of the spectrum for gypsum affected patches compared to non-gypsum sparse grass patches throughout the spectrum (Figure 3.15).

A PLSR model was derived to predict the seven patch types at Vaal River with all calibration sample points (219 samples), including all gypsum-influenced points (Table 3.6). However, this model is weak as it can only account for 28% of the variability in grass patches with 22 components in the model. *Schoenoplectus* patches are also weakly identified by the model

2662 **Table 3.6** Cross-validated RMSEP using LOO estimates for calibration models predicting patch types at Vaal River from ground-based hyperspectral
2663 measurements using all points. The table presents the coefficient of determination (r^2) representing the explained variance for each
2664 variable as one adds components to the model. Values in bold underlined italics reflect models with local RMSEP minima for that patch
2665 type. The bold underlined values reflect variation explained for that variable by three selected models. Patch acronyms are: bsc = biological
2666 soil crust, bg = bare grass, sg = sparse grass, gp = grass patch, tg = tall grass, sch = *Schoenoplectus* patch, and bb = *S. plumosum* patches.

Component	Cross-validated RMSEP							Calibration r^2							
	bsc	bg	sg	gp	tg	sch	bb	Spectra	bsc	bg	sg	gp	tg	sch	bb
1	0.2941	0.3641	0.4962	0.1345	0.1887	0.1320	0.2454	62.34	31.29	10.88	4.12	0.19	0.001	6.98	35.50
2	0.2834	<u>0.3502</u>	0.4587	0.1346	0.1845	<u>0.1319</u>	0.2442	95.30	38.60	18.67	20.25	0.20	6.73	7.91	36.93
3	0.2351	0.3514	0.4053	0.1346	<u>0.1831</u>	0.1323	0.2309	98.02	56.41	18.77	36.68	1.74	8.75	7.95	44.75
4	<u>0.2303</u>	0.3513	0.3922	0.1334	0.1839	0.1324	0.1911	98.93	58.68	19.16	43.38	4.31	8.75	8.29	63.64
5	0.2319	<u>0.3512</u>	<u>0.3707</u>	<u>0.1316</u>	0.1804	<u>0.1317</u>	0.1713	<u>99.34</u>	<u>58.79</u>	<u>20.90</u>	<u>49.02</u>	<u>9.64</u>	<u>13.03</u>	<u>9.20</u>	<u>69.89</u>
6	0.2329	0.3565	0.3738	0.1317	0.1796	0.1321	<u>0.1698</u>	99.77	58.79	21.63	55.12	9.88	23.08	9.22	71.16
7	0.2359	0.3487	<u>0.3306</u>	0.1327	<u>0.1622</u>	0.1322	0.1743	99.86	58.79	24.58	61.97	10.10	33.40	10.87	71.19
8	0.2173	0.3350	0.3308	0.1333	0.1631	0.1319	<u>0.1717</u>	99.89	65.47	31.46	61.99	10.35	33.57	11.99	71.94
9	0.2112	<u>0.3279</u>	0.3391	0.1353	0.1648	<u>0.1302</u>	0.1744	99.91	68.90	39.37	62.94	17.16	36.80	15.64	71.95
10	0.2107	0.3302	<u>0.3216</u>	0.1354	<u>0.1485</u>	0.1304	0.1822	99.93	69.48	41.12	66.04	17.16	47.65	16.44	72.08
11	0.2075	<u>0.3177</u>	0.366	0.1325	0.1581	0.1448	0.2806	99.95	70.35	44.89	66.44	17.20	47.66	17.60	72.86
12	0.2033	0.3264	0.3372	<u>0.1324</u>	0.1583	0.1481	0.2619	<u>99.96</u>	<u>71.81</u>	<u>47.24</u>	<u>69.08</u>	<u>17.32</u>	<u>47.88</u>	<u>17.78</u>	<u>76.44</u>
13	<u>0.2030</u>	0.3187	0.3335	0.1334	0.1530	0.1460	0.2426	99.97	71.91	47.77	69.88	19.70	50.15	20.52	82.83
14	0.2031	<u>0.3174</u>	<u>0.3083</u>	0.1353	<u>0.1477</u>	0.1543	0.2126	99.98	71.96	47.89	71.68	21.19	50.25	20.52	84.54
15	0.2042	0.3361	0.3113	<u>0.1336</u>	0.1513	0.1336	<u>0.2066</u>	99.98	74.69	50.90	72.76	21.74	51.80	31.67	84.55
16	0.247	0.4338	<u>0.3022</u>	0.1406	<u>0.1499</u>	<u>0.1303</u>	0.2191	99.98	75.51	51.22	73.76	21.78	52.06	32.45	85.04
17	0.2381	0.5274	0.3246	0.1323	0.1527	0.1453	0.2092	99.99	77.48	53.18	73.96	24.51	52.06	33.46	85.37
18	0.2419	0.5413	0.3375	<u>0.1322</u>	<u>0.1513</u>	0.1454	0.2091	99.99	78.05	60.13	76.74	25.04	52.39	34.68	85.40
19	0.2059	0.4018	0.3062	0.1358	0.1613	0.1492	0.2079	99.99	78.59	65.54	80.46	25.80	54.09	34.68	85.41
20	<u>0.2044</u>	0.3889	0.3037	0.1372	<u>0.1605</u>	0.1405	0.1930	99.99	78.59	66.07	82.11	25.88	56.74	39.70	86.63
21	0.2077	<u>0.3871</u>	<u>0.2996</u>	<u>0.1367</u>	0.1608	0.1379	<u>0.1883</u>	<u>99.99</u>	<u>79.92</u>	<u>68.92</u>	<u>82.68</u>	<u>26.90</u>	<u>57.24</u>	<u>43.19</u>	<u>88.28</u>
22	0.2121	0.4131	0.3020	0.1367	0.1582	<u>0.1340</u>	0.1887	99.99	80.24	68.94	84.65	28.33	63.26	48.19	88.48

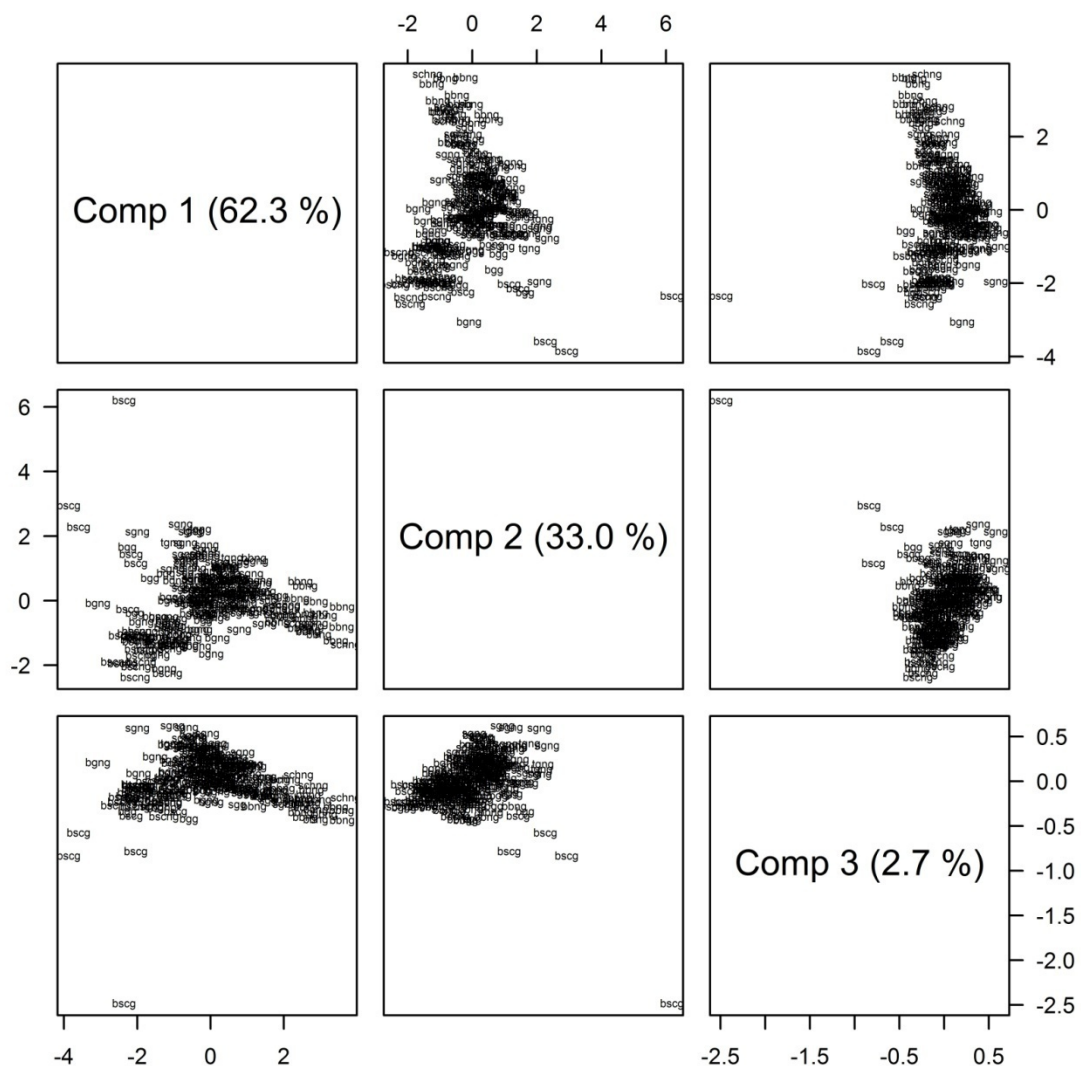


Figure 3.16 Scores plot for the first three components of a PLSR model predicting seven patch types with all points in the calibration dataset included. Labels ending in “ng” indicate a patch with no gypsum while those ending in “g” indicate a patch with gypsum present.

with predictions using calibration data accounting for only 48% of the variability in these patches (Table 3.6). Both these patch types have only 4 samples in the calibration data, but the model also struggles to predict both tall grass (8 samples) and bare grass patches (32 samples). The model found it easier to predict *S. plumosum* patches with a 5-component model accounting for 70% of the variability in calibration predictions. As *S. plumosum* has only 22 samples in the model compared to the 32 for sparse grass, the weakness in the model cannot be entirely due to some patches having very few samples in the model. An

examination of the scores plot for the first three components of this model showed that there were some extreme outliers in the model (Figure 3.16) which were predominantly gypsum-related patches. On this basis it was decided to derive a model with all seven patches but without the gypsum-influenced patches. Removing the gypsum related patches resulted in a slight improvement in the model but the model was still weak overall (results not presented). For instance, a model with twenty components resulted in predictions with calibration data that could account for only 30 and 47% of the variability between predicted and measured values for grass and *Schoenoplectus* patches respectively. It was therefore decided to derive the simplest model possible on the basis of patch types that reflected distinct, unambiguous spectral patterns and then build more complex models to elucidate where the weaknesses in the seven-patch model lay. This was a three-patch model predicting biological soil crust, sparse grass and *S. plumosum* patches with sample points for all other patch types and gypsum-influenced patches removed from the calibration and validation data sets. The resulting calibration data set consisted of 149 samples composed of 24 biological soil crust, 103 sparse grass and 22 *S. plumosum* patches. The three-patch model produced very strong results (not shown) with no obvious outliers in the scores plot (not shown) where the three patch types sorted into distinct groups. The first component, which accounts for 70% of the variability in the spectral data, separates *S. plumosum* patches from the other two patch types. The second component accounts for 27% of the spectral variability and separates sparse grass patches from biological soil crust patches (not shown).

The absolute minimum RMSEP for all three patch types was considerably lower for the three-patch model when compared to the previous seven-patch models with all points or with gypsum associated patches removed. For instance, the minimum RMSEP in the three-patch model for *S. plumosum* was 0.1107 for a 14-component model, whereas the seven-patch model with gypsum had a minimum RMSEP of 0.1593 with a 73-component model, and that for the seven-patch model without gypsum was 0.1261 with a 41-component model (results not shown). For sparse grass patches the 3-patch model produced a minimum RMSEP of 0.1549 with a 16-component model, whereas that for the seven-patch model without gypsum influenced patches was 0.3034 for a 14-component model (results not shown), and the all-point seven-patch model had a minimum RMSEP of 0.2996 for a 21-

component model (Table 3.6). Likewise, biological soil crust patches in the three-patch model had a minimum RMSEP of 0.1041 with an 18-component model, whereas that for the seven patch model without gypsum was 0.1878 for a 13-component model (both results not shown), and the minimum RMSEP for the seven-patch all-points model was 0.2030 for a 13-component model (Table 3.6).

Predictions with a three-patch model and calibration data were also significantly stronger for the three patch types than they were for the same patch types in the seven-patch models. For instance calibration predictions with a 5-component three-patch model resulted in r^2 values of 0.89, 0.83 and 0.84 ($P < 0.0001$ in all cases) for biological soil crust, sparse grass and *S. plumosum* patches respectively (Figure 3.17). For a 5-component, seven-patch, all points model, the same patch types recorded r^2 values of 0.59, 0.49 and 0.70 ($P < 0.0001$; Table 3.6) and 0.58, 0.58 and 0.72 respectively for a 5-component, seven-patch model with no gypsum (results not shown). Predictions with the three-patch model give clear separation between the three patch types with both calibration and validation data. Predictions with a 5-component model with validation data versus measured field data gave r^2 values of 0.80, 0.80 and 0.84 ($P < 0.0001$ in all cases) for biological soil crust, sparse grass and *S. plumosum* patches respectively (results not shown). The only blemish with the 5-component three-patch model is that one sparse grass patch fell on the midpoint (the 0.5 line) of predictions with calibration data (Figure 3.17) which means, in theory at least, it belongs to both the sparse grass patch type and the rest, as a perfect sparse grass patch should be predicted with a value of 1 in the model and all the other patches should have a value of zero. However, from an application perspective, there was clear separation between sparse grass patch predictions and those for the other patch types. Furthermore, all models evaluated with more components, including a 6-component model (not shown), moved this point and all other sparse grass patches, well above the theoretical 0.5 value as a minimum limit for predicting a patch type.

The effect of gypsum was tested by running the same three-patch model but including all the gypsum affected patches that occurred in some biological soil crust and sparse grass patches. The results were unexpected as coefficients of determination with calibration data improved slightly for biological soil crust patches and decreased slightly for sparse grass and

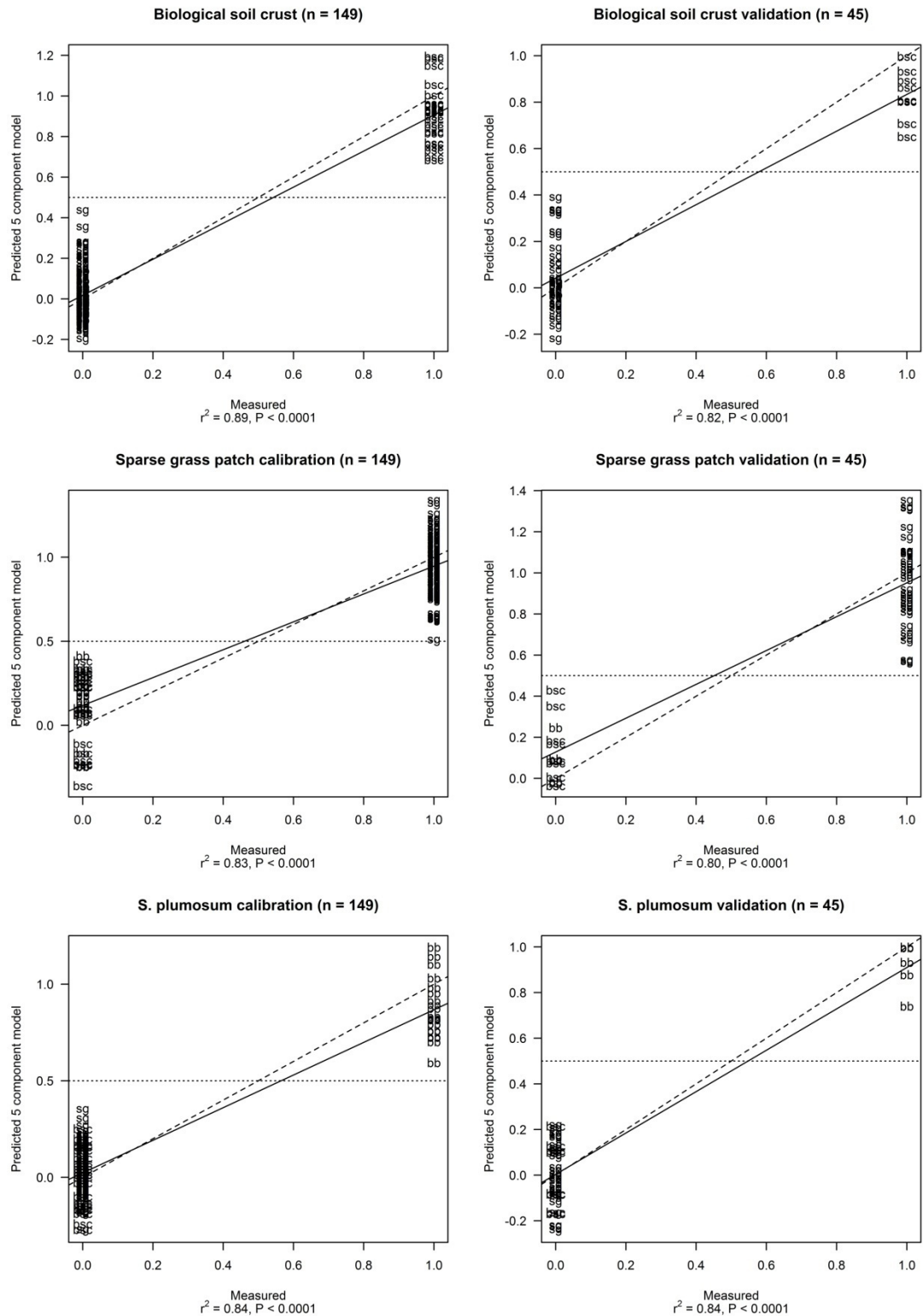


Figure 3.17 Calibration (left column) and validation (right column) predictions for three patch types from Vaal River using a 5-component model. The dashed line represents the 1:1 fit if predicted values matched measured values perfectly. The solid line represents the best simple regression fit between predicted and measured values, and the dotted horizontal line indicates the theoretical midpoint between predicted patch type and the rest of the patches.

S. plumosum patches (results not shown). For biological soil crusts, a 5-component model with all gypsum patches and calibration data increased the r^2 values from 88.8% ($P < 0.0001$) without gypsum to 89.6% ($P < 0.0001$; results not shown) with gypsum. However, for sparse grass patches, the r^2 for calibration predictions decreased from 82.9% ($P < 0.0001$) without gypsum to 81.3% ($P < 0.0001$) and similarly for *S. plumosum*, the calibration r^2 fell from 84.5% ($P < 0.0001$) without gypsum to 79.5% with gypsum (results not shown). Examination of the scores for this three-patch model with gypsum patches showed the same set of extreme outliers as did the original seven-patch model (results not shown). Validation predictions showed that despite the effect of gypsum, all three patches could be clearly separated from the other two patch types (results not shown). So despite some gypsum patches producing outliers in the model, gypsum did not have major negative effects on a three-patch model.

Having found the simplest functional model that achieves the aim of separating the basic patch types that define the extremes and center of the continuum of patch types at Vaal River, namely biological soil crust, sparse grass and *S. plumosum* patches, the next step was to test the other patch types individually in four-patch models together with the three patch types included in the three patch models just discussed. Gypsum influenced samples were not included in the model data sets. Adding the bare grass patches to a model including biological soil crust, sparse grass and *S. plumosum* patches slightly reduced the RMSEP (0.1482 to 0.1436) but also slightly reduced the r^2 values (0.84 to 0.82) for *S. plumosum* patch predictions with a 5-component model (results not shown). However, the bare grass patches had more impact on the statistics for sparse grass patches and especially biological soil crust patches indicating strong spectral confusion between these three patch types in the PLSR modelling. The RMSEP for sparse grass patches increased from 0.2021 for a 5-component three-patch model to 0.2683 for a 5-component four-patch model and the calibration r^2 decreased from 0.83 to 0.73 (results not shown). The RMSEP for biological soil crusts increased from 0.1278 for a 5-component three-patch model to 0.2246 for a 5-component four-patch model, and the calibration r^2 for the same models fell from 0.89 to 0.59 respectively (results not shown). For bare grass patches, the model with four patch types produced an absolute RMSEP minimum of 0.3100 which is quite high when one considers the model is predicting values between 0 and 1. Calibration r^2 values for bare

2778 grass patches was also low with the 5-component four-patch model able to account for only
2779 23% of the variation in bare grass patches, and the model only passing 50% explained
2780 variance for bare grass patches with a 13-component model (53%; results not shown).

2781 Grass patches were only weakly modelled in a four-patch model with a 5-component model
2782 explaining only 15% of the variability in calibration predictions for grass patches and it
2783 required an 18-component model to achieve a coefficient of determination greater than
2784 50% ($r^2 = 0.53$, $P < 0.0001$; results not shown). However, four-patch models with tall grass all
2785 produced stronger coefficients of determination than did four-patch models with grass
2786 patches as an 8-component model explained 51% of the variation in predictions for tall grass
2787 with calibration data (results not shown). In predicting values with validation data, both
2788 grass and tall grass four-patch models give good predictions for biological soil crust and *S.*
2789 *plumosum* patches, but struggle to separate grass and tall grass patches from sparse grass
2790 patches (results not shown). When predicting sparse grass patches with a four-patch model
2791 including grass patches, both calibration and validation predictions for sparse grass patches
2792 were strong ($r^2 = 0.80$, $P < 0.0001$ in both cases), yet in both cases grass patches overlap
2793 with the sparse grass predictions. Furthermore, predictions for grass patches were very
2794 weak and cannot separate grass patches from the rest in both calibration and validation
2795 predictions ($r^2 = 0.15$, $P < 0.0001$ and $r^2 = 0.08$, $P = 0.058$ respectively; results not shown).
2796 Tall grass patches had similar kinds of problems as grass patches although the predictions
2797 were somewhat stronger. This suggests that the sparse grass, grass and tall grass patch
2798 types may be better combined into a single category.

2799 Models for *Schoenoplectus* in a four-patch system where the other patch types were
2800 biological soil crust, sparse grass and *S. plumosum* were also weak with a 5-patch model
2801 with calibration data only achieving an r^2 of 0.14, and to explain 50% of the variation in
2802 calibration predictions required a 13-component model (results not shown). A 10-
2803 component model produced stronger coefficients of determination with validation data ($r^2 =$
2804 0.43 , $P < 0.0001$) then with calibration data ($r^2 = 0.38$, $P < 0.0001$) when predicting
2805 *Schoenoplectus* patches (results not shown) but this is most likely an outcome of there being
2806 only one sample in the validation dataset. What is important in this four patch model
2807 including *Schoenoplectus*, despite the weakness in the model, is that predictions for all four

2808 patch types could be clearly separated from the rest of the patches in each of the four
2809 predictions by the model.

2810 The final step is to put all the results from the four-patch models together in a single model
2811 that is ecologically meaningful and allows accurate prediction of the patch types. It is clear
2812 from the four-patch models that bare grass patches are a transition type between biological
2813 soil crust and sparse grass patches in terms of degree of grass cover. As biological soil crusts
2814 and bare grass patches are at the low end of the ecological functionality continuum and may
2815 represent sequential stages of successional recovery from disturbance in these grasslands
2816 (Macdonald, 1978), it is logical to combine them into a single patch type. The inability to
2817 separate sparse grass, grass patches and tall grass patches suggests that these can be
2818 combined into a single class. *Schoenoplectus* showed affinities with both sparse grass and *S.*
2819 *plumosum* but it did separate from the other three patch types in a four-patch model
2820 despite the model showing weak properties when predicting *Schoenoplectus* patches. Both
2821 six-patch and five-patch models were tested (results not shown) but they suffered from the
2822 same weakness outlined for the four-patch models presented above. Models combining
2823 bare grass patches with sparse grass patches were also tested but these performed
2824 substantially poorer than the final model. The final model presented to predict patch types
2825 is therefore a four-patch model with biological soil crust and bare grass patches combined
2826 into a single patch type (bsbg). The second patch type in this four-patch model was a
2827 combination of sparse grass, grass patches and tall grass patches (sgt), with *Schoenoplectus*
2828 and *S. plumosum* left as separate patch types (Table 3.7).

2829 For the four-patch combination with all gypsum affected patches removed, there are three
2830 possible models: a 5-component model, a 10-component model and a 14-component model
2831 (Table 3.7). The 5-component model was chosen because it was the first RMSEP local
2832 minimum for *S. plumosum* and ended a continuous sequence of > 10% reductions in RMSEP
2833 per component added to the model for *S. plumosum* (Table 3.7). This matched a similar
2834 reduction in RMSEP per added component for the sparse-grass-tall grass combination. The
2835 10-component model was chosen as it matched a long sequence of > 2% reductions in
2836 RMSEP per added component beginning with the 7th component and ending at the 10th
2837 component for *S. plumosum*. The 14-component model was chosen as it represented a local
2838 minimum for all four patch types which also happened to be absolute minima for the

2839 biological soil crust-bare grass combination and the sparse-grass-tall grass combination
2840 (Table 3.7). However, it should be noted in these models that after the second component,
2841 no addition of another component for the biological soil crust-bare grass combination
2842 resulted in a decrease in RMSEP of $\geq 2\%$ with the exception of the 14th component which
2843 reduced the RMSEP by 3% (results not shown). Likewise, for the sparse-grass-tall grass
2844 combination, only the 12th component made a 5% reduction in RMSEP after the first six
2845 components. For the *Schoenoplectus* patch type, no component made a $> 2\%$ decrease in
2846 RMSEP when added to the model with the exception of the 12th (4.6%) and 14th (2.1%)
2847 components (results not shown). Apart from *Schoenoplectus* patches, calibration predictions
2848 for this four-patch combination model were generally quite strong although they weren't so
2849 strong for validation predictions. Of the three chosen models: 5-, 10- and 14-component
2850 models, the 14-component models achieved the strongest r^2 values for validation data ($r^2 =$
2851 0.75 , $P < 0.0001$ for biological soil crust and bare grass combination patches; $r^2 = 0.73$, $P <$
2852 0.0001 for sparse-grass-tall grass combination; $r^2 = 0.34$; $P < 0.0001$ for *Schoenoplectus*
2853 patches and $r^2 = 0.83$, $P < 0.0001$ for *S. plumosum* patches) and the best separation between
2854 predicted patch type and the rest of the patches (Figure 3.18).

2855 However, there is still overlap between predictions of biological soil crust-bare grass
2856 combination patches and sparse-grass-tall grass combination patches (Figure 3.18). For
2857 instance, when predicting the biological soil crust-bare grass combination there are four
2858 bare grass patches that overlap with the rest of the patch types in the calibration
2859 predictions and two bare grass patches in the validation predictions that overlap with
2860 prediction values below 0.5 (Figure 3.18 and 3.19). It is noticeable that these bare grass
2861 patches have a fair degree of grass cover although often the species are different to the
2862 species occurring in the grassland matrix that makes up sparse grass and grass patches
2863 (Figure 3.19). Likewise in predicting the sparse-grass-tall grass combination, there are six
2864 bare grass patches that overlap the lower limit for the sparse-grass-tall grass combination
2865 calibration predictions and three bare grass patches that overlap in the validation
2866 predictions (Figure 3.18). These bare grass patches that overlap with predictions of the
2867 sparse-grass-tall grass combination include the same patches that overlapped when
2868 predicting the biological soil crust-bare grass combination. Between the 0.5 midpoint and
2869 the lower limit for the sparse-grass-tall grass combination predictions are two sparse grass

2870 patches that have low grass cover and relatively high soil surface exposure (Figure 3.19).
2871 Thus, it is clear that this area of overlap between predictions of bare grass and sparse grass
2872 patches represents a transition zone of relatively low grass cover.

2873 Predictions for *Schoenoplectus* patches are considerably lower than the theoretical value of
2874 1, being approximately 0.4 for both calibration and validation data (Figure 3.18). However
2875 these predicted values are well separated from predictions for all the other patches which
2876 tightly bunched around zero with an upper limit of approximately 0.25 for calibration data
2877 and 0.16 for validation data (Figure 3.18). For *S. plumosum* predictions, the 14-component
2878 model clearly separates *S. plumosum* patches from the rest of the patches with all *S.*
2879 *plumosum* prediction values above 0.5 and values for the rest of the patches below 0.5
2880 (Figure 3.18). However, the similarity between *S. plumosum* spectra and *Schoenoplectus*
2881 patches as illustrated in Figure 3.15 is also apparent as *Schoenoplectus* patches congregate
2882 at the upper limit of predictions for the rest of the patches (Figure 3.18).

2883 Thus, the four-patch combination 14-component model is a viable model with the only
2884 question being what value to select in separating the sparse grass patches from the bare
2885 grass patches. Furthermore, bringing the gypsum affected patches back into the model did
2886 not significantly undermine the four-patch combination model and in some cases slightly
2887 improved the prediction statistics (results not shown). For instance, including all sample
2888 points from Vaal River resulted in the r^2 value for biological soil crust-bare grass patches
2889 increasing by 1 percentage point for a 14-component model with validation data, staying
2890 the same for sparse-grass-tall grass patches, falling by 5 percentage points for
2891 *Schoenoplectus* patches and by 2 percentage points for *S. plumosum* patches (Table 3.8 and
2892 Figure 3.20).

Table 3.7 Cross-validated RMSEP using LOO estimates for calibration models predicting patch types at Vaal River from ground-based hyperspectral measurements using combined biological soil crust and bare grass (bsbg), combined sparse, grass and tall grass (sgt), *Schoenoplectus* patches (sch) and *S. plumosum* (bb) patches. All patches that had visible gypsum were removed prior to analysis. The table presents the coefficient of determination (r^2) representing the explained variance for each variable as one adds components to the model. Values in bold underlined italics reflect models with local RMSEP minima for that patch type. The bold underlined values reflect variation explained for that variable by two selected models.

Component	Cross-validated RMSEP				Calibration r^2				
	bsbg	sgt	sch	bb	Spectra	bsbg	sgt	sch	bb
1	0.2853	0.4802	0.1398	0.2690	58.58	61.05	8.74	5.19	30.71
2	0.2334	0.3554	0.1380	0.2375	97.16	73.85	49.55	10.17	45.71
3	0.2212	0.3058	<u>0.1375</u>	0.2038	98.30	76.94	63.71	11.55	60.88
4	0.2200	0.2990	0.1378	0.1806	99.39	77.34	65.43	11.56	69.38
5	0.2169	0.2755	0.1379	<u>0.1609</u>	<u>99.64</u>	<u>78.40</u>	<u>71.43</u>	<u>11.83</u>	<u>76.09</u>
6	<u>0.2107</u>	0.2649	0.1370	0.1625	99.82	80.02	73.88	14.60	76.10
7	0.2110	<u>0.2610</u>	<u>0.1365</u>	0.1556	99.87	80.22	75.11	16.59	79.10
8	0.2119	0.2613	0.1367	0.1477	99.91	80.39	75.26	18.26	82.19
9	0.2163	0.2636	0.1377	0.1407	99.92	80.42	75.74	20.90	86.33
10	0.2187	0.2629	0.1393	0.1375	<u>99.93</u>	<u>81.83</u>	<u>77.97</u>	<u>22.53</u>	<u>86.33</u>
11	0.2183	0.2579	0.1382	0.1374	99.95	82.29	79.64	27.93	86.33
12	<u>0.2162</u>	<u>0.2445</u>	0.1319	0.1345	99.96	82.30	80.72	33.75	86.83
13	0.2211	0.2529	0.1318	0.1332	99.97	82.93	82.26	33.90	87.22
14	<u>0.2143</u>	<u>0.2520</u>	<u>0.1290</u>	<u>0.1322</u>	<u>99.98</u>	<u>83.01</u>	<u>82.31</u>	<u>40.69</u>	<u>87.39</u>
15	0.2220	0.2598	0.1301	0.1334	99.98	85.36	84.63	40.72	87.46
16	0.2244	0.2662	0.1292	0.1342	99.98	86.73	86.16	44.27	88.68
17	<u>0.2224</u>	<u>0.2635</u>	0.1287	<u>0.1325</u>	99.98	87.82	87.36	44.66	88.68
18	0.2292	0.2659	0.1284	0.1345	99.99	88.96	88.63	47.95	89.02
19	0.2314	0.2693	<u>0.1227</u>	<u>0.1304</u>	99.99	90.41	89.32	48.32	89.04
20	<u>0.2303</u>	0.2716	0.1237	0.1341	99.99	91.57	90.18	49.38	89.35
21	0.2342	0.2764	0.1248	0.1340	99.99	91.79	92.16	56.47	89.46
22	0.2375	0.2798	<u>0.1238</u>	0.1334	99.99	92.23	93.02	58.86	90.90
23	0.2402	0.2821	0.1244	<u>0.1320</u>	99.99	92.40	93.65	60.94	92.57
24	0.2491	0.2901	0.1239	0.1321	99.99	93.37	94.21	61.40	92.87
25	0.2535	0.2943	0.1239	0.1318	99.99	94.08	94.76	64.68	93.60
26	0.2578	0.2979	0.1225	<u>0.1296</u>	99.99	95.33	95.29	69.53	93.87
27	0.2642	0.3052	0.1215	0.1300	99.99	95.93	95.78	70.31	94.04
28	0.2643	0.3062	0.1192	0.1313	99.99	96.32	96.05	77.01	94.33
29	0.2650	0.3084	0.1184	0.1315	99.99	96.32	96.13	78.92	95.72
30	0.2703	0.3142	0.1179	0.1313	99.99	96.69	96.67	79.34	95.72

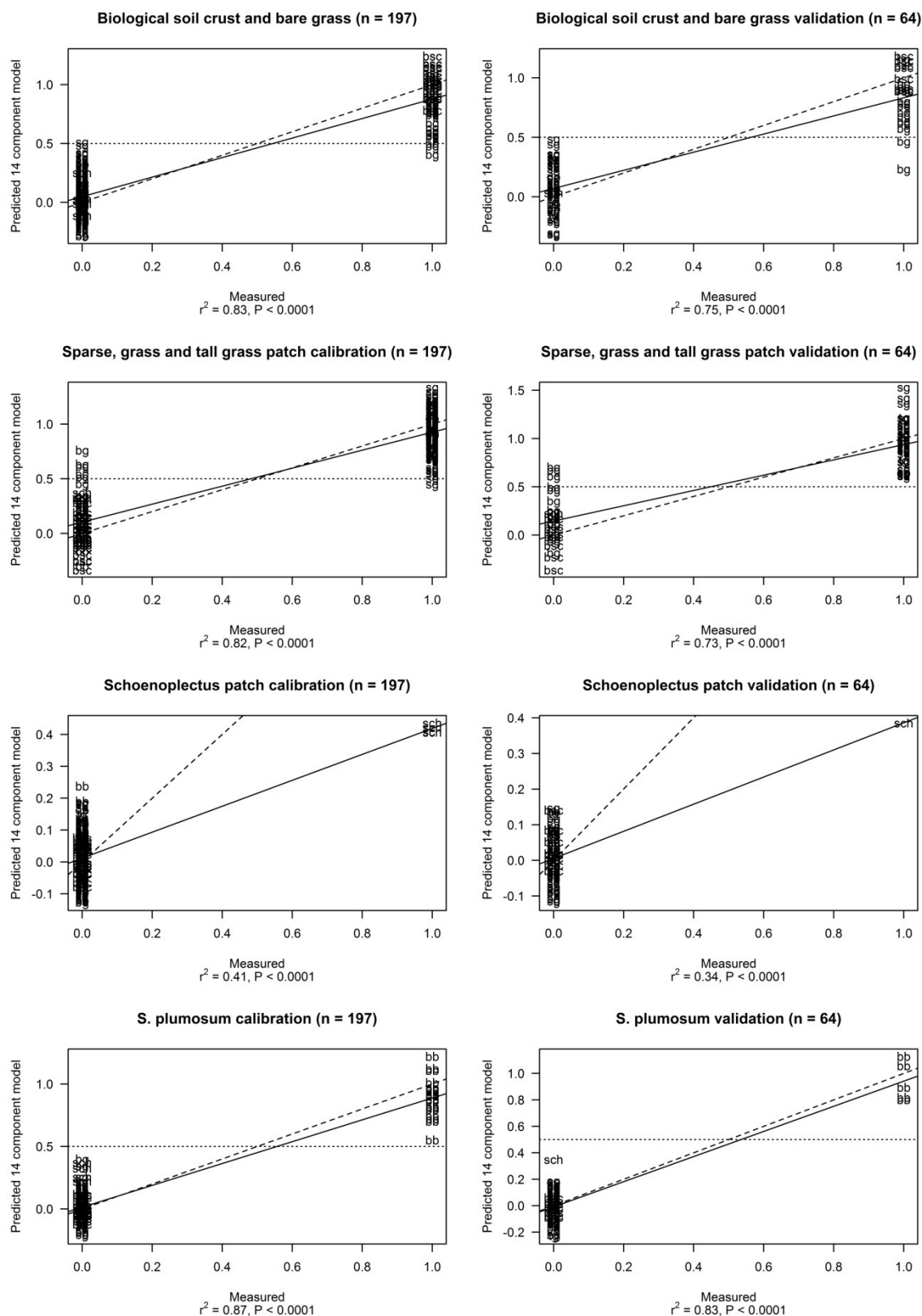


Figure 3.18 Predictions for a 14-component model for Vaal River with no gypsum affected patches and four patch types: biological soil crust and bare grass combined (bsbg), sparse, grass and tall grass patches combined (sgt), *Schoenoplectus* patches (sch) and *S. plumosum* patches (bb).



Figure 3.19 Bare grass patches at Vaal River that overlap with predictions for other patch types. Calibration patches: (a) vr1bg3, (b) vr1bgt75n, (c) vr10bg1 and (d) vr14bg1) and validation patches: (e) vr9bg5 and f) vr10bg2. Sparse grass patches (g) vr7sg3 and (h) vr8sgt15n at the lower limit of the sparse-grass-tall grass combination predictions with calibration data.

2916 **Table 3.8** Prediction calibration and validation results for three forms of the four-patch combination model: (1) All gypsum affected
2917 patches removed, (2) Five outliers removed, and (3) All points including gypsum affected patches included in calibration and
2918 validation data. Patch acronyms are bsbg = biological soil crust-bare grass combination, sgt = sparse-grass-tall grass
2919 combination, sch = *Schoenoplectus* patches and bb = *S. plumosum* patches.

Component Data		bsbg r ² Probability		sgt r ² Probability		sch r ² Probability		bb r ² Probability	
All gypsum removed									
5	Calibration (n = 197)	0.78	< 0.0001	0.71	< 0.0001	0.12	< 0.0001	0.76	< 0.0001
	Validation (n = 64)	0.70	< 0.0001	0.60	< 0.0001	0.17	0.0007	0.69	< 0.0001
10	Calibration (n = 197)	0.82	< 0.0001	0.78	< 0.0001	0.23	< 0.0001	0.86	< 0.0001
	Validation (n = 64)	0.74	< 0.0001	0.66	< 0.0001	0.21	0.0001	0.81	< 0.0001
14	Calibration (n = 197)	0.83	< 0.0001	0.82	< 0.0001	0.41	< 0.0001	0.87	< 0.0001
	Validation (n = 64)	0.75	< 0.0001	0.73	< 0.0001	0.34	< 0.0001	0.83	< 0.0001
4 gypsum outliers removed									
7	Calibration (n = 215)	0.81	< 0.0001	0.74	< 0.0001	0.12	< 0.0001	0.73	< 0.0001
	Validation (n = 71)	0.75	< 0.0001	0.66	< 0.0001	0.15	0.0009	0.63	< 0.0001
11	Calibration (n = 215)	0.83	< 0.0001	0.80	< 0.0001	0.25	< 0.0001	0.84	< 0.0001
	Validation (n = 71)	0.75	< 0.0001	0.73	< 0.0001	0.21	< 0.0001	0.80	< 0.0001
13	Calibration (n = 215)	0.84	< 0.0001	0.81	< 0.0001	0.29	< 0.0001	0.86	< 0.0001
	Validation (n = 71)	0.76	< 0.0001	0.73	< 0.0001	0.25	< 0.0001	0.80	< 0.0001
All points									
5	Calibration (n = 219)	0.79	< 0.0001	0.71	< 0.0001	0.09	< 0.0001	0.71	< 0.0001
	Validation (n = 72)	0.70	< 0.0001	0.61	< 0.0001	0.14	0.001	0.59	< 0.0001
7	Calibration (n = 219)	0.81	< 0.0001	0.75	< 0.0001	0.10	< 0.0001	0.71	< 0.0001
	Validation (n = 72)	0.75	< 0.0001	0.66	< 0.0001	0.13	0.002	0.60	< 0.0001
14	Calibration (n = 219)	0.84	< 0.0001	0.81	< 0.0001	0.29	< 0.0001	0.86	< 0.0001
	Validation (n = 72)	0.76	< 0.0001	0.73	< 0.0001	0.25	< 0.0001	0.81	< 0.0001

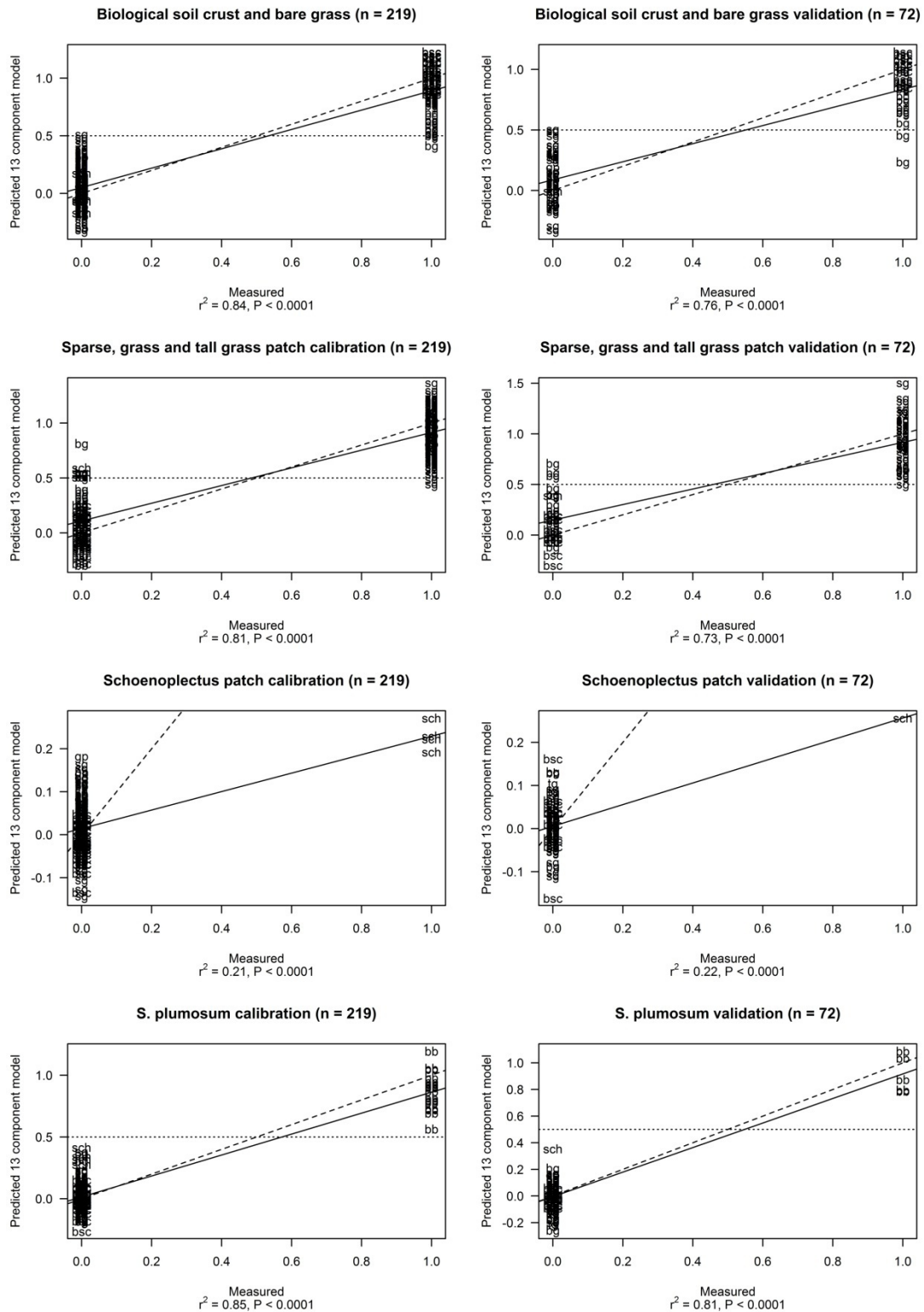


Figure 3.20 Predictions for a 13-component model for Vaal River with all sample points and four patch types: biological soil crust and bare grass combined (bsbg), sparse, grass and tall grass patches combined (sgt), *Schoenoplectus* patches (sch) and *S. plumosum* patches (bb).

3.3.3 Predicting the LFA Stability Index

The models derived to predict the LFA indices of soil stability, infiltration and nutrient cycling give mixed results. Table 3.9 presents the model statistics for a model predicting stability at West Wits and Table 3.10 presents the model predicting stability at Vaal River. As there is now a single numerical response variable, it is feasible to present the r^2 values for both calibration and validation predictions for the first thirty components of the models and these show that selecting the best model is not as straight forward as using local minima to indicate a good model (Mevik and Wehrens, 2007, Mevik and Cederkvist, 2004), or that a component should only be added if it results in the reduction of the RMSEP by 2% (Kooistra *et al.*, 2001). The West Wits model predicting the stability index reaches its first local minima at four components with a RMSEP of 3.018% and an absolute minimum RMSEP of 2.826% at seven components (Table 3.9). These RMSEP values are less than half the standard deviation (6.316%) for the West Wits stability index. Calibration coefficients of determination for the West Wits LFA stability index are 0.81 for the 4-component model (Figure 3.21a) and 0.89 for the 7-component model while the validation coefficients are 0.63 ($P = 0.00013$; Figure 3.21a) and 0.51 ($P = 0.0063$) respectively (Table 3.9). Adding the second, third and fourth components to the model results in a decrease of 29.7%, 15.3% and 12.3% respectively for the RMSEP. However, when examining the validation coefficients of determination, the best model is an 11-component model ($r^2 = 0.69$, $P = 0.00001$; Table 3.9) and the 3-component model results in a slightly stronger validation coefficient of determination ($r^2 = 0.64$, $P = 0.00004$) than did the 4-component model (Table 3.9). Under the assumption of parsimony to avoid over-fitting the model, and taking into account local minima and the 2% rule, the 4-component model is the simplest model with fairly robust predictions despite the 11-component model giving slightly better predictions with validation data. It should be noted that although the probability of these predictions is significant for all components of the model, the values of the probabilities are not as low as that obtained for other variables with strong model predictions tested in this study.

In contrast to the West Wits model, the PLSR model predicting stability values from Vaal River is very weak. Calibration coefficients of determination only pass 50% when the model

2954 **Table 3.9** PLSR models predicting LFA stability for West Wits with calibration data (n =
2955 57) and validation data (n = 19).

Component	Calibration RMSEP (%)	Calibration spectral r^2	Stability calibration r^2	Stability validation r^2	Probability
1	5.793	90.91	21.09	27.93	0.02002
2	4.067	94.91	64.18	54.57	0.00030
3	3.443	98.69	74.27	64.10	0.00004
4	<u>3.018</u>	<u>99.30</u>	<u>81.24</u>	63.48	0.00004
5	3.024	99.63	83.33	<u>64.38</u>	<u>0.00004</u>
6	2.979	99.86	85.88	58.62	0.00013
7	<u>2.826</u>	<u>99.89</u>	<u>88.85</u>	50.64	0.00063
8	3.001	99.92	90.51	57.76	0.00016
9	3.080	99.95	91.69	60.85	0.00008
10	3.124	99.96	93.73	64.95	0.00003
11	3.343	99.96	95.07	<u>69.38</u>	<u>0.00001</u>
12	3.228	<u>99.97</u>	<u>96.11</u>	67.71	0.00002
13	<u>3.210</u>	99.97	97.53	67.66	0.00002
14	3.405	99.97	98.54	65.97	0.00002
15	3.397	99.98	98.74	66.34	0.00002
16	<u>3.249</u>	99.98	99.20	64.69	0.00003
17	3.263	99.99	99.49	64.51	0.00003
18	3.230	99.99	99.57	64.06	0.00004
19	<u>3.225</u>	99.99	99.77	63.70	0.00004
20	3.244	99.99	99.87	64.12	0.00004
21	<u>3.199</u>	99.99	99.93	62.94	0.00005
22	3.205	99.99	99.97	63.70	0.00004
23	3.236	99.99	99.98	63.10	0.00005
24	3.237	99.99	99.99	63.06	0.00005
25	<u>3.240</u>	99.99	100.00	62.76	0.00005
26	3.240	99.99	100.00	62.69	0.00005
27	3.239	99.99	100.00	62.67	0.00005
28	3.239	100.00	100.00	62.63	0.00005
29	3.239	100.00	100.00	62.60	0.00005
30	3.239	100.00	100.00	62.60	0.00005

2956

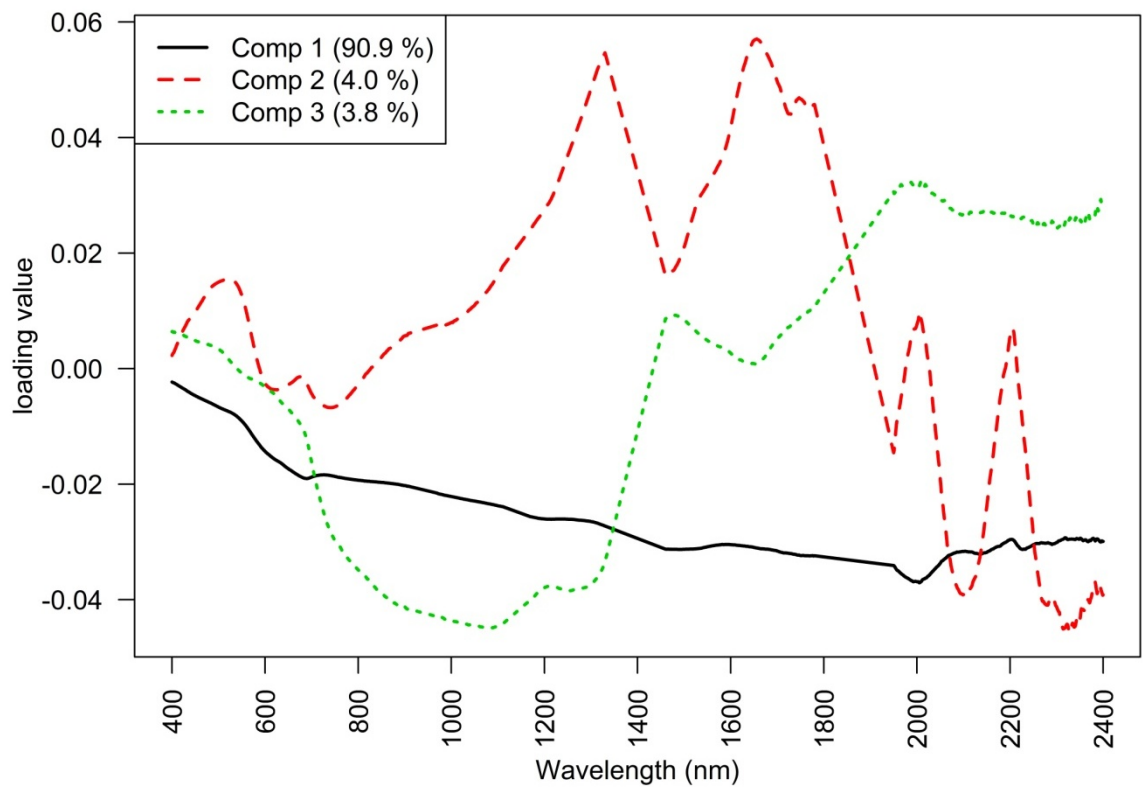
2961 **Table 3.10** PLSR models predicting the LFA stability index for Vaal River from calibration
 2962 data (n = 180) and validation data (n = 59).

Component	Calibration RMSEP (%)	Stability calibration spectral r^2	Stability calibration r^2	validation r^2	Probability
1	3.806	63.52	17.42	21.01	0.0003
2	3.772	96.13	19.46	16.80	0.0013
3	<u>3.656</u>	97.72	27.38	31.97	< 0.0001
4	3.656	98.53	28.83	<u>39.22</u>	< 0.0001
5	<u>3.640</u>	98.78	31.95	29.18	< 0.0001
6	3.677	99.71	32.68	29.73	< 0.0001
7	3.695	99.82	35.05	30.85	< 0.0001
8	<u>3.643</u>	99.85	38.57	33.20	< 0.0001
9	3.709	99.91	39.70	32.25	< 0.0001
10	<u>3.688</u>	99.92	42.22	25.30	< 0.0001
11	3.861	99.93	47.71	29.15	< 0.0001
12	<u>3.777</u>	<u>99.94</u>	<u>52.64</u>	29.71	< 0.0001
13	3.926	99.96	55.68	27.67	< 0.0001
14	<u>3.821</u>	<u>99.96</u>	<u>61.43</u>	31.94	< 0.0001
15	3.920	99.97	64.78	34.28	< 0.0001
16	<u>3.817</u>	<u>99.97</u>	<u>67.51</u>	<u>38.43</u>	< 0.0001
17	3.887	99.98	69.03	35.88	< 0.0001
18	<u>3.876</u>	99.98	73.25	35.51	< 0.0001
19	4.033	99.99	78.38	36.43	< 0.0001
20	4.131	99.99	81.92	38.40	< 0.0001
21	4.205	99.99	84.30	<u>39.79</u>	< 0.0001
22	<u>4.204</u>	99.99	86.43	37.18	< 0.0001
23	4.276	99.99	87.92	37.18	< 0.0001
24	4.392	99.99	89.66	36.67	< 0.0001
25	4.510	99.99	91.05	36.72	< 0.0001
26	4.624	99.99	92.44	35.84	< 0.0001
27	4.807	99.99	93.86	35.08	< 0.0001
28	4.950	99.99	95.59	35.76	< 0.0001
29	5.042	99.99	96.48	36.38	< 0.0001
30	5.090	99.99	97.25	36.36	< 0.0001

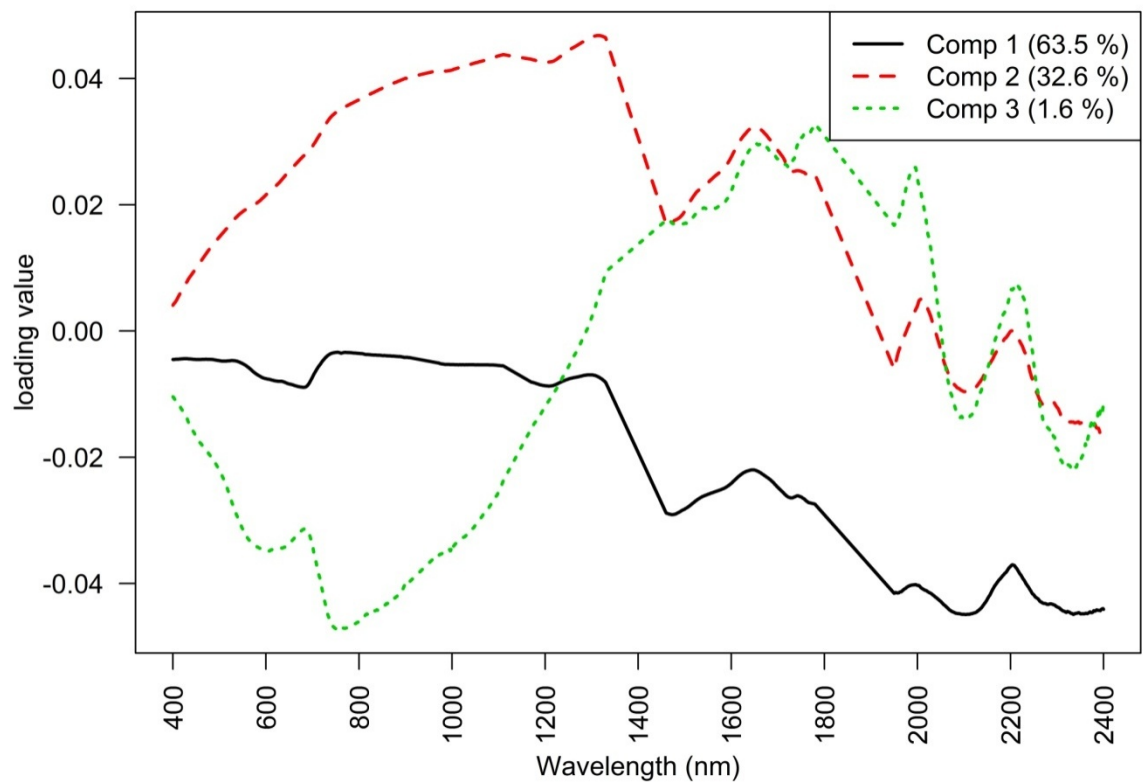
2963

has twelve components with a RMSEP of 3.777 (Table 3.10). The absolute minimum RMSEP was 3.640 for the 5-component model. Validation performed even worse with no version of the model achieving a coefficient of determination greater than 40% (Table 3.10) although this was achieved with a 4-component model (Figure 3.21b). Only the third and twelfth components reduced the RMSEP by greater than 2% (3.07% and 2.17%, respectively). Various transformations did not lead to an improvement in the model, and nor did removing a sample with a very low stability value which was an outlier. In fact removing the outlier produced slightly weaker predictions. Examination of the prediction plots for Vaal River (Figure 3.21b) indicate that there is a large range for predicted values at specific measured values which results in the weakness of model predictions. Furthermore, these predicted values are organized according to patch type with biological soil crust patches predicted with low values and sparse grass and *S. plumosum* patches having high predicted values. This suggests that there is more variability in the spectral measurement of patches than there is in the allocated LFA values.

The first loading for the West Wits PLSR model predicting stability is all negative and in comparison to the second and third loadings, is relatively featureless (Figure 3.22a) although it accounts for 90.9% of the spectral variability and 21.1% of the LFA stability variability in the calibration model. This first loading shows a significant absorption feature located at 689-690 nm which may be related to chlorophyll and the red edge (Cho and Skidmore, 2006, Dawson and Curran, 1998, Rock *et al.*, 1988, Tucker, 1979) and a peak in the SWIR at 2203 nm which may be related to a clay absorption feature (Chabrillat *et al.*, 2002). Correlations between the mean spectra for each patch type and the four loadings used in the PLSR stability model for West Wits are shown in Figure 3.23a. The first loading is most negatively correlated with the bare patches ($r = -0.976$) and bare grass patches ($r = -0.964$), and slightly less negatively correlated with the vegetated patch types: grass patches ($r = -0.871$), sparse grass ($r = -0.839$) and *S. plumosum* ($r = -0.802$; Figure 3.23a), which all suggests that the first loading is a mix of plant and soil features with the soil features slightly more dominant. The second loading shows the inverse pattern with the loading overall being positively correlated with the mean spectra for four of the patch types. This second loading accounted for 4% of the spectral variability and 43.1% of the variability in the LFA stability index for the calibration model and was most correlated with the vegetated patch types: *S. plumosum* ($r =$

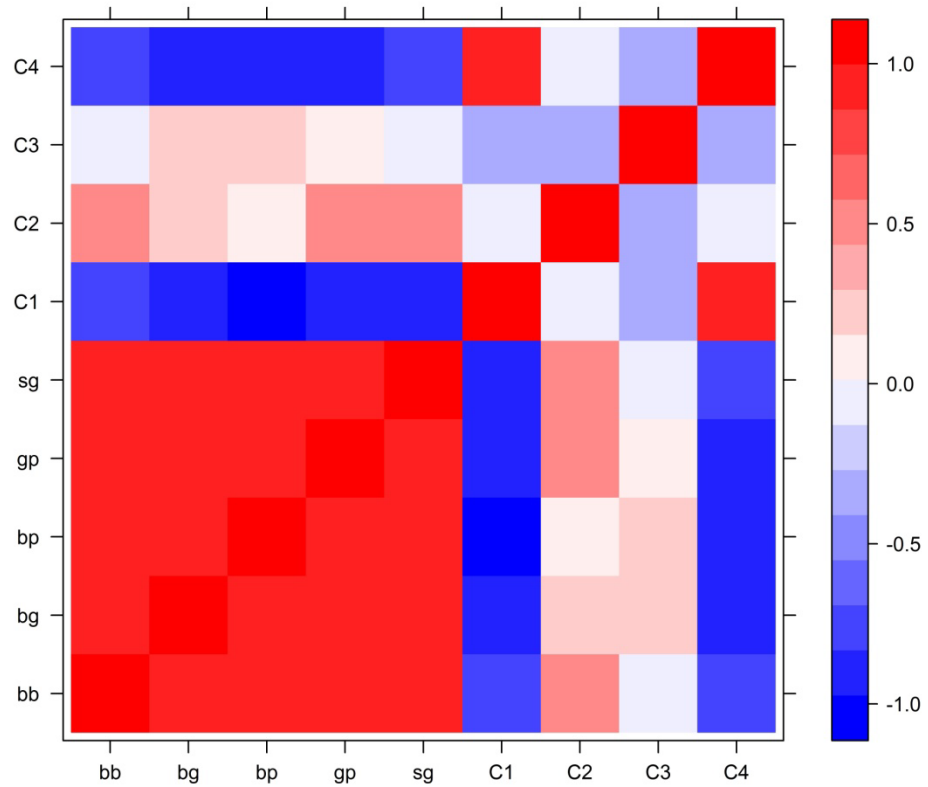


2995 a)

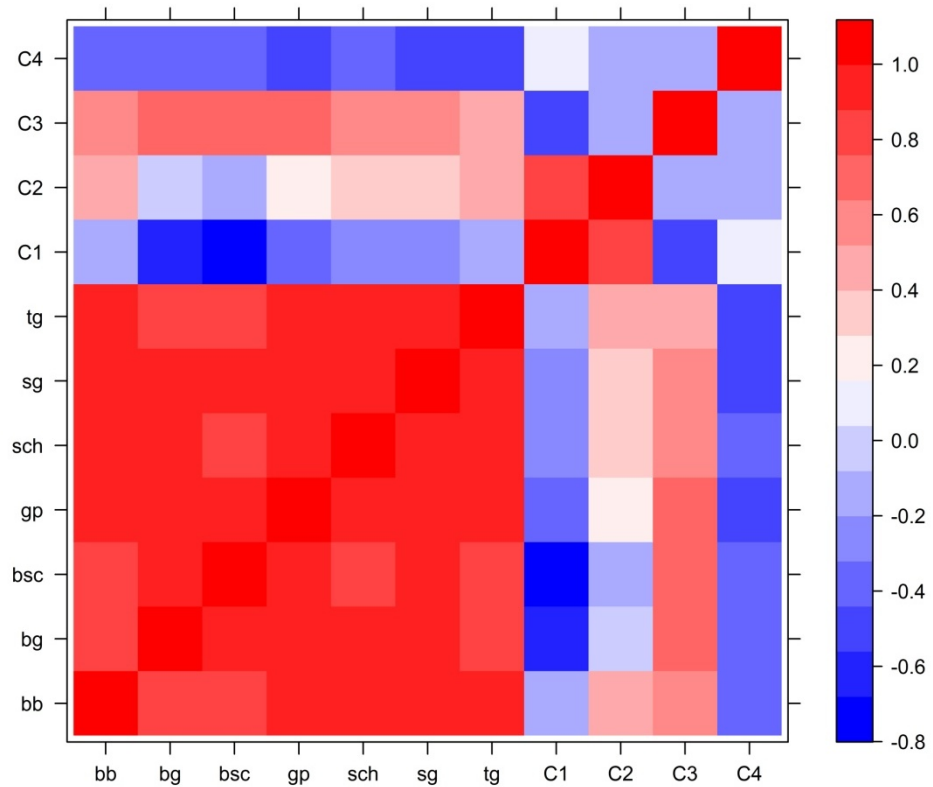


2996 b)

2997 **Figure 3.22** The first three loadings for the models predicting stability at (a) West Wits
2998 and (b) Vaal River



2999 a)



3000 b)

3001 **Figure 3.23** Pearson correlogram between the mean spectra for each patch type and (a)
 3002 the first four loadings for the LFA stability model for West Wits, and (b) the
 3003 first four loadings for the LFA stability model derived from Vaal River.

0.513), sparse grass ($r = 0.469$) and grass patch ($r = 0.465$) and least correlated with the low or no vegetation cover patch types: bare grass ($r = 0.211$) and bare patches ($r = 0.118$; Figure 3.23a). This second loading showed features associated strongly with plants (Figure 3.22a), such as a peak around 550 nm (Gitelson *et al.*, 1996), and large absorption features around 2100 and 2300 nm often associated with cellulose (Daughtry *et al.*, 2004, Daughtry *et al.*, 2003). The third loading, accounting for 10.1% of the LFA stability variability in the calibration model and 3.8% of the spectral variability, was weakly correlated with the high soil patches: bare patch ($r = 0.265$) and bare grass ($r = 0.177$) but had almost no correlation with the vegetated patches (Figure 3.23a). This third loading seemed to be dominated by NIR values between 900 – 1100 nm and SWIR values between 2000 – 2400 nm (Figure 3.22a). The fourth loading, which accounted for only a small percentage of the spectral variability, was very similar to the first loading in its correlations with the mean spectra for each patch type (Figure 3.23a).

The loadings for the PLSR model predicting LFA stability at Vaal River differ significantly from those predicting LFA stability at West Wits where most of the spectral variability was explained by the first component (90.9%; Figure 3.22a). At Vaal River the first component explained 63.5% of the spectral variability in the calibration model and the second component accounted for 32.6% of the spectral variability (Table 3.10 and Figure 3.22b). However, the Vaal River model predicting LFA stability was weak and the first component explained only 17.4% of the variability in the calibration data for the LFA stability measurements with the second accounting for 2.0% and the third 7.9% (Table 3.10). Despite a strong absorption feature centered on 680 nm in the first loading (Figure 3.22b), which may be related to the presence of chlorophyll (Cho and Skidmore, 2006, Tucker, 1979), strong absorption features centered on 2200 and 2300 nm, which may be associated with cellulose (Daughtry *et al.*, 2004, Daughtry *et al.*, 2003), and a minor absorption center at 1728 – 1731 nm which may be related to lignin (Serrano *et al.*, 2002), the first loading is negatively correlated with the mean spectra for all seven patch types found at Vaal River (Figure 3.23b). It is most negatively correlated with the predominantly soil-related patch types: biological soil crust ($r = -0.683$) and bare grass patches ($r = -0.617$) and only weakly, negatively correlated with the vegetation-dominated patch types (Figure 3.23b). The second loading shows stronger positive correlations with the vegetated patches, of which tall grass

($r = 0.471$) and *S. plumosum* patches ($r = 0.452$) are the most correlated, and weak negative correlations with the two soil-dominated patch types: biological soil crust and bare grass (Figure 3.23b). The third loading shows a different pattern with all patch types showing a positive correlation with the third loading with grass patches ($r = 0.729$) and bare grass patches ($r = 0.717$) having the strongest correlation and tall grass patches the weakest ($r = 0.496$; Figure 3.23b).

3.3.4 Predicting the LFA Infiltration Index

In contrast with models predicting the LFA stability index, models predicting the LFA infiltration indices were more robust for both West Wits and Vaal River sites. The first minimum RMSEP for the model predicting LFA infiltration at West Wits was a 5-component model with a RMSEP of 3.893. This was also the absolute minimum RMSEP for the model (Table 3.11) and compares favorably with the standard deviation for field-measured LFA infiltration values ($\bar{x} = 35.313\%$, $SD = 7.998\%$) from West Wits. This 5-component model had strong coefficients of determination for both calibration ($r^2 = 0.83$, $P < 0.0001$) and validation data ($r^2 = 0.75$, $P < 0.0001$; Table 3.11 and Figure 3.24a) and corresponded with the end of a sequence of major reductions in RMSEP per added component (not shown). However the best validation coefficients of determination ($r^2 = 0.77$, $P < 0.0001$) were achieved with a 6-component model (Table 3.11) which resulted in an increase in RMSEP from that of the 5-component model for the extra component of 1.72% and therefore the 5-component model was preferred.

Model predictions for LFA infiltration from Vaal River were stronger than model predictions for LFA stability from Vaal River but not as robust as predictions for LFA infiltration from West Wits. The first minimum for the RMSEP was a 4-component model with a value of 4.597 (Table 3.12). However the absolute minimum RMSEP was 4.330 for a 17-component model (Table 3.12) which is closer to the standard deviation for the field-measured LFA infiltration values ($\bar{x} = 31.595\%$, $SD = 7.258\%$) than the minimum RMSEP for the West Wits model predicting infiltration values was. The calibration coefficient of determination for a 4-component model was 0.63 ($P < 0.0001$; Table 3.12). However, for validation the coefficient of determination was 0.54 ($P < 0.0001$) which had actually fallen slightly from 0.55 ($P < 0.001$) for a 3-component model (Table 3.12). For validation predictions, two models stand

3065 **Table 3.11** PLSR models predicting LFA Infiltration index for West Wits with calibration
3066 data (n = 57) and validation data (n = 19).

3067

Component	Calibration RMSEP (%)	Calibration spectral r^2	Infiltration calibration r^2	Infiltration validation r^2	Probability
1	6.218	91.09	42.53	55.92	0.0002
2	5.324	95.47	60.46	64.45	< 0.0001
3	4.979	98.66	66.63	68.75	< 0.0001
4	4.573	99.26	73.20	71.70	< 0.0001
5	<u>3.893</u>	<u>99.34</u>	<u>82.90</u>	75.29	< 0.0001
6	3.960	99.87	83.21	<u>76.82</u>	<u>< 0.0001</u>
7	4.031	99.89	85.59	72.16	< 0.0001
8	<u>3.987</u>	99.94	86.66	72.78	< 0.0001
9	4.364	99.95	90.13	<u>76.38</u>	<u>< 0.0001</u>
10	4.010	99.95	92.75	73.95	< 0.0001
11	<u>3.946</u>	99.96	93.52	73.53	< 0.0001
12	4.055	99.97	94.82	70.70	< 0.0001
13	<u>4.043</u>	99.97	96.45	74.66	< 0.0001
14	4.450	99.98	97.20	74.34	< 0.0001
15	4.490	99.98	97.74	72.10	< 0.0001
16	4.610	99.98	98.75	73.46	< 0.0001
17	4.701	99.99	99.18	71.78	< 0.0001
18	4.685	99.99	99.47	69.52	< 0.0001
19	4.669	99.99	99.71	68.76	< 0.0001
20	4.573	99.99	99.85	67.55	< 0.0001
21	4.507	99.99	99.91	67.24	< 0.0001
22	4.474	99.99	99.95	67.63	< 0.0001
23	4.447	99.99	99.98	68.35	< 0.0001
24	<u>4.431</u>	99.99	99.99	68.38	< 0.0001
25	4.445	99.99	100.00	68.72	< 0.0001
26	4.451	99.99	100.00	68.68	< 0.0001
27	4.464	99.99	100.00	68.63	< 0.0001
28	4.468	99.99	100.00	68.66	< 0.0001
29	4.467	100.00	100.00	68.65	< 0.0001
30	4.466	100.00	100.00	68.63	< 0.0001

3068

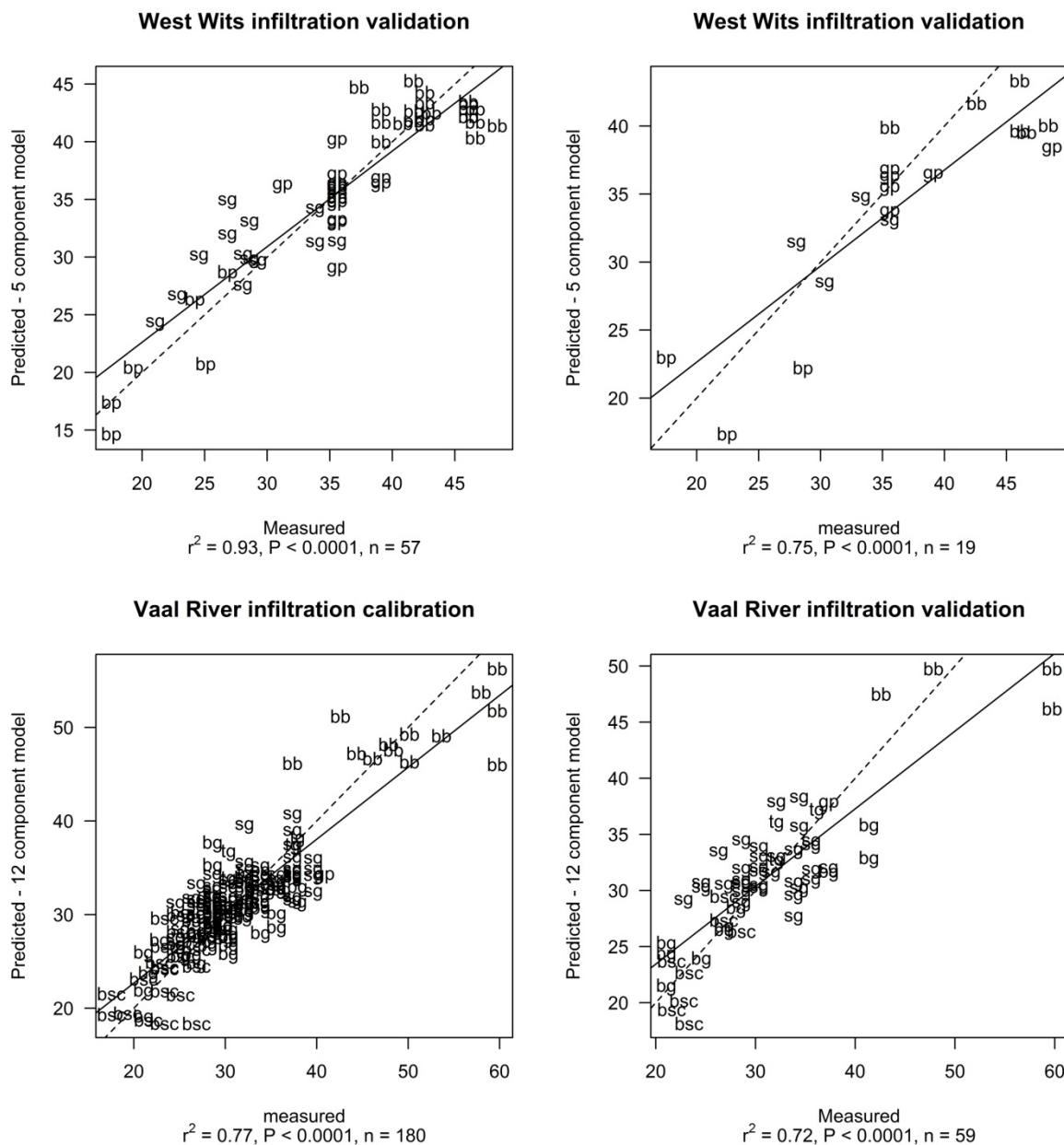


Figure 3.24 PLSR model predictions for the LFA infiltration index with calibration (left column) and validation (right column) data from West Wits (top row) and Vaal River (bottom row).

3073 **Table 3.12** PLSR models predicting the LFA infiltration index for Vaal River from
 3074 calibration data (n = 180) and validation data (n = 59).

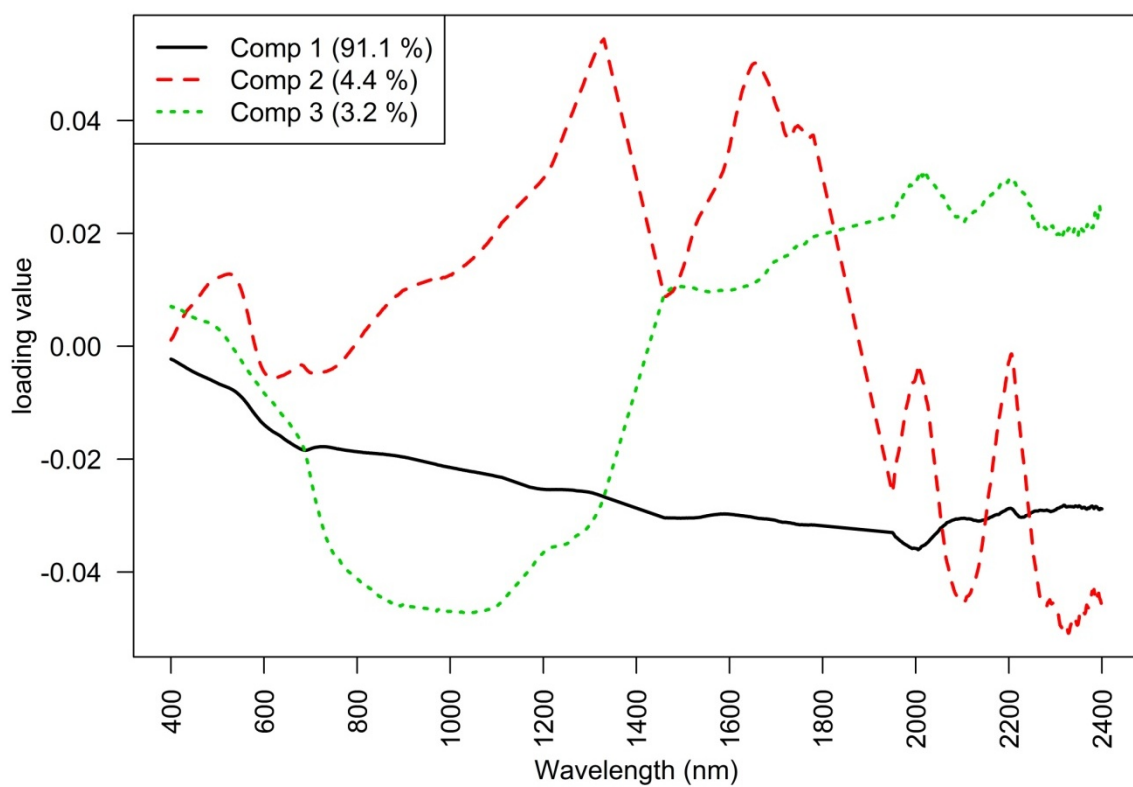
3075

Component	Calibration RMSEP (%)	Infiltration calibration spectral r^2	Infiltration calibration r^2	Infiltration validation r^2	Probability
1	5.681	64.84	40.74	52.46	< 0.0001
2	5.603	96.30	42.98	49.84	< 0.0001
3	5.236	97.45	52.22	54.76	< 0.0001
4	<u>4.597</u>	<u>97.66</u>	<u>63.30</u>	53.65	< 0.0001
5	4.637	98.69	63.62	54.93	< 0.0001
6	4.654	99.72	64.11	56.28	< 0.0001
7	4.718	99.80	69.07	57.97	< 0.0001
8	<u>4.495</u>	<u>99.84</u>	<u>70.34</u>	59.31	< 0.0001
9	4.568	99.89	71.63	61.36	< 0.0001
10	<u>4.465</u>	<u>99.92</u>	<u>72.96</u>	64.50	< 0.0001
11	4.536	99.93	75.51	70.94	< 0.0001
12	4.551	99.95	76.89	<u>72.21</u>	<u>< 0.0001</u>
13	<u>4.491</u>	99.96	79.21	68.85	< 0.0001
14	4.510	99.97	82.02	67.28	< 0.0001
15	4.516	99.98	83.03	65.82	< 0.0001
16	4.495	99.98	84.93	69.43	< 0.0001
17	<u>4.330</u>	99.98	87.55	70.64	< 0.0001
18	4.473	99.98	88.42	72.69	< 0.0001
19	4.567	99.99	90.65	72.60	< 0.0001
20	4.529	99.99	92.35	<u>74.98</u>	<u>< 0.0001</u>
21	<u>4.486</u>	99.99	93.52	74.39	< 0.0001
22	4.512	99.99	95.02	72.04	< 0.0001
23	4.517	99.99	95.82	70.73	< 0.0001
24	4.582	99.99	96.39	69.60	< 0.0001
25	4.597	99.99	97.07	67.81	< 0.0001
26	4.690	99.99	97.51	66.99	< 0.0001
27	4.696	99.99	97.89	65.87	< 0.0001
28	4.746	99.99	98.36	65.09	< 0.0001
29	4.782	99.99	98.72	64.33	< 0.0001
30	4.830	99.99	98.95	63.35	< 0.0001

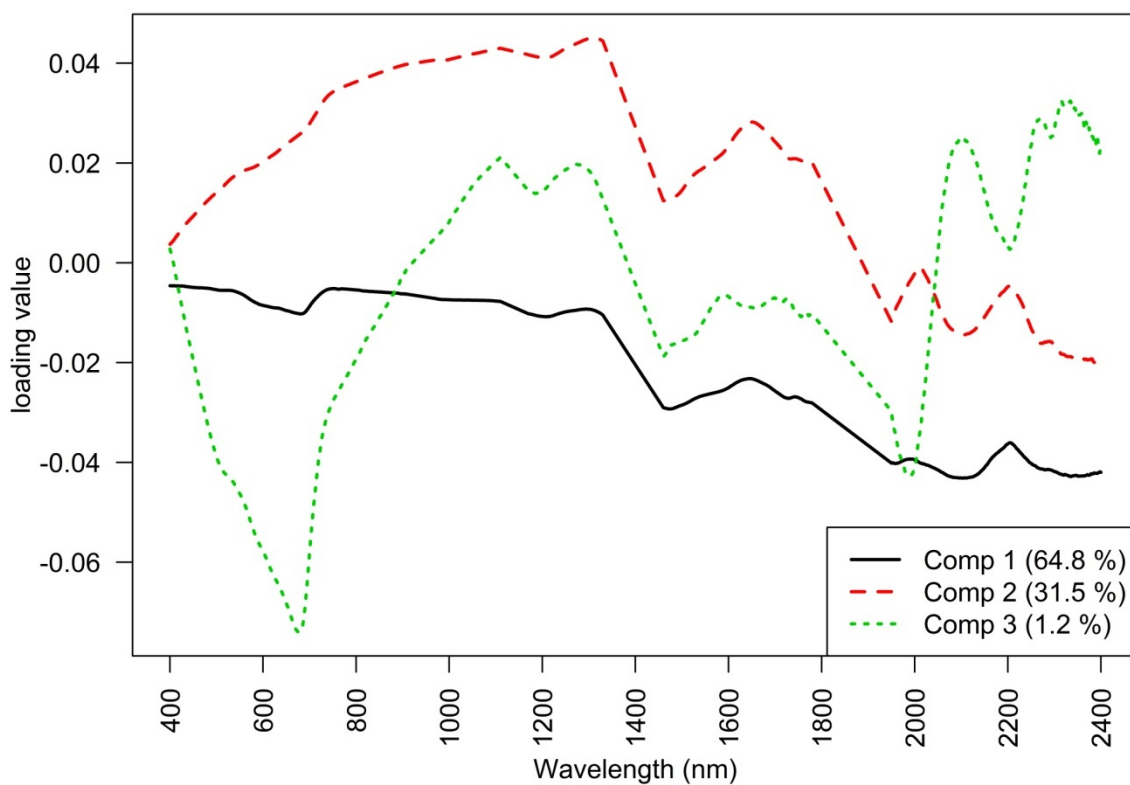
3076

out, a 12-component model ($r^2 = 0.72$, $P < 0.0001$) and a 20-component model ($r^2 = 0.75$, $P < 0.0001$; Table 3.12). Neither of these two models are associated with RMSEP minima and nor are they associated with components that reduced the RMSEP by greater than 2%. In fact from the addition of the 4th component to the model which accounted for a 12% reduction in RMSEP, only the 8th (4.7%), 10th (2.3%) and 17th (3.7%) components resulted in a RMSEP reduction of greater than 2% when added to the model. Thus in choosing the best model, the situation is quite challenging. From a parsimonious perspective, only adding a component if it makes a greater than 2% reduction in RMSEP (Cho *et al.*, 2007, Kooistra *et al.*, 2004), a 4-component model would be optimum. However, this only gives a coefficient of determination of 0.54 with validation data. From a prediction perspective, a 12-component model would give coefficients of determination of 0.72 which is markedly better than the 4-component model. But the 12-component model doesn't fit either the RMSEP condition of being a local minimum (Mevik and Wehrens, 2007) or the greater than 2% rule (Cho *et al.*, 2007, Kooistra *et al.*, 2004) for including components in the model, nevertheless, it was the model with the strongest fit for validation data and was therefore preferred.

The first loading for the infiltration model derived for West Wits was all negative values (Figure 3.25a) and showed a minor feature at 689 nm that may be associated with chlorophyll (Cho and Skidmore, 2006, Dawson and Curran, 1998, Rock *et al.*, 1988, Tucker, 1979) and a small peak and trough in the SWIR located at 2203 and 2220 nm that may be associated with ALOH clay (Chabrilat *et al.*, 2002). These features suggest that the first component, which accounted for 91% of the spectral variation and 42.5% of the LFA infiltration variation, is a mix of soil and plant features. However, Pearson correlations between the mean spectra for each patch type and the first five loadings (Figure 3.26a) show that the first loading was most negatively correlated with bare patches ($r = -0.978$) and bare grass ($r = -0.966$) and slightly less so with the vegetated patches: grass patches ($r = -0.878$), sparse grass ($r = -0.846$) and *S. plumosum* patches ($r = -0.809$). This shows that the first loading is strongly influenced by all five patch types but is dominated by the soil-related patch types. The second loading accounts for 4.4% of the spectral variability and 17.9% of the variability in the LFA infiltration index (Figure 3.25a). Pearson correlations show that this second loading is positively correlated with the three vegetated patch types: *S. plumosum* ($r = 0.456$), sparse grass ($r = 0.417$) and grass patches ($r = 0.385$), but has almost no correlation

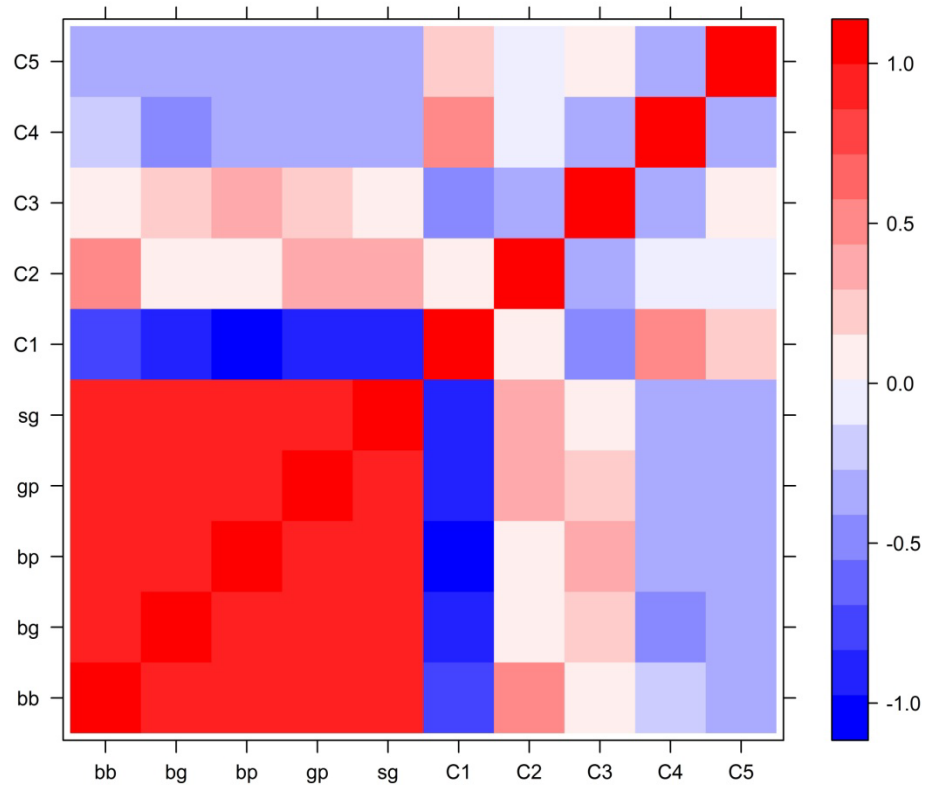


3108 a)

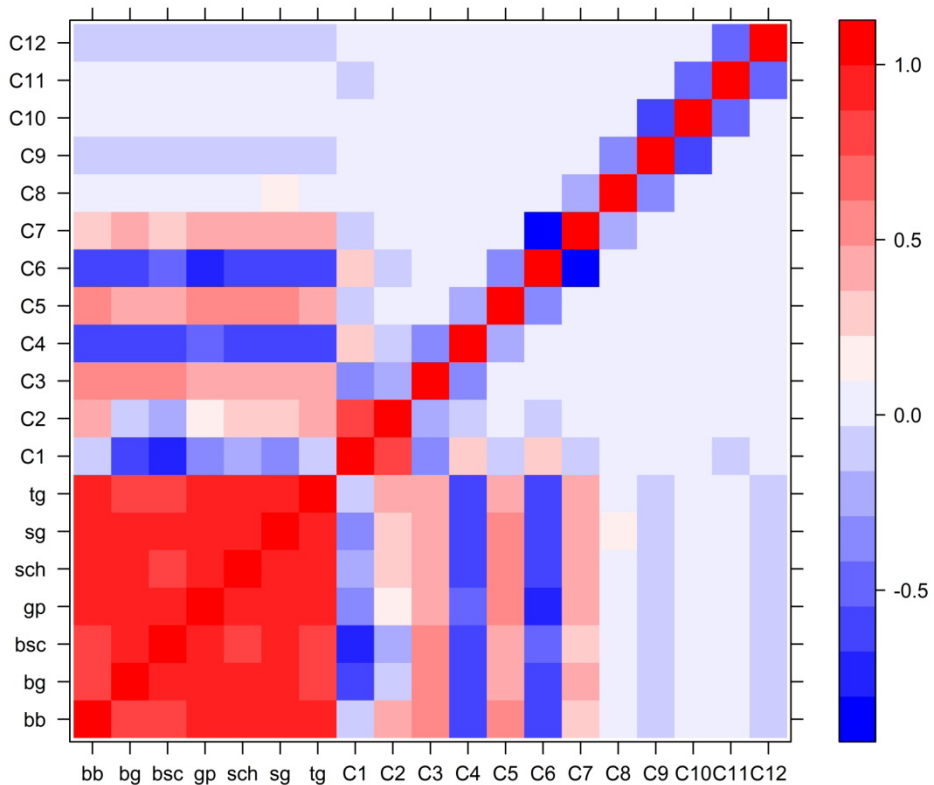


3109 b)

3110 **Figure 3.25** Loadings from the PLSR model predicting infiltration at (a) West Wits and (b)
3111 Vaal River



3112 a)



3113 b)

3114 **Figure 3.26** Pearson correlations between the mean spectra for each patch type and (a)
 3115 the first five loadings for the LFA infiltration model for West Wits, and (b) the
 3116 the first twelve loadings for the infiltration model derived from Vaal River.

3117 with the soil-related patch types: bare grass ($r = 0.135$) and bare patch ($r = 0.036$, Figure
3118 3.26a). This suggests that this second component is dominated by plant-related spectral
3119 features. The third loading is most correlated with bare patches ($r = 0.338$), bare grass patch
3120 ($r = 0.255$) and grass patches ($r = 0.218$) and has virtually no correlation with *S. plumosum* (r
3121 $= 0.098$) and sparse grass patches ($r = 0.048$; Figure 3.26a) implying a dominance of soil-
3122 related features and the presence of some plant related features. It is noticeable though,
3123 and odd, that sparse grass and grass patches, which have such similar mean spectra, have
3124 such different correlations to this third loading.

3125 The first loading for the PLSR derived model predicting infiltration at Vaal River accounts for
3126 64.8% of the spectral variability and 40.7% of the variability in the LFA infiltration calibration
3127 data (Figure 3.25b). The first loading has a strong absorption feature centered around 681
3128 nm (Figure 3.25b) that is possibly related to chlorophyll absorption (Cho and Skidmore,
3129 2006, Dawson and Curran, 1998, Tucker, 1979), a smaller feature at 1208 nm, and a small
3130 feature at 1729/30 nm that may relate to cellulose (Fourty *et al.*, 1996), and a strong peak
3131 centered on 2204/5 nm that may be related to clay (Chabrillat *et al.*, 2002). This first loading
3132 is negatively correlated with the mean spectra for all patch types (Figure 3.26b) and most
3133 negatively correlated with the biological soil crust patches ($r = -0.711$) followed by bare
3134 grass patches ($r = -0.648$). There is a weak negative correlation between the first loading and
3135 the two grass dominated patches: grass patches ($r = -0.407$) and sparse grass ($r = -0.315$),
3136 and a very weak negative correlation with *Schoenoplectus* patches ($r = -0.257$), tall grass
3137 patches ($r = -0.143$) and *S. plumosum* patches ($r = -0.141$; Figure 3.26b). These correlations
3138 suggest that the first component is strongly dominated by soil related characteristics and
3139 less so by vegetation related spectral characteristics. The second loading is mostly positive
3140 with the exception of the SWIR region above 1950 nm which is all negative and has features
3141 that compare strongly with plant reflectance in the SWIR (Figure 3.25b). This second loading
3142 is most positively correlated with tall grass patches ($r = 0.415$) and *S. plumosum* patch ($r =$
3143 0.398), with weaker correlations with *Schoenoplectus* patches ($r = 0.293$) and sparse grass
3144 patches ($r = 0.246$; Figure 3.26b). Both biological soil crust patches ($r = -0.224$) and bare
3145 grass patches ($r = -0.137$) are negatively correlated with the second loading (Figure 3.26b).
3146 So, the second loading is predominantly a plant dominated loading of the PLSR infiltration
3147 model for Vaal River. The third loading from the Vaal River infiltration model accounts for

1.2% of the spectral variation and 9.2% of the variability in the LFA infiltration values (Figure 3.25b) and is positively correlated with the mean spectra of all seven patch types from Vaal River (Figure 3.26b). However, it is most strongly correlated with the non- or low-vegetated patch types of biological soil crust ($r = 0.605$) and bare grass ($r = 0.553$), or the chlorophyll dominated spectra of *S. plumosum* ($r = 0.517$; Figure 3.26b). This suggests that the third loading is dominated by soil-related spectral characteristics and least influenced by the spectral characteristics of winter-senesced grass.

3.3.5 Predicting the LFA Nutrient Cycling Index

Like the infiltration models for West Wits, models predicting the LFA nutrient cycling index were robust. The first RMSEP local minimum (4.076%) was at six components (Table 3.13) and there were two more local minima at the 13th component (3.832%) and the 21st component (3.727%; Table 3.13) before the model reached its minimum RMSEP at 33 components (3.718%, result not shown). All these RMSEP values compare favorably with the standard deviation for the nutrient cycling index ($\bar{x} = 30.21\%$, $SD = 9.00\%$). The 6-component model gives strong predictions with both calibration ($r^2 = 0.85$, $P < 0.0001$) and validation data ($r^2 = 0.73$, $P < 0.0001$; Table 3.13 and Figure 3.27a). However, the 5-component model gives stronger validation predictions ($r^2 = 0.75$, $P < 0.0001$) than the 6-component model and the best validation predictions are for a 16-component model ($r^2 = 0.83$, $P < 0.0001$; Table 3.13). Neither of these two models is associated with a RMSEP minimum although the 5th component caused a reduction in the RMSEP of 10.3% and the 16th component a reduction of 5.4%. The 6th component was the end of a sequence of reductions in RMSEP value of greater than 3% per added component. Under the assumption of parsimony, the 5-component model is the preferred model (Table 3.13 and Figure 3.27a).

The strength of the Vaal River PLSR model predicting the LFA nutrient cycling index lies between the strength of the infiltration model for Vaal River, which was fairly strong, and the soil stability model for Vaal River, which was weak. The first local, and also the absolute minimum RMSEP (4.675%), is for the 9th component (Table 3.14) and is quite close to the standard deviation for the Vaal River nutrient cycling values ($\bar{x} = 24.63\%$, $SD = 6.46\%$). This 9-component model has a relatively low calibration and validation coefficient of determination ($r^2 = 0.58$, $r^2 = 0.44$ respectively, $P < 0.0001$ in both cases; Table 3.14). The

3178 **Table 3.13** PLSR models predicting LFA nutrient cycling index for West Wits with
3179 calibration data (n = 57) and validation data (n = 19).

Component	Calibration RMSEP (%)	Calibration spectral r ²	Nutrient cycling calibration r ²	Nutrient cycling validation r ²	Probability
1	7.470	91.07	35.19	46.78	0.0012
2	5.809	94.79	64.07	62.63	< 0.0001
3	5.225	98.68	70.80	68.55	< 0.0001
4	4.721	99.27	77.46	71.45	< 0.0001
5	4.236	99.36	84.40	75.04	< 0.0001
6	<u>4.076</u>	<u>99.87</u>	<u>84.99</u>	72.72	< 0.0001
7	4.131	99.89	88.04	65.55	< 0.0001
8	4.309	99.93	89.51	68.88	< 0.0001
9	4.584	99.95	92.52	76.39	< 0.0001
10	4.192	99.96	94.85	78.33	< 0.0001
11	4.092	99.96	95.71	78.28	< 0.0001
12	3.890	99.97	96.65	76.96	< 0.0001
13	<u>3.832</u>	<u>99.97</u>	<u>97.78</u>	80.09	< 0.0001
14	4.133	99.98	98.24	80.64	< 0.0001
15	4.140	99.98	98.84	80.61	< 0.0001
16	3.918	99.99	99.19	82.82	< 0.0001
17	3.911	99.99	99.47	82.63	< 0.0001
18	3.774	99.99	99.64	81.17	< 0.0001
19	3.750	99.99	99.80	80.99	< 0.0001
20	3.732	99.99	99.90	80.23	< 0.0001
21	<u>3.727</u>	99.99	99.95	79.07	< 0.0001
22	3.731	99.99	99.97	78.64	< 0.0001
23	3.738	99.99	99.99	78.32	< 0.0001
24	3.727	99.99	100.00	78.45	< 0.0001
25	3.726	99.99	100.00	78.53	< 0.0001
26	3.722	99.99	100.00	78.55	< 0.0001
27	3.721	99.99	100.00	78.50	< 0.0001
28	3.721	99.99	100.00	78.50	< 0.0001
29	<u>3.720</u>	100.00	100.00	78.50	< 0.0001
30	3.720	100.00	100.00	78.48	< 0.0001

3180

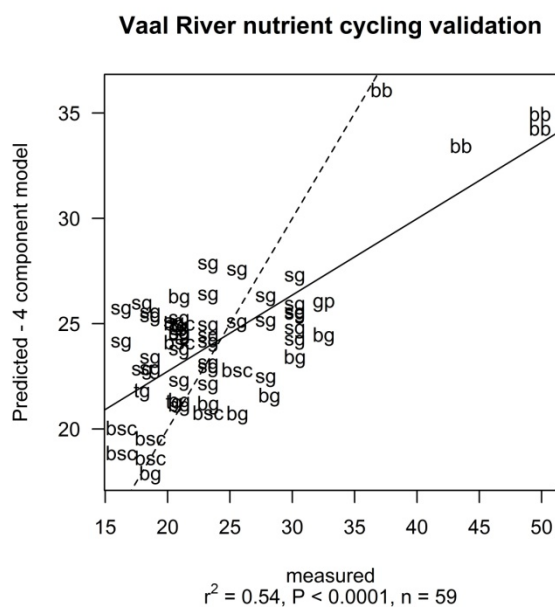
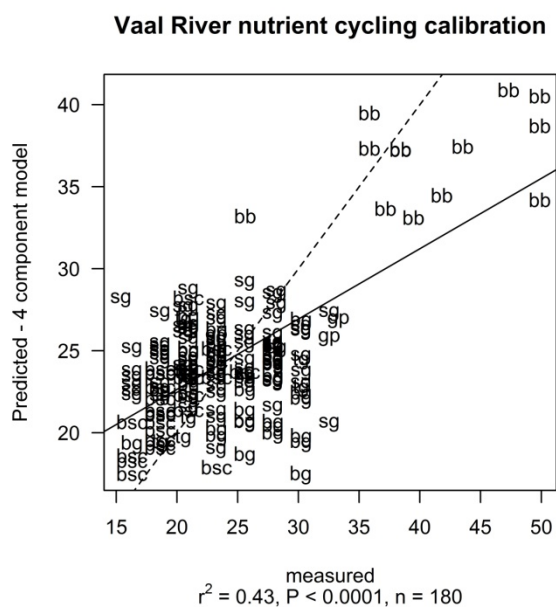
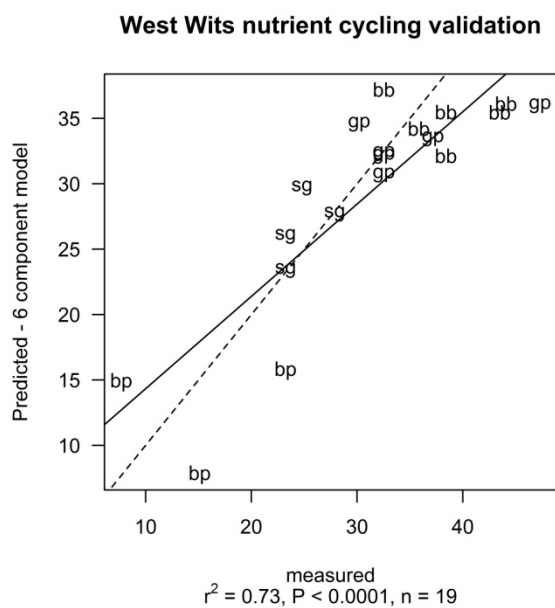
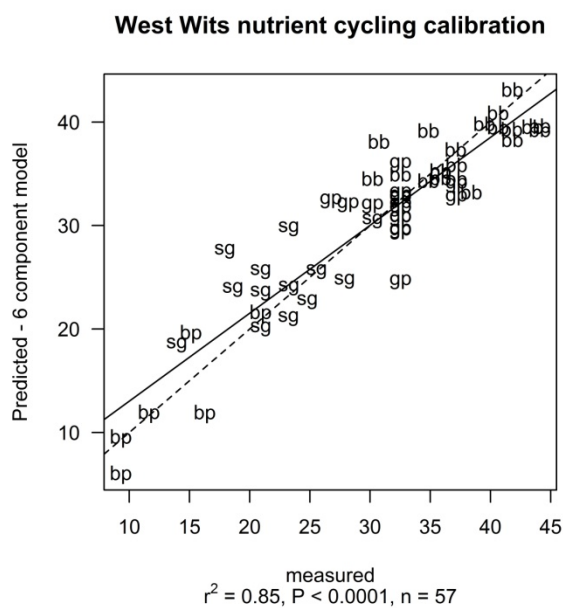


Figure 3.27 PLSR model predictions for LFA nutrient cycling index with calibration (left column) and validation (right column) data from West Wits (top row) and Vaal River (bottom row)

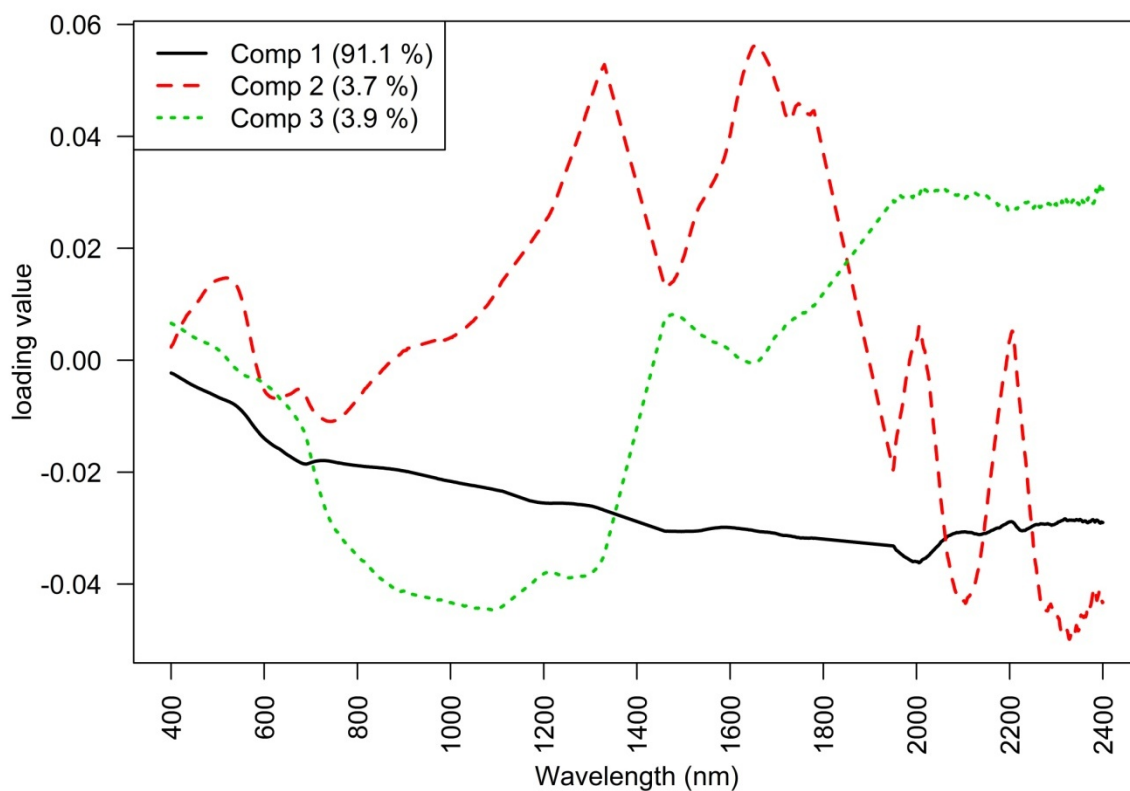
3185 **Table 3.14** PLSR models predicting the LFA nutrient cycling index for Vaal River from
3186 calibration data (n = 180) and validation data (n = 59).

Component	Calibration RMSEP (%)	Spectral calibration r^2	Nutrient cycling calibration r^2	Nutrient cycling validation r^2	Probability
1	5.745	65.62	23.44	37.76	< 0.0001
2	5.698	95.33	25.13	34.16	< 0.0001
3	5.322	97.52	37.51	51.43	< 0.0001
4	5.134	98.30	42.86	<u>53.97</u>	<u>< 0.0001</u>
5	4.940	99.04	47.07	50.03	< 0.0001
6	4.891	99.77	48.18	47.18	< 0.0001
7	4.794	99.82	53.79	44.29	< 0.0001
8	4.681	99.85	56.99	44.24	< 0.0001
9	<u>4.675</u>	<u>99.91</u>	<u>57.97</u>	43.57	< 0.0001
10	4.731	99.93	60.23	40.17	< 0.0001
11	4.786	99.94	63.97	44.10	< 0.0001
12	4.831	99.95	65.40	48.35	< 0.0001
13	<u>4.790</u>	<u>99.96</u>	<u>68.39</u>	<u>51.01</u>	<u>< 0.0001</u>
14	4.851	99.96	73.80	45.55	< 0.0001
15	4.974	99.97	74.90	47.43	< 0.0001
16	4.975	99.98	77.15	46.23	< 0.0001
17	5.110	99.98	81.03	51.98	< 0.0001
18	<u>5.081</u>	99.98	83.17	48.69	< 0.0001
19	5.127	99.99	86.42	51.86	< 0.0001
20	<u>4.974</u>	99.99	89.56	<u>52.42</u>	<u>< 0.0001</u>
21	5.055	99.99	91.33	51.29	< 0.0001
22	<u>4.946</u>	99.99	92.47	49.73	< 0.0001
23	5.030	99.99	93.51	47.31	< 0.0001
24	5.064	99.99	94.50	46.45	< 0.0001
25	5.187	99.99	95.20	46.45	< 0.0001
26	5.215	99.99	96.10	44.11	< 0.0001
27	5.289	99.99	96.79	42.32	< 0.0001
28	5.382	99.99	97.51	40.58	< 0.0001
29	5.499	99.99	98.20	41.41	< 0.0001
30	5.502	99.99	98.54	40.26	< 0.0001

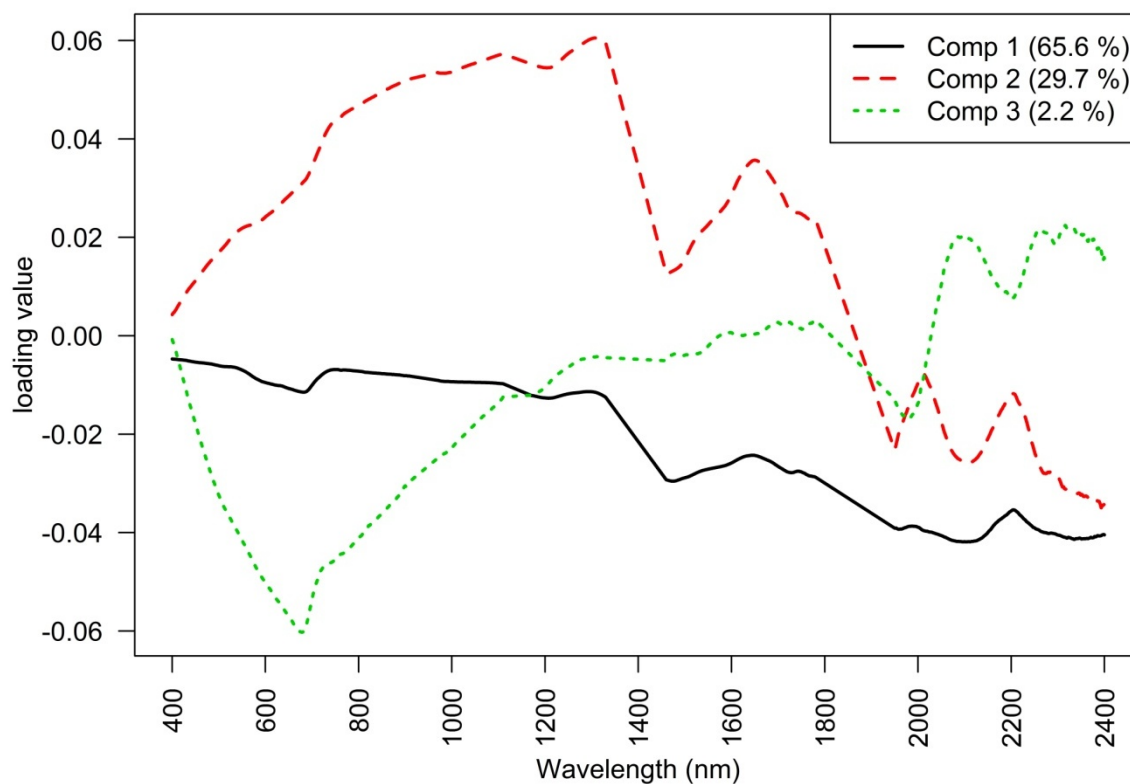
3187

highest validation coefficient of determination ($r^2 = 0.54$, $P < 0.0001$) was for a 4-component model but the calibration coefficient of determination was low ($r^2 = 0.43$, $P < 0.0001$; Table 3.14 and Figure 3.27b). For the first five components, all but the addition of the second component resulted in a reduction in the RMSEP of greater than 2%. However, regardless of the number of components in the model, the performance of the 5-component model is weak and for this reason and through parsimony and stronger coefficient of determination with validation data, the 4-component model is preferred (Table 3.14 and Figure 3.27b).

The first loading for the West Wits nutrient cycling model explained 91.1% of the variation in the spectral data but only 35% of the variation in the LFA nutrient cycling index (Table 3.13 and Figure 3.28a), and like previous LFA indices, had entirely negative values with a general increase on model influence from visible wavelengths to SWIR wavelengths. Generally this first PLSR loading had few standout identifiable features with the most apparent being around 688 - 690 nm (Figure 3.28a) which presumably is associated with the chlorophyll absorption center and associated red-edge (Cho and Skidmore, 2006, Dawson and Curran, 1998). This first loading was also strongly negatively correlated with all five vegetation patch types with the strongest negative correlation with bare patches (-0.978), closely followed by bare grass patches (-0.966; Figure 3.29a). The vegetation patches were only slightly less negatively correlated than the mainly exposed soil patches with the order being grass patches (-0.877), then sparse grass patches (-0.846), and *S. plumosum* patches (-0.809) the least correlated with the first loading (Figure 3.29a). The second loading, which shows much more variability than the first loading and accounts for 3.7% of the spectral variability and 29% of the variability in the LFA nutrient cycling index (Figure 3.28a), had peaks at 519 and 678 nm which may be plant pigment related (Cho and Skidmore, 2006, Gitelson *et al.*, 2002, Dawson and Curran, 1998, Tucker, 1979), and absorption features at 1729, 1772 and 2099 that may be related to lignin and/or cellulose (Fourty *et al.*, 1996). Correlations between the 2nd loading and the mean spectra from the five patch types showed a fairly strong positive correlation with *S. plumosum* patches (0.501) and slightly less so with grass patches (0.453) and sparse grass patches (0.452) but weakly correlated with bare grass (0.195) and bare patches (0.105) which all suggests that the 2nd loading is dominated by plant features rather than soil features (Figure 3.29a). The third loading accounts for 3.9% of the spectral variability and 6.7% of the variability in the LFA

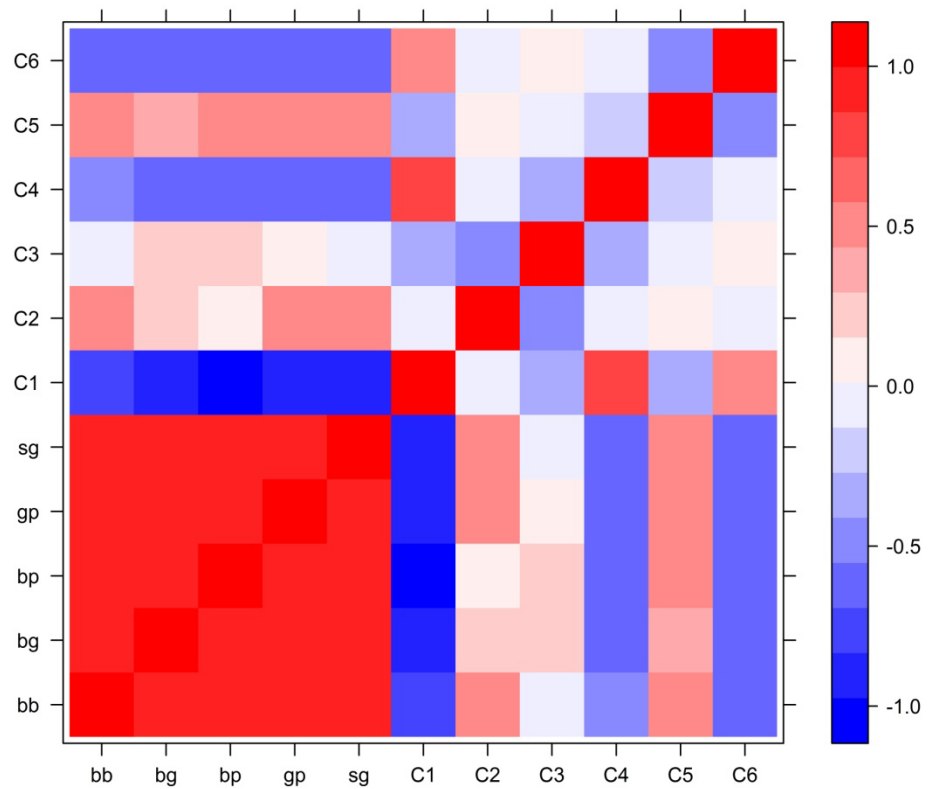


3219 a)

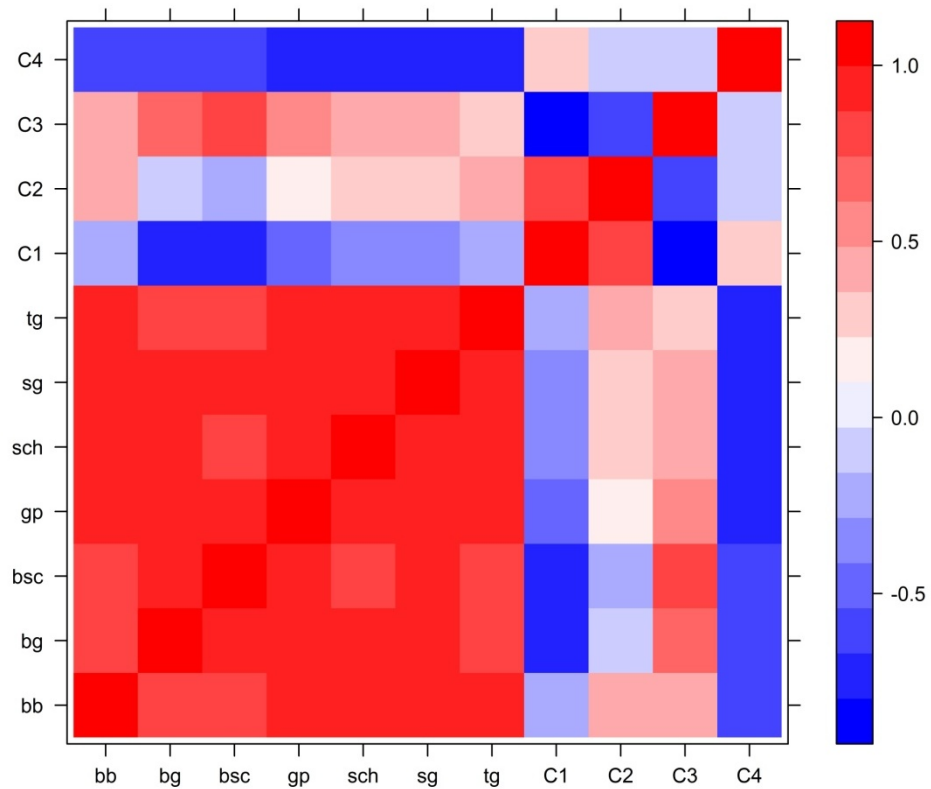


3220 b)

3221 **Figure 3.28** Nutrient cycling loadings for (a) West Wits and (b) Vaal River



3222 a)



3223 b)

3224 **Figure 3.29** Nutrient cycling at a) West Wits and (b) Vaal River

3225 nutrient cycling index and is dominated by a very broad negative feature from about 750 nm
3226 to 1100 nm and a very strong positive SWIR plateau between 2000 and 2400 nm that is
3227 relatively featureless (Figure 3.28a). This third loading for the West Wits PLSR model
3228 predicting LFA nutrient cycling has very weak correlations with all patch types being most
3229 correlated with bare patches (0.276) and bare grass patches (0.183) but has almost no
3230 correlation with the vegetated patch types (Figure 3.29a).

3231 The first three loadings for the PLSR model predicting LFA nutrient cycling with ground-
3232 based spectral data has quite different patterns to the loadings for the West Wits PLSR
3233 model (Figure 3.28b). The first loading accounts for 65.9% of the variation in the spectral
3234 data and 23.4% of the variation in the LFA nutrient cycling index and has a clear absorption
3235 feature centered at 680 – 684 nm which is probably associated with chlorophyll (Cho and
3236 Skidmore, 2006, Dawson and Curran, 1998, Tucker, 1979), a feature of unknown origin at
3237 1207 – 1209 nm, a small feature at 1727 – 1729 nm and a broad feature centered on 2103
3238 nm, both of which may be related to cellulose (Daughtry *et al.*, 2004, Fourty *et al.*, 1996).
3239 This first loading is negatively correlated with the mean spectra for all seven patch types at
3240 Vaal River (Figure 3.29b) and most strongly so with biological soil crust (-0.737) and bare
3241 grass patches (-0.677), both of which have predominantly soil surface exposure. Grass (-
3242 0.441) and sparse grass (-0.351) patches have very similar and moderate mean spectral
3243 correlation to the first loading, while *Schoenoplectus* (-0.294), tall grass (-0.181) and *S.*
3244 *plumosum* (-0.177) patches have very low mean spectral correlation with the first loading
3245 (Figure 3.29b). So, despite many of the features identified in the profile of the first loading
3246 being related to plant characters, it seems that soil characters have a great influence on the
3247 first loading.

3248 The second loading accounted for 29.7% of the spectral variability in the PLSR model and
3249 only 1.7% of the variability in the predictions for the nutrient cycling index (Table 3.14). This
3250 second loading has a broad feature between 580 and 680 nm (Figure 3.28b) but it unclear if
3251 this is a plant pigment related feature although the 680 nm inflection may be related to
3252 plant chlorophyll and the red edge (Cho and Skidmore, 2006, Dawson and Curran, 1998,
3253 Tucker, 1979). There is a small feature at 1734 nm and a major feature centered on 2107 nm
3254 in the second loading (Figure 3.28b) that may both be related to plant cellulose (Daughtry *et al.*
3255 *al.*, 2004, Fourty *et al.*, 1996). Correlations between the mean spectra for each patch and

the second loading showed two distinct groups with biological soil crust (-0.237) and bare grass (-0.151) patches being negatively correlated with the second loading while tall grass (0.401), *S. plumosum* (0.389), *Schoenoplectus* (0.281), sparse grass (0.232) and less so grass patches (0.126) were all positively correlated with the second loading. This suggests that this second loading is a mix of vegetation and soil characters but these are having opposite effects in the model. The third loading only accounted for 2.2% of the variability in the spectra but 12.4% of the variability in the LFA nutrient cycling index (Table 3.14) and has a deep feature centered on 678/9 nm which presumably would be a chlorophyll feature, and two noisy absorption features at 2197 and 2292 nm (Figure 3.28b). However, in contradiction to the strong 678/679 nm feature, the third loading is most correlated with the mean spectra from the biological soil crust (0.760) and bare grass (0.705) patch types, and least correlated with *S. plumosum* (0.408) and tall grass (0.310) patches, with grass (0.527), *Schoenoplectus* (0.451) and sparse grass (0.464) patches lying between (Figure 3.29b). This suggests that the third loading is more correlated with soil surface features and less correlated with plant pigments.

3.4 Predicting Ecological Variables with PLSR and Spectral Data

Partial Least Squares regression was used to explore how the spectral measurements collected at ground level at both West Wits and Vaal River related to and predicted the various soil variables and vegetation biomass measured during this study. The results presented in Table 3.15 and 3.16 are the best models at predicting the respective value with validation data. Only models of fresh biomass or oven-dried dry above-ground standing biomass produced fair models at both sites. The results were very similar for these two treatments of biomass where fresh consists of winter senesced vegetation and evergreen *S. plumosum* above-ground biomass as collected in the field, and oven-dried means the same samples were oven-dried at 50°C until no more loss of weight. At West Wits the RMSEP for fresh above-ground biomass was 0.483 kg/m² for a 5-component model and the calibration coefficient of determination ($r^2 = 0.70$, $P < 0.0001$) was slightly stronger than that for validation ($r^2 = 0.55$, $P = 0.0003$; Table 3.15 and Figure 3.30a). At Vaal River, a 9-component model with a RMSEP of 0.651 kg/m² produced slightly stronger predictions than those at West Wits. However, the calibration coefficient of determination ($r^2 = 0.67$, $P < 0.0001$) was

3286 **Table 3.15** PLSR Model calibration and validation statistics for various ecological variables for above-ground biomass and soil sampled from
3287 West Wits.

Variable	Mean	SD	Model components	Calibration			Validation		
				n	RMSEP	Spectral r^2	Response r^2	n	Response r^2 Probability
Fresh above-ground biomass (kg/m ²)	1.19	0.74	5	57	0.483	99.43	69.50	19	55.29 0.0003
Dry above-ground biomass (kg/m ²)	1.01	0.60	5	57	0.398	99.44	68.24	19	50.15 0.0007
Soil organic matter (%)	4.92	1.66	10	57	1.903	99.95	68.56	17	31.57 0.019
Soil moisture (%)	9.97	2.75	7	57	1.960	99.89	69.99	19	44.84 0.0017
pH	5.28	0.34	5	57	0.343	99.38	29.72	19	56.73 0.0002
Soil solution EC (μS/cm)	41.84	19.61	9	57	22.57	99.94	55.24	19	10.12 0.18
soil-extractable NH ₄ ⁺ (μg/g)	1.38	0.85	3	57	0.938	97.08	5.72	19	22.87 0.038
Soil-extractable NO ₃ ⁻ (μg/g)	1.72	0.78	7	57	0.857	99.89	36.12	19	60.84 0.00008
Soil-extractable organic nitrogen (μg/g)	29.88	8.36	6	57	8.250	99.70	49.91	19	25.32 0.028
Nutrient cycling ratio	11.30	7.00	9	57	9.035	99.95	39.12	19	21.16 0.048

3288

3289 **Table 3.16** PLSR Model calibration and validation statistics for various ecological variables for above-ground biomass or soil sampled from
3290 Vaal River.

Variable			Model components	Calibration			Validation			
	Mean	SD		n	RMSEP	Spectral r ²	response r ²	n	response r ²	P
Fresh above-ground biomass (kg/m ²)	0.51	0.82	13	180	0.651	99.96	67.10	59	78.78	< 0.0001
Dry above-ground biomass (kg/m ²)	0.42	0.66	13	180	0.516	99.96	67.00	59	78.54	< 0.0001
Fresh root biomass (kg/m ²)	1.19	1.42	9	180	1.239	99.89	39.92	58	19.77	0.0004
Soil organic matter (%)	4.12	1.12	11	177	0.865	99.94	55.07	59	49.66	< 0.0001
Soil moisture (%)	3.10	1.46	8	180	1.013	99.85	59.00	59	67.81	< 0.0001
Soil solution pH	6.48	0.41	18	180	0.470	99.98	63.38	59	21.63	0.0002
Soil solution EC (μS/cm)	126.5	310.5	9	180	260.1	99.90	52.56	59	63.03	< 0.0001
Dry gravel (g/cm ³)	0.13	0.15	6	180	0.131	99.73	32.37	59	51.11	< 0.0001
Core depth	8.91	2.14	7	180	1.822	99.72	38.22	59	45.28	< 0.0001
Soil bulk density (g/cm ³)	1.55	0.29	6	180	0.256	99.77	29.67	59	37.32	< 0.0001

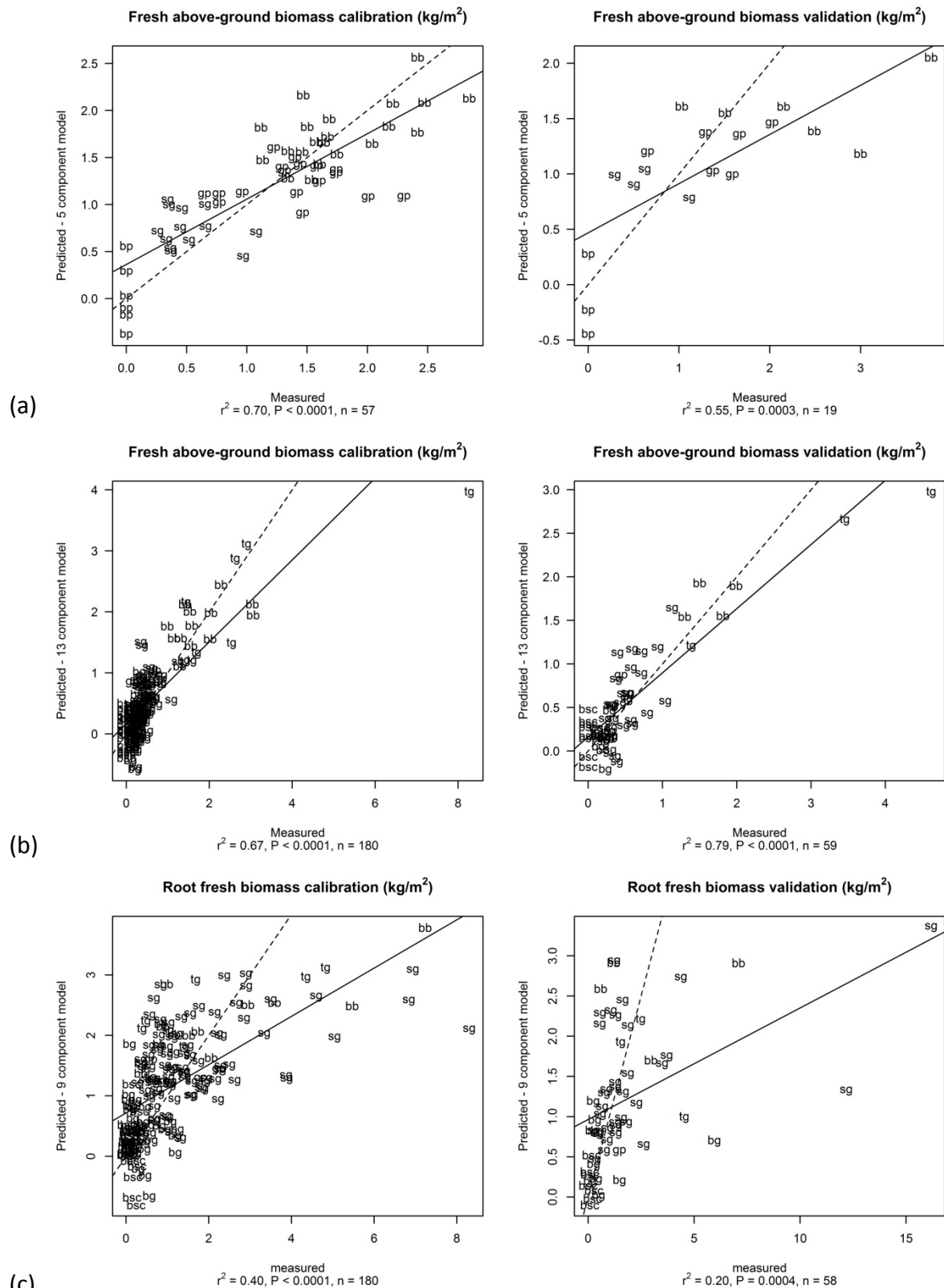


Figure 3.30 Calibration (left column) and validation (right column) predictions of (a) fresh above-ground biomass for West Wits and (b) Vaal River, and (c) fresh root biomass for Vaal River

3336 lower than that for validation ($r^2 = 0.79$, $P < 0.0001$) at Vaal River (Table 3.16 and Figure
3337 3.30b)

3338 Root biomass collected as the root material remaining in a 2 mm sieve was only quantified
3339 for Vaal River and was tested as kg/m^2 , kg/m^3 of soil, and g/g soil, both fresh and oven-
3340 dried, and all of these variables produced weak models. Therefore, only the best model is
3341 presented, which is fresh root on a kg/m^2 basis. The RMSEP for 9-component model was
3342 1.24 kg/m^2 which was almost as large as the standard deviation ($\bar{x} = 1.19 \pm 1.42 \text{ kg/m}^2$;
3343 Table 3.16). For calibration, the coefficient of determination was low ($r^2 = 0.40$, $P < 0.0001$)
3344 and validation was lower ($r^2 = 0.20$, $P = 0.0004$; Table 3.16 and Figure 3.30c). Various
3345 transformations of the response variable did not make any difference to the model
3346 outcomes although in the validation prediction, one of the gypsum-affected samples that
3347 was previously identified as an outlier was removed from the root predictions. It should be
3348 noted though that predictions from the PLSR model showed considerable compression of
3349 the range of values compared to the measured values (Figure 3.30c).

3350 The prediction of soil organic matter (SOM) from spectral data was mixed. West Wits had
3351 the best predictions with a 10-component model (Table 3.15) and Vaal River with an 11-
3352 component model (Table 3.16). West Wits had a higher RMSEP (1.90%) compared to Vaal
3353 River (0.87%) which followed the mean and standard deviation for SOM measurements
3354 which were also higher at West Wits ($\bar{x} = 4.92 \pm 1.66\%$; Table 3.15) compared to Vaal River
3355 ($\bar{x} = 4.12 \pm 1.12\%$; Table 3.16). Calibration predictions with West Wits data were
3356 considerably stronger ($r^2 = 0.69$, $P < 0.0001$) than validation predictions ($r^2 = 0.32$, $P = 0.019$)
3357 which were only just significant (Table 3.15 and Figure 3.31a). Vaal River had weaker
3358 predictions than did West Wits for calibration ($r^2 = 0.55$, $P < 0.0001$) but validation
3359 predictions ($r^2 = 0.50$, $P < 0.0001$; Table 3.16 and Figure 3.31b) were significantly stronger
3360 than validation predictions for SOM at West Wits.

3361 Soil moisture produced fairly robust models with spectral data with Vaal River producing
3362 stronger models than West Wits. A 7-component model at West Wits had a RMSEP of 1.60%
3363 which was greater than the standard deviation for soil moisture measurements ($\bar{x} = 3.11 \pm$
3364 1.39% ; Table 3.15). At Vaal River, an 8-component model had a RMSEP of 1.01% which was
3365 lower than the standard deviation for soil moisture ($\bar{x} = 3.10 \pm 1.46\%$; Table 3.16)

measurements at Vaal River. Calibration predictions ($r^2 = 0.70$, $P < 0.0001$) at West Wits were stronger than validation predictions ($r^2 = 0.45$, $P < 0.0017$; Table 3.15 and Figure 3.32a). At Vaal River, calibration predictions ($r^2 = 0.59$, $P < 0.0001$) were slightly weaker than validation predictions ($r^2 = 0.68$, $P < 0.0001$; Table 3.16 and Figure 3.32b).

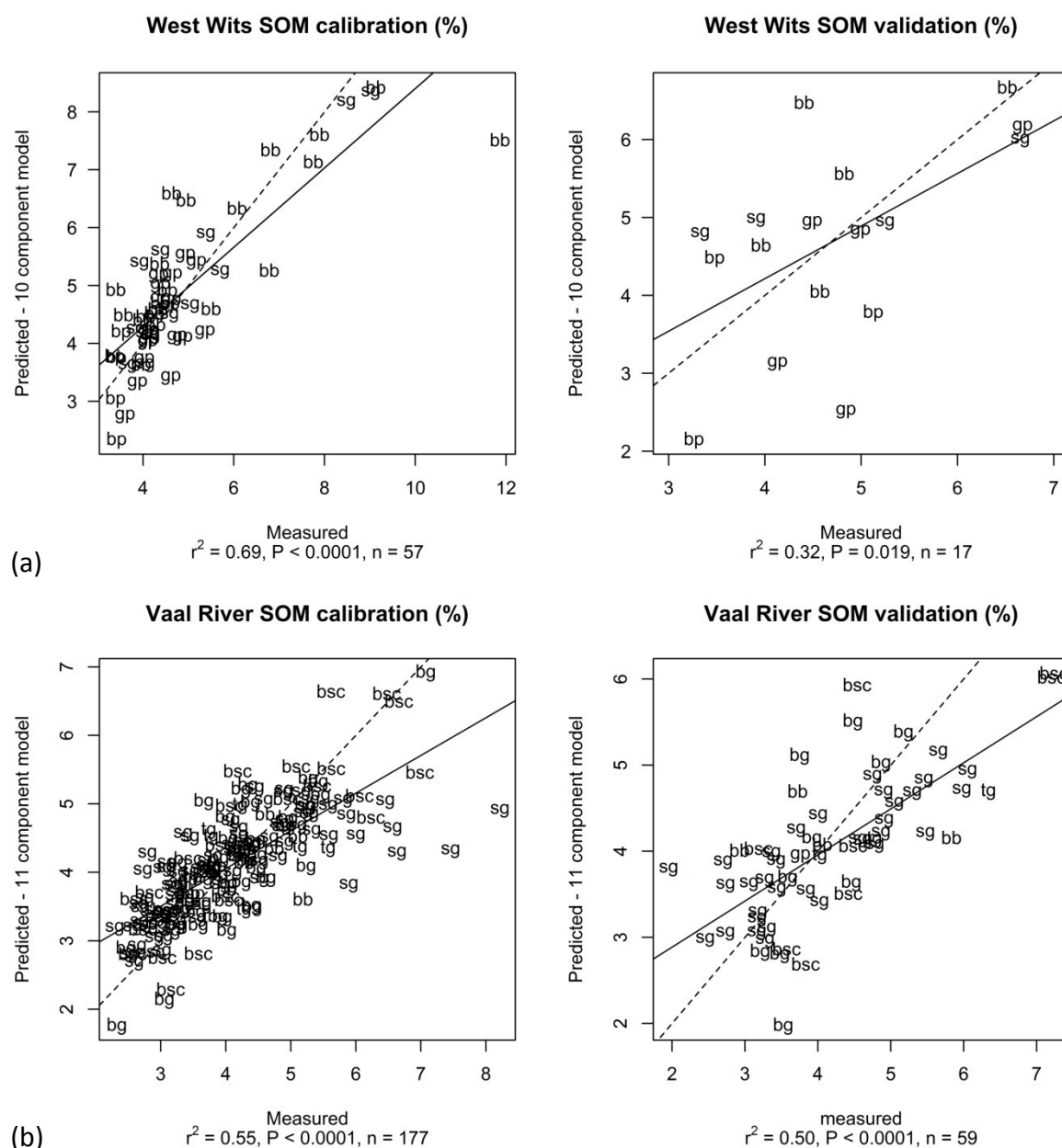


Figure 3.31 Calibration (left column) and validation (right column) predictions for SOM from (a) West Wits and (b) Vaal River

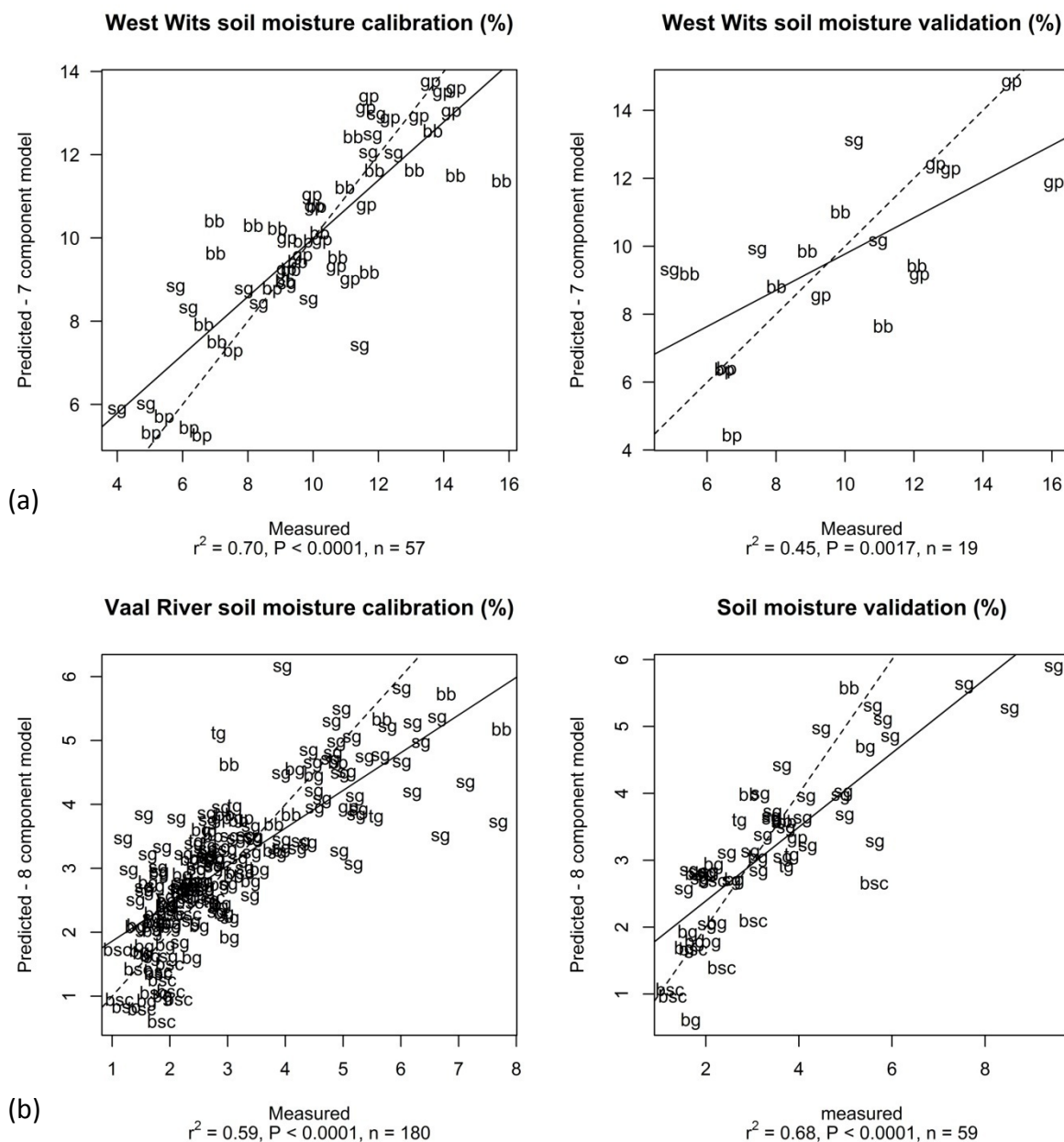


Figure 3.32 Calibration (left column) and validation (right column) predictions for soil moisture from (a) West Wits and (b) Vaal River

Soil solution pH for both West Wits and Vaal River produced contradictory results. At West Wits the RMSEP for the 5-component model was 0.34 which was identical to the standard deviation ($\bar{x} = 5.28 \pm 0.34$; Table 3.15) and the calibration predictions were weak ($r^2 = 0.30, P < 0.0001$) whereas the validation predictions were substantially stronger ($r^2 = 0.57, P = 0.0002$; Table 3.15). Vaal River showed the reverse pattern with strong calibration predictions ($r^2 = 0.63, P < 0.0001$) and weak validation predictions ($r^2 = 0.22, P = 0.0002$) for an 18-component model (Table 3.16). The RMSEP at Vaal River for the 18-component model was 0.47 which was greater than the standard deviation ($\bar{x} = 6.48 \pm 0.41$; Table 3.16) for soil

solution pH. Models testing the prediction of soil solution electrical conductivity from spectral data were also weak at West Wits but in the case of Vaal River they were fairly strong through the influence of a few very high values. At West Wits, a 9-component model had a RMSEP of 22.57 $\mu\text{S}/\text{cm}$ which was higher than the standard deviation ($\bar{x} = 41.84 \pm 19.61 \mu\text{S}/\text{cm}$) and there were no extreme values. The calibration coefficient of determination was moderately strong ($r^2 = 0.55$, $P < 0.0001$) but the validation coefficient of determination was very weak and not significant ($r^2 = 0.10$, $P = 0.18$; Table 3.15). At Vaal River the RMSEP (260.1 $\mu\text{S}/\text{cm}$) was lower than the standard deviation ($\bar{x} = 126.5 \pm 310.5 \mu\text{S}/\text{cm}$; Table 3.16). The calibration coefficient of determination was 0.53 ($P < 0.0001$) and the validation coefficient of determination was slightly stronger ($r^2 = 0.63$, $P < 0.0001$; Table 3.16 and Figure 3.33a).

Inorganic and organic-extractable soil nitrogen was only measured with samples from West Wits. Models predicting soil-extractable ammonium were very weak (Table 3.15) and will not be discussed further. Soil-extractable nitrate predictions with calibration data were fairly weak ($r^2 = 0.36$, $P < 0.0001$) and the RMSEP (0.857 $\mu\text{g}/\text{g}$ soil) for a 7-component model was larger than the standard deviation of the measurements ($\bar{x} = 1.72 \pm 0.78 \mu\text{g}/\text{g}$ soil; Table 3.15). However, validation predictions with the same 7-component model were considerably stronger ($r^2 = 0.61$, $P = 0.00008$; Table 3.15 and Figure 3.33b). Extractable organic nitrogen produced moderate calibration predictions ($r^2 = 0.50$, $P < 0.0001$) but the validation predictions were weak ($r^2 = 0.25$, $P = 0.028$; Table 3.15). The nutrient cycling ratio, which is calculated by dividing extractable organic nitrogen by inorganic nitrogen, also performed weakly with calibration predictions ($r^2 = 0.39$, $P < 0.0001$) being slightly better than validation predictions ($r^2 = 0.21$, $P = 0.48$; Table 3.15).

Gravel, defined as the portion of stone remaining in a 2 mm sieve, was weighed fresh and dry and converted to values of g/cm^3 soil, kg/m^2 and mg/g soil. None of these measurements of gravel content of soils produced good models with strong predictions. As representative of the group, the model predicting dry gravel on a gram per cubic centimetre of soil is presented (Table 3.16). The RMSEP is 0.13 g/cm^3 which is less than the standard deviation ($\bar{x} = 0.13 \pm 0.15 \text{ g}/\text{cm}^3$ soil; Table 3.16). The calibration predictions were weaker ($r^2 = 0.32$, $P < 0.0001$) than the validation predictions ($r^2 = 0.51$, $P < 0.0001$; Table 3.16 and Figure 3.33c), and although log, square root and arcsin transformation of the response

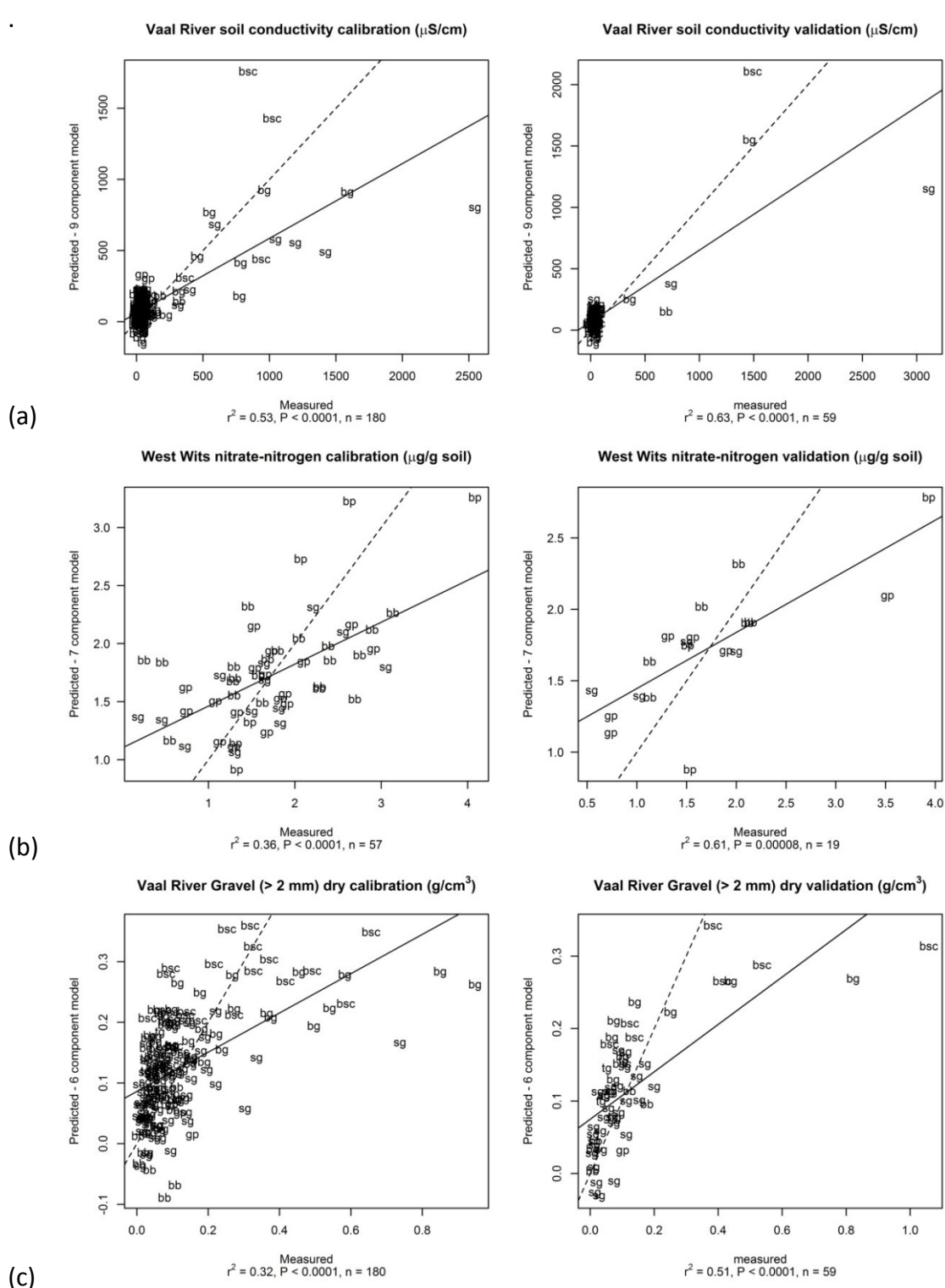


Figure 3.33 Calibration (left column) and validation (right column) predictions for (a) soil electrical conductivity at Vaal River, (b) organic soil nitrogen at West Wits, and (c) dry gravel content of soil at Vaal River.

3524 variable improved the calibration predictions, they did not improve the validation
3525 predictions (not shown). Core depth also produced weak models (Table 3.16) and possibly
3526 might have been more realistic if the measurement had been depth of penetration rather
3527 than limited to 10 cm as a large proportion of the variability in predictions was related to
3528 the maximum core depth of 10 cm. Soil bulk density also produced weak models with
3529 spectral data and will not be discussed further (Table 3.16).

3530 3.4 **Discussion**

3531 This discussion will explore the results for the patch models and the various ecological
3532 variables measured at West Wits. It discusses selecting the best model under the section on
3533 the PLSR models predicting the soil stability, infiltration and nutrient cycling indices (section
3534 3.4.2), but it will not include a discussion of the results for the three LFA indices of soil
3535 stability, infiltration and nutrient cycling, as well as those for above-ground biomass. These
3536 were also modelled and predicted using Hyperion hyperspectral data which is the subject of
3537 Chapter 4 and therefore the results from the ground-based models and the Hyperion results
3538 are discussed together in Chapter 4, sections 4.4.2 and 4.4.4 respectively.

3539 3.4.1 **Patch Models**

3540 The results explored in depth the PLSR models that predicted patch types for both sites
3541 because these models illustrate some of the weaknesses and strengths of using ground-
3542 based spectral data and PLSR modeling to predict ecological function and processes, and
3543 because most other properties of a site, such as biomass or nutrient cycling, are often
3544 dependent on or influenced by the type of patch that occupies that site (Daryanto *et al.*,
3545 2013, Li *et al.*, 2013, Jobbágy and Jackson, 2001, Jobbágy and Jackson, 2000, Aguiar and
3546 Sala, 1999, Burke *et al.*, 1998, Schlesinger *et al.*, 1996, Charley and West, 1975). The initial
3547 models predicting four patch types at West Wits and seven patch types at Vaal River were
3548 weak (Table 3.2 and 3.6). The *S. plumosum* patch type with its chlorophyll feature (Cho and
3549 Skidmore, 2006, Dawson and Curran, 1998, Tucker, 1979) dominating its spectrum (Figure
3550 3.3 and 3.14) stood out in both models as being relatively strongly modelled, and slightly
3551 less strongly were the bare patches at West Wits and the biological soil crust patches at Vaal
3552 River. The various winter-senesced grass patches were very weakly modelled at both sites,
3553 as were *Schoenoplectus* patches at Vaal River (Table 3.2 and 3.6). As was shown in Figures
3554 3.4 and Figure 3.14, the spectra for these various grass patches with weak predictions
3555 tended to be very similar. A simplified model predicting patch types with clearly
3556 differentiated spectral profiles such as biological soil crust, sparse grass and *S. plumosum*
3557 produce a strong model, and showed that where the different patch types had clearly
3558 different spectral forms, PLSR modelling could clearly separate the patch types (Figure 3.17).
3559 However, PLSR struggled to separate patch types with very similar spectra such as sparse

3560 grass and grass patches, implying that the spectra for these patch types were spectrally
3561 inseparable using PLSR.

3562 Furthermore, patch types having a mix of spectral profiles, such as bare grass, which
3563 consists of exposed soil surface with a range of grass cover from almost nothing to almost
3564 100%, can significantly weaken the model (Tables 3.2 and 3.6). It is clear from an ecological
3565 perspective that bare grass patches represent something different to sparse grass patches
3566 through features such as species composition, plant size and possibly annual versus
3567 perennial habit (Macdonald, 1978), and are also different to bare patch/biological soil crust
3568 patches through the presence of some form of vegetation cover. Spectroscopy, on the other
3569 hand, “sees” these three patch types as a gradation between two extremes: exposed soil
3570 surface and dry senesced vegetation, where bare grass patches fall into either
3571 bare/biological soil crust patches or sparse grass patches depending on the specific patch
3572 sampled and the degree of vegetation cover. This shows clearly in the different prediction
3573 plots with both calibration and validation data (Figures 3.2, 3.10, 3.11, 3.18 and 3.20). These
3574 bare grass patches are important ecologically as they delineate a degraded area from
3575 functional grasslands (Macdonald, 1978). Examination of the prediction values for samples
3576 that occurred in the overlap of predictions showed that they were a mix of healthy sparse
3577 grass patches with low grass abundance, and bare grass patches, without any clear
3578 differentiation in ranking between the two patch types. So simply setting a range of values
3579 when predicting sparse grass patches or bare/biological soil crust patches for instance,
3580 wherein any patch within that value range could be considered a bare grass patch, and
3581 those above the range as the predicted patch type, and those below the range as the non-
3582 predicted patch types, would not adequately indicate true bare grass patches. These bare
3583 grass patches are essentially mixed pixels and a possible approach to the problem could be
3584 spectral mixture analysis (SMA) or any spectral unmixing or spectral feature fitting method.

3585 The SMA technique attempts to resolve in various ways the spectrum from a target
3586 consisting of a mixture of components, in this case a patch consisting of senesced grass and
3587 bare exposed soil surface, into pure spectra of each component and these pure spectra are
3588 called endmembers (Wang *et al.*, 2014, Rudorff *et al.*, 2009, Numata *et al.*, 2007, Numata *et al.*,
3589 2003, Roberts *et al.*, 1998, Smith *et al.*, 1990a, Smith *et al.*, 1990b, Huete, 1986). The
3590 most basic way to approach SMA is to regard the spectrum from a mixed pixel as equal to a

weighted sum of the pure spectra from each component making up the pixel, and the weights for each component are equal to their respective area in the pixel (Roberts *et al.*, 1998). Implicit is the assumption that the mixing of pure spectra is linear. Endmembers may be image endmembers whereby they are derived directly from pixels from the image itself, or they may be derived from field spectra, or from a library source such as those available from the USGS (Clark *et al.*, 2007) and are referred to reference endmembers (Roberts *et al.*, 1997). Ustin *et al.* (1994) used SMA to predict dry grass biomass with AVIRIS and Landsat TM over an 80 km area in California. They derived three endmembers for green vegetation (GV), non-photosynthetic vegetation (NPV) which included senesced grass, and soil. Shade is often included as an endmember and these four endmembers are used to derive a fraction image per endmember which presents the prevalence of that endmember in each pixel, and an image describing the error in the linear model (Smith *et al.*, 1990a). Roberts *et al.* (1998) elaborated on the method to map vegetation using multiple endmembers of spectra representing pure components and spectra representing variation in components modelled as linear mixtures of the pixel spectrum. They referred to this method as multiple endmember spectral mixture analysis (MESMA) whereby the number of endmembers and the types of endmembers vary across the scene (Roberts *et al.*, 1998). Thus, to differentiate bare patches, bare grass patches and sparse grass patches, two endmembers could be a pure soil spectrum and a pure sparse grass spectrum. However, this approach might not separate bare grass patches from sparse grass patches with a low canopy cover. As described in chapter 2, there is some evidence that the causes of these bare patches may be due to a variety of factors such as intermittent drainage lines, areas of pooling of run-off, very shallow soil profiles due to near-surface rock outcrops, or grazing impacts and recovery there from (Figure 2.11). Furthermore, these bare grass patches seem to be occupied by a variety of grass species that are not found in the matrix grasslands, and from field observations, seem to be specific to the possible causes of these bare grass patches. Thus it would be prudent to include a variety of endmembers that cover this range in types of bare grass patch and species composition. Other endmembers would include the minor patch types found at Vaal River such as tall grass, *Schoenoplectus* and *S. plumosum* patches.

A different approach might be to focus on the differences in floristic composition of these two patch types. Schmidtlein and Sasson (2004) used PLSR and AVIS-2 airborne hyperspectral

data to model and map floristic gradients in Bavaria. Their spectral data covered only the visible and NIR region (400 – 874 nm) with pixels of 2 x 2 m. A transformation of the reflectance to $\log(1/R)$ was used in their PLSR modelling and an initial PLSR to select a subset of bands from the loadings for their final PLSR. As response variables in their models, they used Detrended Correspondence Analysis (DCA) scores, three Ellenburg indicators for water supply, nitrogen supply and soil pH, as well as cover values for four species. For the DCA scores and Ellenburg indices, the final models produced validation coefficients of determination of between 0.66 and 0.76 whereas the species cover performed less robustly with r^2 values between 0.22 and 0.55. These coefficients of determination are generally slightly weaker than the final patch model for West Wits where combining bare grass and bare patches into one group, sparse grass and grass patches into another and *S. plumosum* as a patch type on its own produced r^2 values of 0.78, 0.78 and 0.86 respectively for a 6-component three-patch model. Similarly, the final four-patch model for Vaal River with biological soil crusts and bare grass forming one group, another group composed of sparse grass, grass patches and tall grass patches, and *Schoenoplectus* and *S. plumosum* as individual patch types had coefficients of determination of 0.76, 0.73, 0.25 and 0.81 for a 14-component model with some gypsum encrusted patches included in the model as a confounding variable. One major difference between Schmidtlein and Sassin's study and this study's patch modelling is that their floristic gradients were continuous whereas the patch types in this study, especially at the resolution of the ground-based spectral measurements, tend to be discrete entities with the exception of the grass matrix. Schmidtlein and Sassin (2004) repeat Lewis's (1998) claim that species composition is difficult to discern from spectral data due to the influence on plant reflectance of factors such as structure, physiology and phenological aspects of plants.

3.4.2 **Predicting Soil Stability, Infiltration and Nutrient Cycling**

Models predicting the LFA indices mark a move away from a matrix of response variables coded with dummy values (0 or 1) for each patch type, to a single numeric response variable on an interval scale. This has the advantage that the predictions with validation data are a single outcome and not a matrix of patch types where one patch type may do well at a certain level of components while another may do poorly but well at a different number of components in the model. Thus, we can now explore the issues around selecting the

number of components for an optimum model from both the model fit and cross-validated RMSEP with calibration data, and the outcome of all possible models through the predictions with validation data. Mevik and Wold (2007) make the point that no “generally applicable test” exists for efficient model selection but the norm is to select the model in which there is no further reduction in the cross-validation error or RMSEP. However, with the complex data modelled in this research the outcome of LOO cross-validation was often a number of local RMSEP minima in which the first may not be the lowest. The RMSEP value itself does not indicate whether a model is strong or weak as seen in the West Wits model predicting the LFA infiltration index (Table 3.18), a relatively strong model with a good fit between predicted and measured values, where the RMSEP values were higher than those for the LFA stability index at Vaal River which was a weak model with a low fit between predicted and measured values (Table 3.17).

It should be noted at this point that regardless of the RMSEP, adding a component always leads to a better prediction coefficient of determination with the calibration data (Table 3.18 and 3.19), presumably because the PLSR modelling process “knows” the prediction value it is aiming for. This improvement continues until the number of components reaches some point where no more variability is left in the data and adding more components has no effect. For predictions with validation data there is usually a maximum value in the coefficient of determination at some number of components and then the r^2 value may drop off or may oscillate before reaching another local maximum before falling to some lower but stable value with only minor oscillations as components are added to the model (Table 3.18 and 3.19). So the key is to find the number of components in the model that will correspond with this maximum r^2 value for validation data as this will give the strongest predictions with new data. The calibration coefficients of determination do, however, provide evidence of whether the initial model is a strong or weak model. If the model calibration r^2 values start low and the increment per added component is small, as shown in the Vaal River model predicting stability (Table 3.17) where even after 30 components the model hasn’t managed to explain all the variability in the stability index, then the model is likely to be weak. This is borne out in the validation coefficients of determination where the maximum ($r^2 = 0.398$) is only reached after 21 components although an alternative model with 4 components reached a local maximum of 0.392. The infiltration model for West Wits illustrates the

general pattern if the model is strong (Table 3.18), with the first component having a relatively high r^2 value (0.425) with calibration data and large increases for the addition of each of the next few components with all the variability in the infiltration index explained by the 25th component.

In the presentation of the PLSR model results for the LFA indices, the coefficients of determination for validation predictions are included for each component. It is clear that the maximum r^2 value for validation doesn't always coincide with a RMSEP minimum or with a 2% decrease in RMSEP per added component. For instance, when predicting the LFA infiltration index for West Wits, the best model from calibration data was a 5-component model that fulfilled both the criteria of being a RMSEP minima and reducing the RMSEP by 14.9% (Table 3.18). However, for validation data, the best model was a 6-component model which was not a RMSEP minimum and increased the RMSEP by 1.7% (Table 3.18). Similarly, the best model for predicting the LFA nutrient cycling index with calibration data from West Wits was a 6-component model which again was a RMSEP minima and each of the six components reduced the RMSEP by more than 2%, but a strong validation model was a 5-component model ($r^2 = 0.75$) with the strongest model a 16-component model ($r^2 = 0.83$, $P < 0.0001$, Table 3.20). The 5- or 6-component model is statistically better because it is more parsimonious but the examples do illustrate that the two methods (RMSEP minima and 2% per component) provide direction towards a good model but not necessarily the optimum model.

3.4.3 The PLSR Models Predicting the Ecological Variables

In selecting the optimum model for the PLSR models predicting the ecological variables, Table 3.22 and 3.23 for West Wits and Vaal River respectively allows the comparison of the RMSEP for the best model for each variable with the standard deviation for the measurements of that variable. Generally, the stronger models had RMSEP values lower than their standard deviation (Fresh above-ground biomass at West Wits $\bar{x} = 1.19 \pm 0.74$ kg/m², RMSEP = 0.48 kg/m²; Table 3.22), but so did some models with weak predictions such as the Vaal River model predicting fresh root biomass ($\bar{x} = 1.19 \pm 1.42$ kg/m², RMSEP = 1.24 kg/m²; Table 3.23). However, the general trend for weak models was to have a RMSEP closer to or greater than the standard deviation (Table 3.22 and 3.23).

3714 Of the various vegetation and soil variables that were measured in this study and apart from
3715 above-ground biomass, only soil moisture produced relatively strong models for both sites.
3716 At West Wits, a 7-component model was strongest (calibration $r^2 = 0.70$, $P < 0.0001$, $n = 57$;
3717 validation $r^2 = 0.45$, $P = 0.00017$, $n = 19$; Table 3.15 and Figure 3.32a) while for Vaal River, an
3718 8-component model was strongest (calibration $r^2 = 0.59$, $P < 0.0001$, $n = 180$; validation $r^2 =$
3719 0.68 , $P < 0.0001$, $n = 59$; Table 3.16 and Figure 3.32b). Ben-Dor *et al.* (2002) mapped a
3720 number of soil properties in Israel using an airborne DIAS-7915 sensor and the Visible and
3721 Near Infrared Analysis (VNIRA) approach. The VNIRA approach uses forward multiple
3722 regression analysis to select wavelengths giving the strongest prediction values for the soil
3723 property concerned (Ben-Dor *et al.*, 2002, Ben-Dor and Banin, 1995, Ben-Dor and Banin,
3724 1994). They found their models could predict soil moisture in lab conditions with
3725 coefficients of determination of 0.85 which dropped to 0.65 in field conditions (Ben-Dor *et*
3726 *al.*, 2002). Their field results compare favorably with the ground-based field results from
3727 both Vaal River and West Wits. However, their models used two wavelengths for field
3728 models (739 and 1650 nm) and a single wavelength (689 nm) with intercept term for
3729 laboratory-based models (Ben-Dor *et al.*, 2002) whereas both West Wits and Vaal River
3730 were full-spectrum models. Furthermore, their study only focused on soil pixels as they
3731 masked out vegetation related pixels in their airborne data whereas these models predicting
3732 soil moisture in this study included all samples whether soil only or various degrees and
3733 types of vegetation cover.

3734 Generally, soil moisture has been mapped using microwave-based remote sensing (Lakshmi
3735 *et al.*, 2001, Engman, 1984) as microwaves are little affected by atmospheric conditions and
3736 vegetation, and has high contrast between wet and dry soils. Generally, increasing the
3737 moisture content of soil reduces overall reflectance but the shape of the spectrum remains
3738 unaltered (Wessman, 1991). However, above-ground vegetation components would
3739 obscure soil surfaces from remote sensing sensors but these same vegetation components
3740 are in intimate contact with below-ground processes and thus vegetation canopy spectral
3741 properties may indirectly reflect conditions in soil (Ji and Peters, 2003, Wessman, 1991). For
3742 instance, many studies have shown a direct relationship between NDVI and rainfall (Zribi *et*
3743 *al.*, 2010, Paruelo and Lauenroth, 1995, Nicholson and Farrar, 1994) although these
3744 relationships are complex and affected by many factors such as vegetation type (Richard

3745 and Pocard, 1998, Peters *et al.*, 1997, Paruelo and Lauenroth, 1995), soil type (Farrar *et al.*,
3746 1994, Nicholson and Farrar, 1994), and spatial and temporal scales (Eklundh, 1998).
3747 Furthermore, Giraldo *et al.* (2008) showed that there was a strong correlation at small
3748 spatial scales (30 x 30 m plots) between soil moisture and land use/ land cover categories
3749 that was both spatially and temporally consistent. The fact that the West Wits models are
3750 stronger than the Vaal River models may reflect the influence of season. The West Wits data
3751 was collected in the earlier winter dry season when the grasses had lost their pigments and
3752 just entered their senescent phase, and thus may still have reflected soil conditions in their
3753 moisture content. Vaal River sites were sampled mid to late winter dry period and may thus
3754 have been fully cured by that stage, and therefore were more decoupled from soil
3755 conditions. This all supports the results of the PLSR models predicting soil moisture from
3756 ground-based portable spectroscopy despite the land cover being a mix of green
3757 photosynthetic canopy (*S. plumosum*), senescent canopy (grass) and soil surface with and
3758 without biological soil crusts.

3759 The PLSR models predicting soil organic matter (SOM) were fairly strong at both West Wits
3760 and Vaal River. At West Wits, Calibration models were considerably stronger ($r^2 = 0.69$, $P <$
3761 0.0001 , $n = 57$) than validation models ($r^2 = 0.32$, $P = 0.019$, $n = 17$; Table 3.15 and Figure
3762 3.31a). Whereas at Vaal River, calibration models were slightly weaker ($r^2 = 0.55$, $P < 0.0001$,
3763 $n = 177$) than those at West Wits, but the validation model for Vaal River was stronger ($r^2 =$
3764 0.50 , $P < 0.0001$, $n = 59$; Table 3.16 and Figure 3.31b) than the West Wits validation results.
3765 Numerous studies have examined the feasibility of using remote sensing to predict SOM
3766 (Wight *et al.*, 2016, Gomez *et al.*, 2008, Stevens *et al.*, 2008, Stevens *et al.*, 2006, Ben-Dor *et*
3767 *al.*, 1997). Fidêncio *et al.* (2002) was able to predict organic matter in soil samples with a
3768 validation r^2 of 0.95 with 30 samples. Peltre *et al.* (2011) also used PLSR to predict soil
3769 organic matter ($r^2 = 0.80$) from NIR spectroscopy in a laboratory experiment with soils
3770 organic matter amendments, and Wetterlind *et al.* (2010) was able to predict SOM with
3771 PLSR with r^2 values between 0.81 and 0.89. Kusmumo *et al.* (2011) predicted soil organic
3772 carbon (SOC) using PLSR with a calibration r^2 of 0.86, and Viscarra Rossel and Behrens (2010)
3773 tested six different methods to predict SOC of which PLSR produced the second best results
3774 ($r^2 = 0.82$) for the stand-alone methods. In all these studies, the models were derived from
3775 spectra collected from soil samples under laboratory conditions. Whereas in the Vaal River

and West Wits sites, spectra were collected under field conditions with the soil surface often covered at least partially by biological soil crusts, or with a vegetation canopy of either senesced grass standing litter or in the case of the shrub *S. plumosum*, a green photosynthetically active canopy. Therefore, although the results achieved in this study are substantially lower than those quoted in other studies, under field circumstances in which measurements were made, the predictions of SOM by PLSR can be considered reasonable.

Electrical conductivity (EC) is a measure of the salt content of soils and models predicting EC at West Wits with ground-based spectroscopy had moderate calibration statistics ($r^2 = 0.55$, $P < 0.0001$, $n = 57$) but very weak validation statistics ($r^2 = 0.10$, $P = 0.18$, $n = 19$; Table 3.15). Vaal River on the other hand, had quite strong PLSR models predicting EC with calibration statistics ($r^2 = 0.53$, $P < 0.0001$, $n = 180$) being slightly weaker than validation statistics ($r^2 = 0.63$, $P < 0.0001$, $n = 59$; Table 3.16 and Figure 3.33a). These differences between the two mining regions are a result of the presence of visible salt crusts or evaporate on some transects at Vaal River and its absence at West Wits. Values for EC from soil extracts at West Wits were very low ($18 - 143 \mu\text{S}\cdot\text{cm}^{-1}$), conversely the range at Vaal River was much more extreme ($12 - 2546 \mu\text{S}\cdot\text{cm}^{-1}$) with all transects having higher mean values (Table 2.4) than the maximum range at West Wits and the high values being strongly associated with transects where evaporite was visible. No chemical examination was carried out on the evaporate so its chemical constitution is unknown but is assumed to be gypsum or $\text{CaSO}_4\cdot 2\text{H}_2\text{O}$ (Herrero *et al.*, 2009). However, as the underlying surface geology is dolomite, these evaporites may include calcium and magnesium carbonates, as well as other salts, although a one-off field test with acid did not indicate the presence of carbonates. It was also noted that the presence of evaporite caused a major increase in the overall brightness in the visible, NIR and SWIR portions of the spectrum up to 1750 nm for biological soil crust patches (Figure 3.15). Presumably this effect was somewhat masked in the bare grass and sparse grass patches which were affected by evaporate crusts due to the intervening grass canopy between sensor and soil surface where evaporite crusts develop. However, Nawar (2014) showed the opposite effect with CaCO_3 affected soils in laboratory conditions with increasing EC values led to a reduction in spectral brightness across the spectrum.

Many studies have used hyperspectral remote sensing to predict and map EC (Farifteh *et al.*, 2010, Dutkiewicz *et al.*, 2009b, Bannari *et al.*, 2008, Dutkiewicz *et al.*, 2006) and some have

3807 applied PLSR (Nawar *et al.*, 2014, Weng *et al.*, 2008, Farifteh *et al.*, 2007) to predict EC in
3808 soils. Nawar *et al.* (2014) used PLSR to predict EC in the Sinai, Egypt, and had coefficients of
3809 determination of 0.71 but they were dealing with a range of EC values much higher than
3810 those at Vaal River and the amount of vegetation cover is not clear but probably
3811 considerably less than that at West Wits and Vaal River. Likewise, Weng *et al.* (2008)
3812 predicted EC with PLSR on the Yellow River delta and achieved a calibration r^2 of 0.89 and
3813 validation r^2 of 0.57. Similarly, Farifteh *et al.* (2007) used PLSR models to predict EC at a
3814 variety of sites and achieved r^2 values of between 0.6 and 0.98 depending on the salt and
3815 the soil. They found that PLSR performed better over sandy soils compared to sandy loam
3816 soils. These results are all stronger than those obtained from West Wits and Vaal River with
3817 ground-based spectral measurements although all these studies are in areas with known
3818 salinization issues, a suggested ground cover lower than that at the two mining sites, and EC
3819 values considerably higher than those at either mining region.

3820 Predicting pH at West Wits with PLSR resulted in weak calibration models ($r^2 = 0.30$, $p <$
3821 0.0001 , $n = 57$) and moderately strong validation results ($r^2 = 0.57$, $P = 0.0002$, $n = 19$; Table
3822 3.15) whereas Vaal River had stronger calibration statistics ($r^2 = 0.63$, $P < 0.0001$, $n = 180$)
3823 but weak validation statistics ($r^2 = 0.23$, $P = 0.0002$, $n = 59$; Table 3.16). This suggests that
3824 the models for predicting pH at the two mining regions were not very robust. A number of
3825 studies have attempted to characterize pH through spectral means in either mining (Ong
3826 and Cudahy, 2014, Mars and Crowley, 2003, Swayze *et al.*, 2000, López-Pamo *et al.*, 1999) or
3827 agricultural contexts (Wetterlind *et al.*, 2010, Viscarra Rossel *et al.*, 2006, Ben-Dor *et al.*,
3828 2002) but pH itself does not have a spectral profile. These studies have either focused on
3829 identifying mineral changes associated with changes in pH (Swayze *et al.*, 2000, López-Pamo
3830 *et al.*, 1999, Bigham *et al.*, 1996) or directly calibrated models from soil spectra in which the
3831 pH was known but the exact mechanism behind the spectral response was unknown
3832 (Wetterlind *et al.*, 2010, Viscarra Rossel *et al.*, 2006, Ben-Dor *et al.*, 2002).

3833 Viscarra Rossel *et al.* (2006) showed that the visible portion of the spectrum had the
3834 weakest PLSR model predictions ($r^2 = 0.36$, $P < 0.001$), followed by the NIR portion of the
3835 spectrum ($r^2 = 0.54$, $P < 0.001$) with the mid-infrared portion of the spectrum or a
3836 combination of the three having the strongest ($r^2 = 0.75$, $P < 0.001$ or $r^2 = 0.73$, $P < 0.001$
3837 respectively), which, apart from the MIR or combination models, are in the same range of

3838 the results achieved at West Wits and Vaal River. Ben-Dor *et al.* (2002) using multiple linear
3839 regression achieved coefficients of determination of 0.53 and 0.88 when predicting pH using
3840 a two-wavelength (722 and 2118 nm) model with intercept term with both field and
3841 laboratory spectra respectively. However, all these studies have focused on exposed soil in
3842 the lab, or a combination of exposed soil in the field and laboratory. Both West Wits and
3843 Vaal River have some exposed soil patches but the majority of the two areas have some
3844 degree of vegetation cover, generally in a winter senesced state, obscuring the soil surface.
3845 Wessman (1991) makes the point that soil processes *per se* cannot be detected by spectral
3846 remote sensing methods through a canopy, but as plants are in intimate contact with sub-
3847 surface soil processes, they may reflect conditions in the soil. At Vaal River, there was no
3848 significant correlation between pH and any other variable, whereas at West Wits there was
3849 a strong correlation between pH and fresh or dry above-ground biomass ($r = 0.63$ and 0.65
3850 respectively; Figure 2.8). This difference may reflect differences in geology where the West
3851 Wits sites were underlain by shale and quartzite, while Vaal River sites were situated over
3852 dolomites with chert bands. However, it does reflect the somewhat mixed results of the
3853 PLSR models in predicting pH at patch scales in the two mining regions.

3854 3.5 Conclusion

3855 The purpose of this research was to test if LFA indices could be predicted from spectral
3856 measurements of grasslands with PLSR modelling. The objective of this chapter was to test
3857 this with ground-based spectra collected from two mines: West Wits and Vaal River mining
3858 regions. It was shown that ground-based spectral measurements and PLSR could
3859 successfully predict patch types at both mining sites, although the spectra and PLSR
3860 struggled to separate bare grass patches from bare or biological soil crust patches and
3861 sparse grass patches. The reason for this is that these three patch types do not form distinct
3862 spectral types, but are rather on a continuum with bare or biological soil crust patches at
3863 one end and sparse grass patches at the other, with bare grass patches in between. The
3864 PLSR models consistently separated *S. plumosum* patches from other patch types. However,
3865 models were not convincing in separating tall grass and *Schoenoplectus* patches although
3866 there were very few replicates of these patch types which probably contributed to this
3867 weakness in the models. The results for the PLSR models predicting the three LFA indices of
3868 soil stability, infiltration and nutrient cycling from ground-based hyperspectral
3869 measurements are presented in this chapter, but as these indices are also derived from
3870 PLSR models using satellite derived hyperspectral data, these results are discussed together
3871 with the results of the satellite derived models in the next chapter. Likewise, the results for
3872 PLSR models predicting biomass from ground-based spectra are presented in this chapter
3873 but discussed together with the results from the PLSR models predicting biomass from
3874 Hyperion hyperspectral data in the next chapter.

3875 Amongst the ecological variables, PLSR was able to predict soil moisture with fairly strong
3876 results. The models were stronger with West Wits data and this may have been because
3877 West Wits was sampled early in the dry winter season while Vaal River was sampled late in
3878 the dry winter season. This difference may have meant the senesced grasses at West Wits
3879 were, to some degree, still reflecting the below ground conditions. Whereas, the Vaal River
3880 grasses may have been fully cured and hence, more decoupled from soil moisture
3881 conditions, and as the senesced grass canopy would mask the soil from the hyperspectral
3882 sensors perspective, this may have lead to weaker models for Vaal River sites. Models
3883 predicting soil organic matter or SOM were mixed with West Wits giving fairly strong
3884 calibration models but weak validation models and Vaal River producing moderately strong

3885 models with both calibration and validation data. It is not clear why West Wits produced
3886 weak validation results but the presence of vegetation canopies over most sites may have
3887 partially masked the soil surface and contributed to weakness in the models.

3888 Models predicting soil solution EC at West Wits were very weak whereas those at Vaal River
3889 were quite strong. This is believed to be due to the presence of evaporite crusts, presumed
3890 to be gypsum, at Vaal River which had a much larger range of EC values than did West Wits
3891 sites which were generally very low. Models predicting soil solution pH were generally weak
3892 although the calibration model for Vaal River was moderately strong but the validation
3893 results were weak. Models for other ecological variables like gravel content, soil density or
3894 any of the soil nitrogen measurements were weak, probably because none of these have a
3895 direct spectral effect and so predicting these variables would require some sort of spectral
3896 proxy.

3897 Going forward, it would be interesting to see how the patch analysis would play out with
3898 summer data. This would bring plant pigments into the spectra and may produce more
3899 subtle differences between species, especially between bare grass and sparse grass patches
3900 where different disturbance regimes result in fundamental differences between grass
3901 species occupying these two niches which may be masked by winter senescence conditions.
3902 In terms of mapping patches it would also be better not to follow the LFA transect structure,
3903 but to ensure good replication of patch types and representivity of all patch types in the
3904 sampling. Spectral preprocessing before PLSR modelling would also be a useful exercise.
3905 Two preprocessing techniques that may be useful are continuum removal (Clark and Roush,
3906 1984) which tends to make the absorption features far more explicit, and first derivative of
3907 the spectra may also be useful as these highlight the inflection points in the spectra.
3908 Nonetheless, this research into the use of PLSR models to predict LFA indices has shown that
3909 in general and with a useful degree of accuracy, PLSR can successfully predict the LFA
3910 indices, patch structure and some edaphic or ecological variables.

3911

4. Predicting LFA Indices from Hyperion Hyperspectral Imagery



3912
3913

3914 4.1.1 The Hyperion Hyperspectral Sensor

3915 The Earth Observation 1 satellite (EO-1) was launched by NASA on 21 November 2000 with
3916 three experimental sensors aboard (Ungar *et al.*, 2003, Ungar, 2001). These were the
3917 Hyperion sensor, the Advanced Land Imager (ALI), and the Linear Etalon Imaging
3918 Spectrometer Array (LEISA) Atmospheric Corrector (LAC) (Pearlman *et al.*, 2001). Of these
3919 sensors, the Hyperion sensor is the first high spatial resolution spaceborne hyperspectral
3920 sensor deployed in space (Ungar, 2001). The other sensors aboard EO-1, being broadband
3921 sensors, do not form part of this study although the ALI can capture a LandSat type image at
3922 the same time of the same area that Hyperion captures a hyperspectral image. The
3923 Hyperion sensor is a pushbroom-type sensor, also referred to as an along-track scanner,
3924 with a single telescope and two spectrometers: a visible/near infrared (VNIR) spectrometer
3925 and a shortwave infrared (SWIR) spectrometer (Pearlman *et al.*, 2001). Pushbroom sensors
3926 have a linear array of light detectors arranged perpendicular to the direction of sensor
3927 travel. The telescope focuses the image through a slit such that the instantaneous field of
3928 view (FOV) is 0.624° which, at 705 km away, creates a swath width on the earth's surface of
3929 7.5 km (Pearlman *et al.*, 2001). The spatial resolution is 30 m with a spectral resolution of 10
3930 nm and a spectral range from 400 – 2500 nm across 220 bands for continuous spectral
3931 coverage (Ungar *et al.*, 2003). Hyperion produces an image composed of square pixels
3932 where each pixel covers approximately 900 m² of the Earth's surface with an image width of
3933 7.5 km and a minimum length of 42 km of the Earth's surface. Within these horizontal
3934 dimensions, each pixel has its own spectrum (Kumar *et al.*, 2010) composed of a stack of
3935 220 image layers at 10 nm intervals. Thus, each Hyperion image is a three dimensional
3936 datacube with the spatial component constituting the horizontal dimensions, and the
3937 intensity of reflectance at each wavelength constituting the vertical component.

3938 The EO-1 satellite was originally designed with a life of two years for a one-year validation
3939 mission of the next generation imaging technologies (Ungar, 2001). Since the EO-1 mission
3940 objectives have all been accomplished (Ungar *et al.*, 2003) and the satellite is still collecting
3941 images, the United States Geological Survey (USGS) Earth Resources Observation and
3942 Science (EROS) center have made Hyperion data freely available to the public since the
3943 5 August 2009. Archival images are available by searching the database and new images can
3944 be requested through a Data Acquisition Request (DAR) filed over the internet. The DAR

3945 requires a latitude, longitude, and period within which to fulfill the request. The USGS
3946 makes no guarantee that DARs will be successful, nor does it guarantee the quality of the
3947 images acquired.

3948 HRS and PLSR provides a robust set of techniques for deriving models from spectra to
3949 predict ecosystem processes such as soil stability, infiltration and nutrient cycling. The
3950 spectral characteristics of minerals, soil and plants discussed briefly in the previous chapter
3951 have been used to study many environmental aspects and processes. Foster *et al.* (2002)
3952 used Hyperion to predict total carbon, above-ground carbon, carbon in wood, understory
3953 carbon and soil carbon in the Bolivian Amazon using multiple regression approaches. Huete
3954 *et al.* (2003) applied a number of techniques including Vis, spectral mixture analysis, spectral
3955 derivatives and albedo with Hyperion hyperspectral imagery to characterize land
3956 degradation patterns in the floristically diverse Monte Desert region of Argentina and found
3957 that spectral mixture analysis performed best. Dutkiewicz *et al.* (2009a, 2006) compared
3958 Hyperion hyperspectral imagery with two airborne hyperspectral sensors to map symptoms
3959 of salinization in drylands in Australia and found that both spectral and spatial resolution
3960 was important to accurately characterize saline features in these landscapes. White and
3961 Abuelgasim (2010) used a 2003 Hyperion image to monitor the vegetation health of
3962 rehabilitation programs instituted since the 1970s to mitigate the impacts of 100 years of
3963 mining and smelting around the Sudbury area of Ontario, Canada. They showed that
3964 equivalent water thickness and leaf area index of rehabilitated vegetation increased with
3965 distance from Sudbury. Campbell *et al.* (2013) used Hyperion hyperspectral imagery over a
3966 number of years to track the seasonal profiles and ecosystem carbon exchange at two sites
3967 in North America: a tallgrass prairie and a hardwood forest, and a third site which was
3968 Miombo Woodland in Zambia, Africa. Wang *et al.* (2014) used an unsupervised multiple
3969 endmember approach with Hyperion imagery to identify and map poppy fields in
3970 Afghanistan which resulted in an over-prediction of 12% of pixels.

3971 4.2 **Methods and Materials**

3972 4.2.1 **Preparation of the LFA Data**

3973 The LFA data was collected as outlined in detail in chapter 2. The use of a tape measure and
3974 GPS points to define the LFA transects allowed the transects to be laid over the Hyperion
3975 image. Each 180 m transect was segmented into sub-transects at the point where pixel
3976 boundaries intersected the transects (Figure 4.1). As some transects ran diagonally across
3977 portions of pixels, any sub-transect length less than two thirds (20 m) of the length of its
3978 corresponding pixel (30 m) was discarded from the analysis, as was the corresponding pixel.
3979 This decision was made as it was assumed that short LFA transects (< 20 m long) that
3980 bisected the corners of a pixel might not adequately represent the landscape variability in a
3981 30 x 30 m area captured in the spectrum of a pixel (Figure 4.1). As a result, only pixels and
3982 sub-transects which had greater than 20 m transect overlap were used in the study (red
3983 sub-transects in Figure 4.1). Once the sub-transects were defined, the LFA field data that
3984 corresponded with that sub-transect was entered into the LFA Microsoft Excel spreadsheet
3985 (Tongway and Hindley, 2004). Thus landscape organization index (LOI) and patch area index
3986 (PAI) values were calculated for each pixel corresponding to a LFA transect.

3987 The three LFA soil surface indicators (SSI): soil stability, infiltration and nutrient cycling, were
3988 calculated for each patch type found within the pixel transect using three replicates for each
3989 patch type. Replicates refers to the fact that a maximum of five patches of the same patch
3990 type along the 180 m length of the transect were sampled for calculating SSI values with the
3991 exception of the matrix patch type (generally sparse grass) where there were generally a
3992 few more patches sampled as explained in Chapter 2. However, a pixel did not always have
3993 three sampled points for each patch type on it. Therefore, the sampled patches that did fall
3994 on the pixel were used first, and if needed, the rest of the triplet made up from the sampled
3995 patches nearest to that pixel. A further issue with the SSIs is that they are calculated for a
3996 short (50 cm), essentially a pure length of patch, and do not reflect the patch type's
3997 contribution to the whole transect. A weighting for each patch type was thus derived based
3998 on the percentage of the sub-transect occupied by a specific patch type. By multiplying the
3999 mean SSI for the pure patch type by its weighting value, a weighted mean SSI for each patch
4000 type on the sub-transect was calculated. Finally, to calculate the overall SSI value for the

sub-transect, the weighted SSI values for each patch type occurring on the sub-transect were summed to give a single value for soil stability, infiltration and nutrient cycling for each sub-transect. Biomass measurements had similar issues and therefore a similar process was used to determine the total biomass on each pixel.

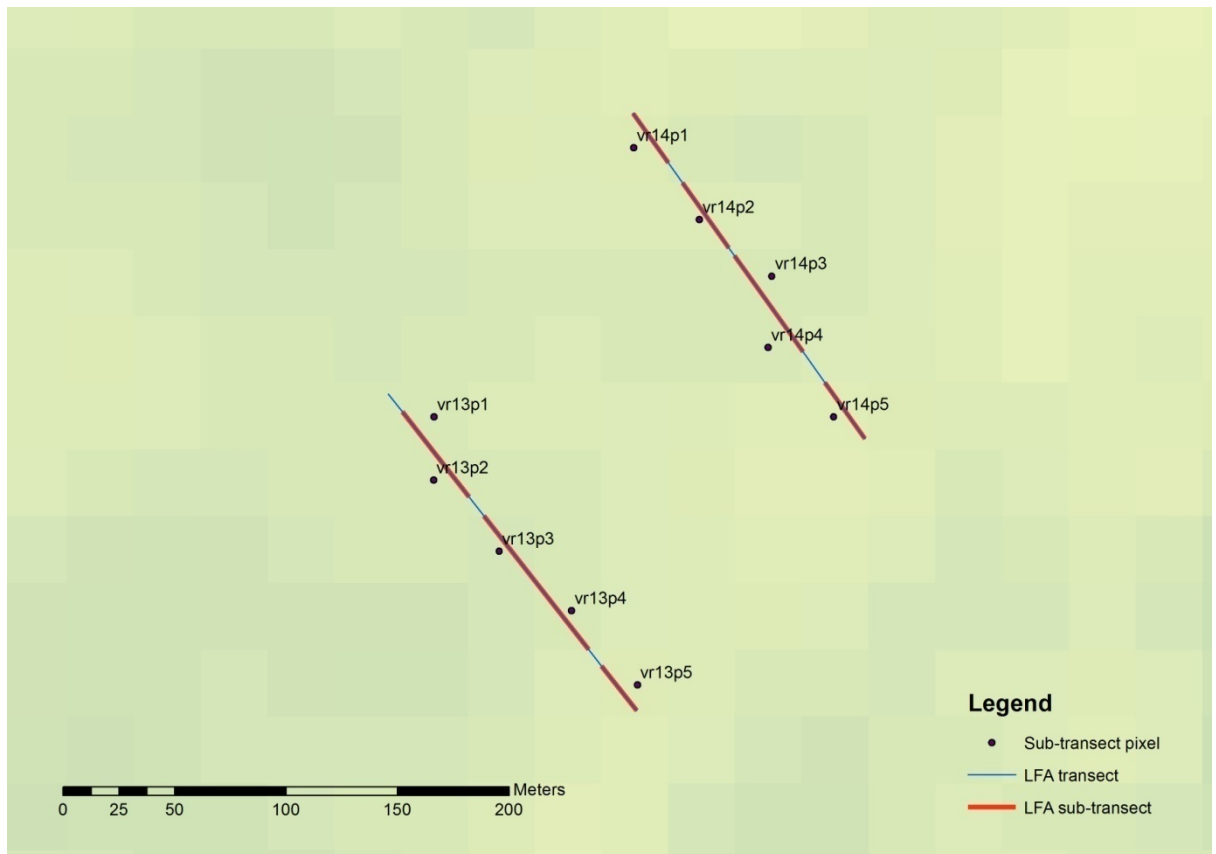


Figure 4.1 Partitioning of LFA transects into sub-transects corresponding to pixels.

4.2.2 Preparation of the Hyperion Hyperspectral Data

In 2013, USGS provided two Hyperion hyperspectral images for Vaal River in response to a data acquisition request (DAR) submitted in 2011. The first image just missed the transects by a couple of hundred metres and was not used in this study. The second image covered 6 transects from Vaal River but missed the other five transects (Figure 4.2). This image was acquired on the 2nd of November 2013 and downloaded from <http://glovis.usgs.gov/> by the 29 December 2014. This image was stacked into a data cube in ENVI 4.7 using Hyperion tools 2.0 (White, 2011), FLAASH corrected to convert at sensor radiance to surface reflectance, and reprojected from Transverse Mercator 35 S to UTM 27 S. The reprojection meant the Hyperion image and LFA data were in the same projection space and could be

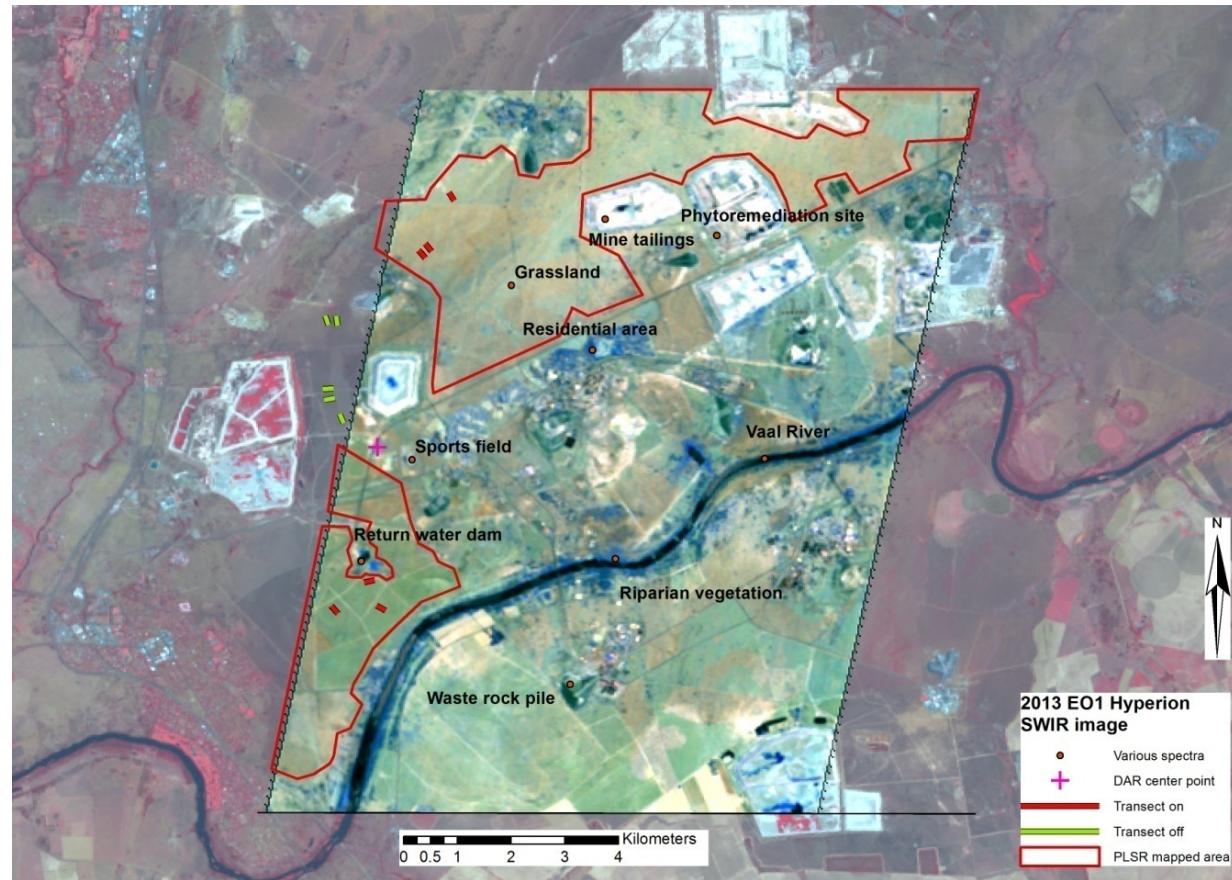


Figure 4.2

November 2013 Hyperion image of a portion of Vaal River mining region displayed as a SWIR image using the bands at 2194, 1649 and 1074 nm, and laid over a false infrared Landsat 8 image taken 2 days after the Hyperion image. The Hyperion image represents the cropped datacube used in the PLSR modelling. The pink cross is the coordinates from the DAR for the center of the Hyperion image. Red transects represents sample points on the Hyperion image and green transects represent sample points off the Hyperion image. Labeled orange points represent pixels selected from the Hyperion image to explore the reflectance spectra of various features in the Hyperion image (Figure 4.3) to ascertain the validity of the FLAASH correction. The red polygons represent grassland areas mapped to predict LFA indices and biomass using the respective PLSR models.

verified against a known, spatially defined data set, which was the South African Surveyor General's 2010 colour aerial orthophotos (National Geo-spatial Information, 2010). Testing the spatial accuracy between the two projections showed that the Vaal River Hyperion image had been accurately georeferenced by the USGS and required no re-registration to properly match the data (Figure 4.2). Overlapping bands, bad bands, and those affected by atmospheric gases, were removed from the original 242 bands, which left 155 bands in the data cube. These remaining bands were individually inspected by visual means for striping and other artifacts. Any band deemed of poor quality as a result of striping, smearing or speckling was then removed from the data cube resulting in 128 bands making up the z-axis of the data cube. However it should be noted that many of the bands included in the final data cube, especially those in the SWIR, still showed some faint striping in some area of the image. The final Vaal River Hyperion data cube was cropped to 399 columns and 448 rows (Figure 4.2) for importing into R for PLSR modelling and covers a representative area of the mining region. However, the spectra were still very noisy (Figure 4.3) so a smoothing algorithm, EOSap_smoothing from the Colibri spectral tools package (Bertels, 2012), was applied to all spectra in the data cube. The smoothing algorithm calculates a weighted mean for the wavelength of interest and its immediate neighbours and the smoothing is increased by iteratively applying the algorithm. One to four iterations were applied and the results from PLSR models using the original unsmoothed, "raw" Hyperion spectra and the smoothed spectra were compared (Figure 4.4).

The Hyperion image was acquired by NASA and the USGS two years after LFA sampling. The scene captures early summer whereas the LFA data was collected through winter into early spring. The rangeland seems to be in a relatively unchanged condition at the resolution of the Hyperion image although there is an indication of burning on transects 13 and 14, which just misses transect 15 (Figure 4.1). This was verified with Landsatlook georeferenced images sourced from the USGS and taken on July 7th, August 16th, September 1st, October 11th and November 4th which showed that the area around these two transects burnt between July 7th and August 16th 2013. There is no evidence that transects 1, 2 and 3, the southern transects in the RG Williams Reserve, were burnt during the 2013 season. However, an area about 750m southwest of transect 2 was burnt before June 5th, the section north of this and west of transect 2 by about 300 m was burnt by August 16th, and

between October 11th and November 4th a further small section was burnt about 500 m southwest of transect 2. There are no signs of a major change in land use around the 6 transects between 2011 and 2013. Although the early summer recovery (3 months and possible spring rains) from burning may have smoothed some of the fire induced differences between transects 13 and 14 and the other transects, the fire scar is still visible. However, taking into account the burn history, the two-year three-month difference between LFA ground sampling and image acquisition, and the slightly different phenological phases between ground sampling and image acquisition, the image was deemed suitable to provide a relative test of the application of PLSR modelling to predict LFA indices.

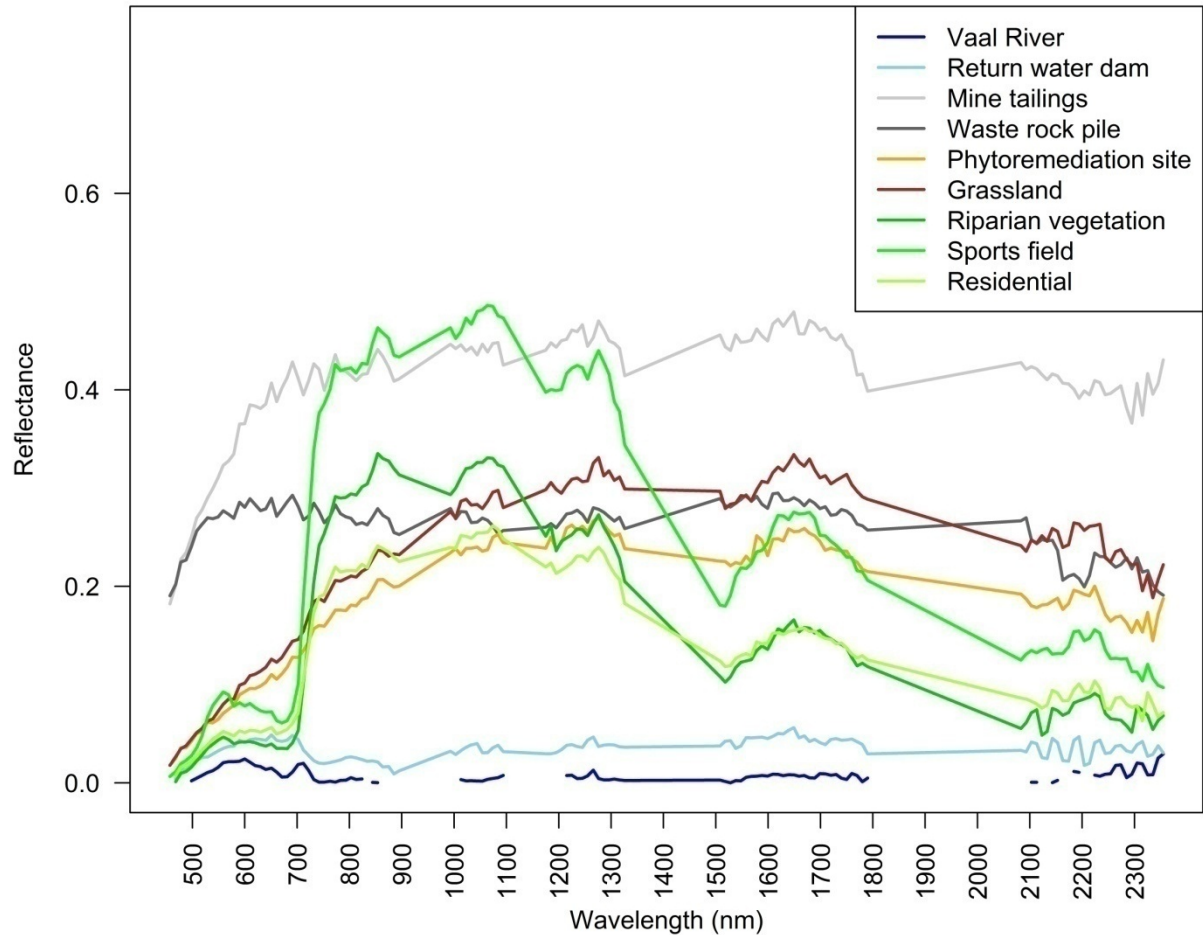


Figure 4.3 Surface reflectance spectra from a variety of features identified in the Hyperion image in Figure 4.2. The missing reflectance values in the spectrum from the Vaal River are due to the conversion of all negative values to NA during the preparation of the data. These areas of negative values correspond with the bad bands visually identified and removed from the data cube before initial analysis, and are identifiable as straight lines connecting spectral groups.

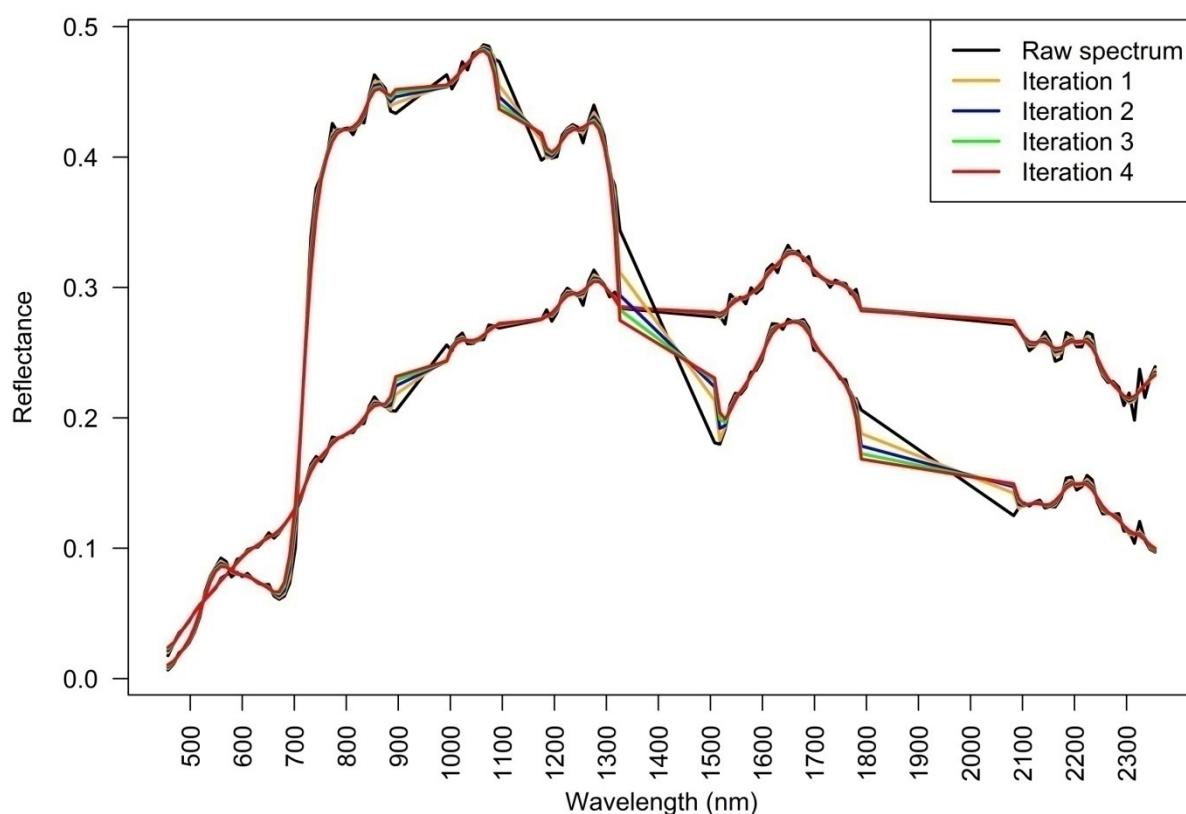


Figure 4.4 The effect of applying the Colibri smoothing algorithm at 1 to 4 iterations on two spectra extracted from the Hyperion image. The spectrum with the strong red edge at 700 nm was a pixel from the sports field in the Oppenheimer Stadium, and the flatter spectrum was a grassland pixel on transect 15.

4.3 PLSR Statistical Modelling and Mapping

The cropped Hyperion data cube was imported into R (R Core Team, 2013) using the raster package (Hijmans, 2014). Pixels in the cropped image with negative values or values greater than the scale factor determined during the FLAASH atmospheric correction were converted to not-a-number (NA). These were mostly confined to the river and a few pixels in the residential area that may have represented a tin roof or a swimming pool. The respective transect associated pixels were extracted, using points determined on the basis of a transect crossing at least 20 metres of a pixel, and converted into a matrix. This resulted in 31 pixels for the LOI and PAI index but herdsman set the rangeland near transect 15 alight while the ASD was being set up for spectral sampling and the entire section of grassland burnt out through the rest of the day. Thus for the LFA indices of stability, infiltration and nutrient

cycling, and above-ground biomass, the number of available pixels was reduced to 25. The LFA data was imported and both the Hyperion matrix and the LFA data split into validation and calibration data sets on a 1:2 basis. The result was 20 (16 after the fire in transect 15) pixels in the calibration data and 11 (9) in the validation data. The LFA data was then combined with the Hyperion matrices to form two data frames at each site for calibration and validation of the models. PLSR models were derived using the pls package (Mevik and Wehrens, 2007) with calibration data, and tested on the validation data. Once the number of components in the model was finalized, values were predicted for all pixels using the spectra from the entire Hyperion data cube and converted to a single-layer raster for display.

Table 4.1 Number of transects and sample points falling on Hyperion imagery from Vaal River, the maximum number of bands in the Hyperion data cube used for PLSR modelling and prediction mapping. The numbers in brackets reflect what was left after the fire lit by herdsman in the far corner of the rangeland burnt down transect 15 halfway through sampling. So 31 pixels were used for modelling LFA transect indices such as LOI and PAI, but only 25 pixels for other LFA indices and biomass.

Site	Transects	LFA points	Pixels on transect	Total Hyperion pixels in data cube	Number of pixels in grassland prediction area
Vaal River	6 (5)	178 (151)	31 (25)	178 353	29621

4110 4.3 **Results**

4111 4.3.1 **The Pixel Spectra Used in PLSR Modelling**

4112 The spectra from the 31 pixels used in the PLSR modelling group into three clusters
4113 according to the transects they corresponded to (Figure 4.5). Transects 1, 2 and 3 were a
4114 minimum of 460 m apart but form one group with very similar spectra and reflect their
4115 position within the R.G. Williams reserve in the southwestern portion of the Hyperion
4116 image. This area is located over a broad sub-surface seepage area seeping through the flood
4117 plain into the Vaal River and draining the West Complex Tailings Storage Facilities and the
4118 Black Reef Ridge a few hundred metres north of transects 10, 11, 13, 14 and 15 (Figure 4.2).
4119 Transects 1, 2 and 3 were not burnt through the winter of 2013 and therefore, their spectra
4120 should closely resemble the spectra of the winter field campaign of 2011, albeit with some
4121 adjustment for the fact that the spectra were collected during early summer where some
4122 greening of the grassland may have occurred. Thus, it is not entirely surprising that these
4123 three transects form a common spectral group.

4124 Transects 13 and 14 were only about 60 m apart from each other and form a second group.
4125 This area was burnt approximately four months before the Hyperion image was acquired
4126 and although it has had some time to recover, it is not surprising from both their nearness
4127 to each other and their similar history that the spectra from the ten pixels on these two
4128 transects form a distinct group separate from the other pixel spectra. The burn from this
4129 area is clearly visible in the Hyperion SWIR image (Figure 4.2) as a tan arc-shape around the
4130 northern edge of these two transects and ending at the dirt road just south of transect 15.
4131 The six pixel spectra from transect 15 form the third spectral grouping and this area was not
4132 burnt during the 2013 year so it is expected that it's pixel spectra would form a group
4133 different to transects 13 and 14. Furthermore, during the sampling period of 2011, transect
4134 15 was severely overgrazed with grass cropped almost to ground level. It was observed that
4135 cattle were herded onto it every afternoon that it was in view while sampling transects 13
4136 and 14. None of the other transects showed this degree of grazing pressure although all
4137 transects on the Hyperion image experienced some degree of grazing.

4138

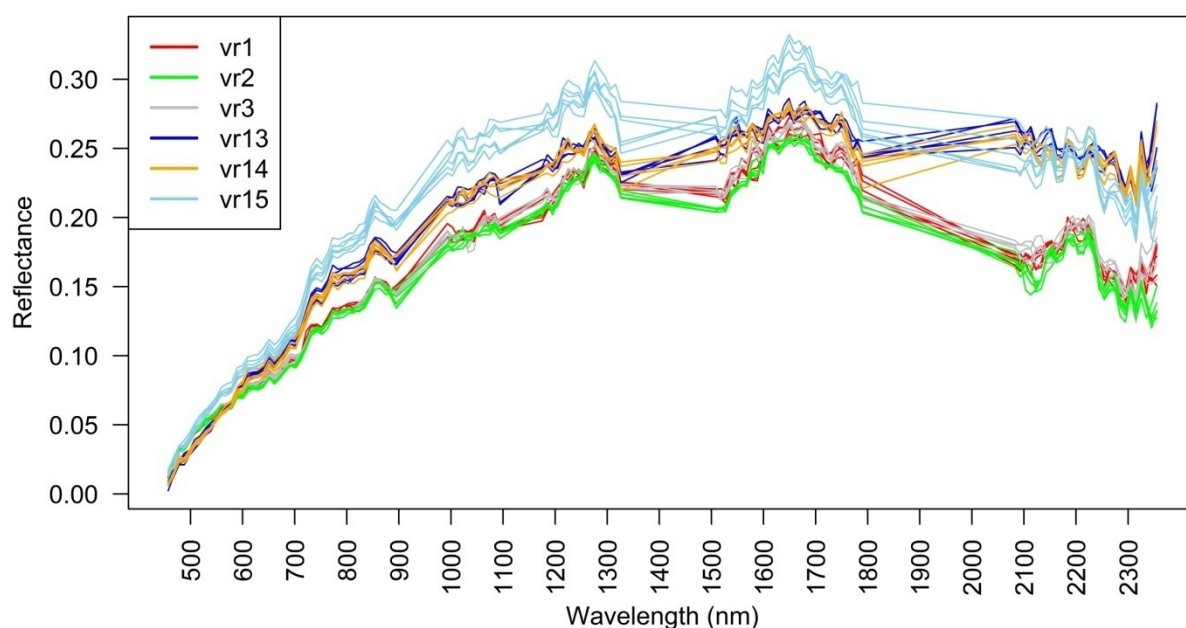
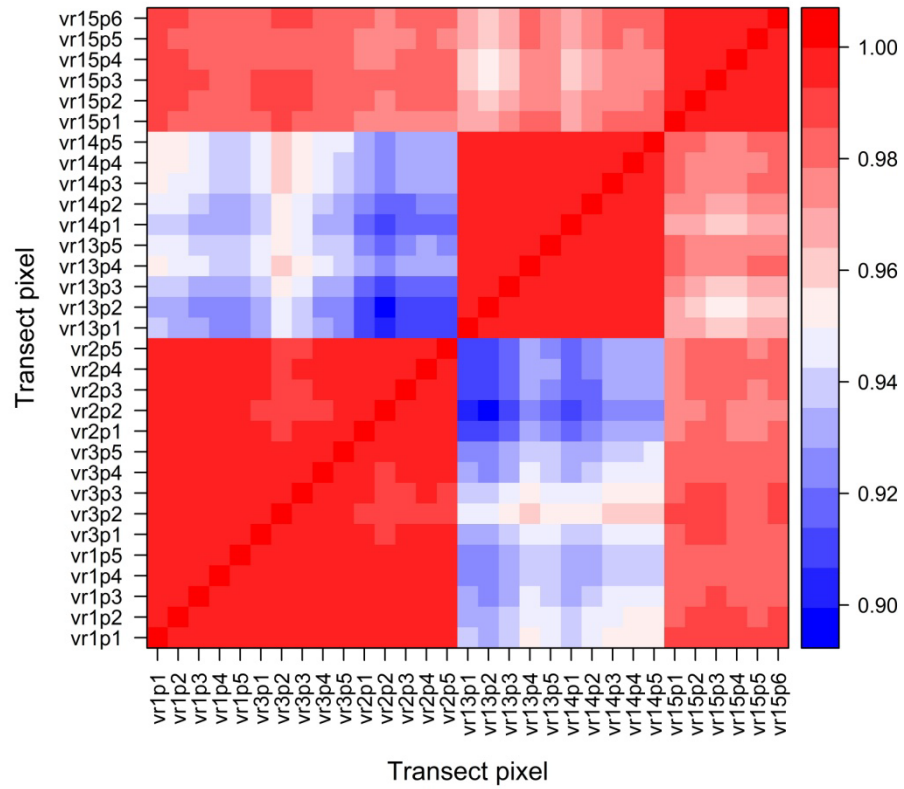


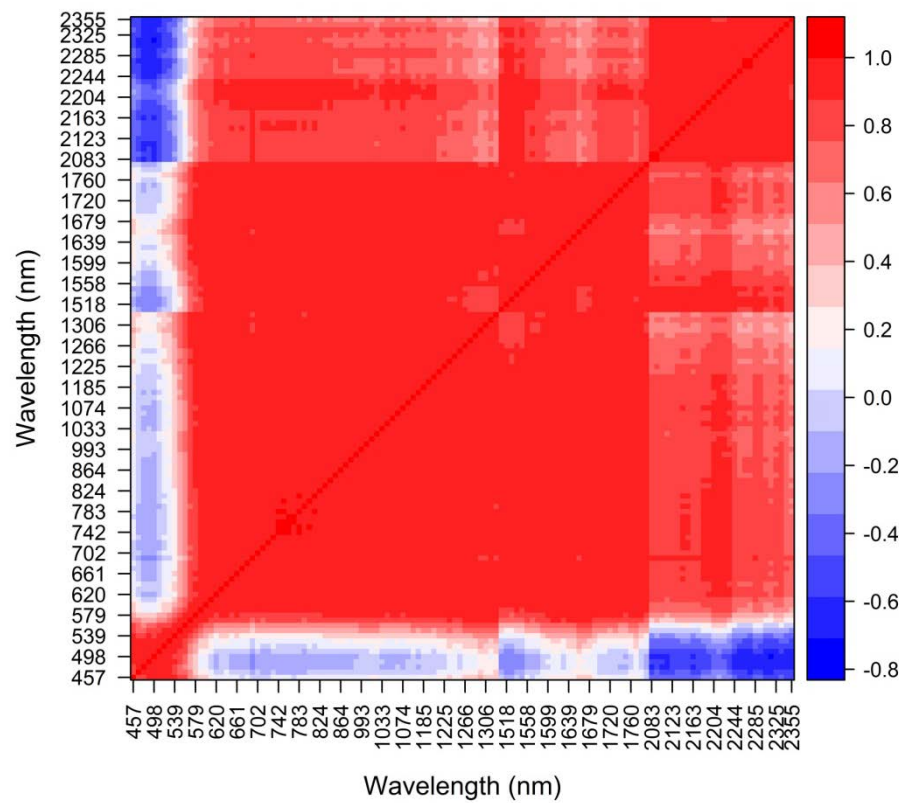
Figure 4.5 Spectra from pixels in the Vaal River 2013 Hyperion hyperspectral image corresponding with the LFA transects and used in the PLSR modelling. The spectra are composed from the 128 bands selected from the original 220 by removing bands with stripes or smearing.

A correlogram of the correlation between spectra from pixels associated with LFA transects (Figure 4.6a) supports the groupings described previously and illustrated in Figure 4.5. All pixel spectra are highly correlated with the minimum correlation coefficient between the 31 pixels from six transects being 0.90 (Figure 4.6a). However, within this high correlation, it can be seen that transects 13 and 14 form a definite group which is distinctly different to transects 1, 2 and 3 and less different to transect 15 (Figure 4.6a). It can also be seen that transects 1, 2 and 3 form a distinct group, as do the pixels in transect 15, and that these two groups are different from each other. Some pixels in the transect 1, 2 and 3 grouping, in particular pixel vr3p2 and less so pixels vr1p1, vr1p2 and vr1p3 show a stronger correlation with some pixels in the transect 13 and 14 group, and also with some pixels in the transect 15 group than does the rest of the pixels associated with the transect 1, 2 and 3 group.

Using the same spectra but comparing correlations between bands (wavelengths) from the 128-band data cube shows that the majority of the bands between 579 nm and 1790 nm are strongly correlated (Figure 4.6b). The far SWIR region between 2083 nm and 2355 nm is also strongly correlated, but slightly less correlated as a group with the central group first described (Figure 4.6b). The visible bands between 457 nm and approximately 579 nm also



a)



b)

Figure 4.6

(a) Correlogram showing correlation between 128-band unsmoothed Hyperion spectra from pixels associated with the LFA transects. (b) Correlogram showing correlation between wavelengths (bands) in the 128-band unsmoothed Hyperion data cube.

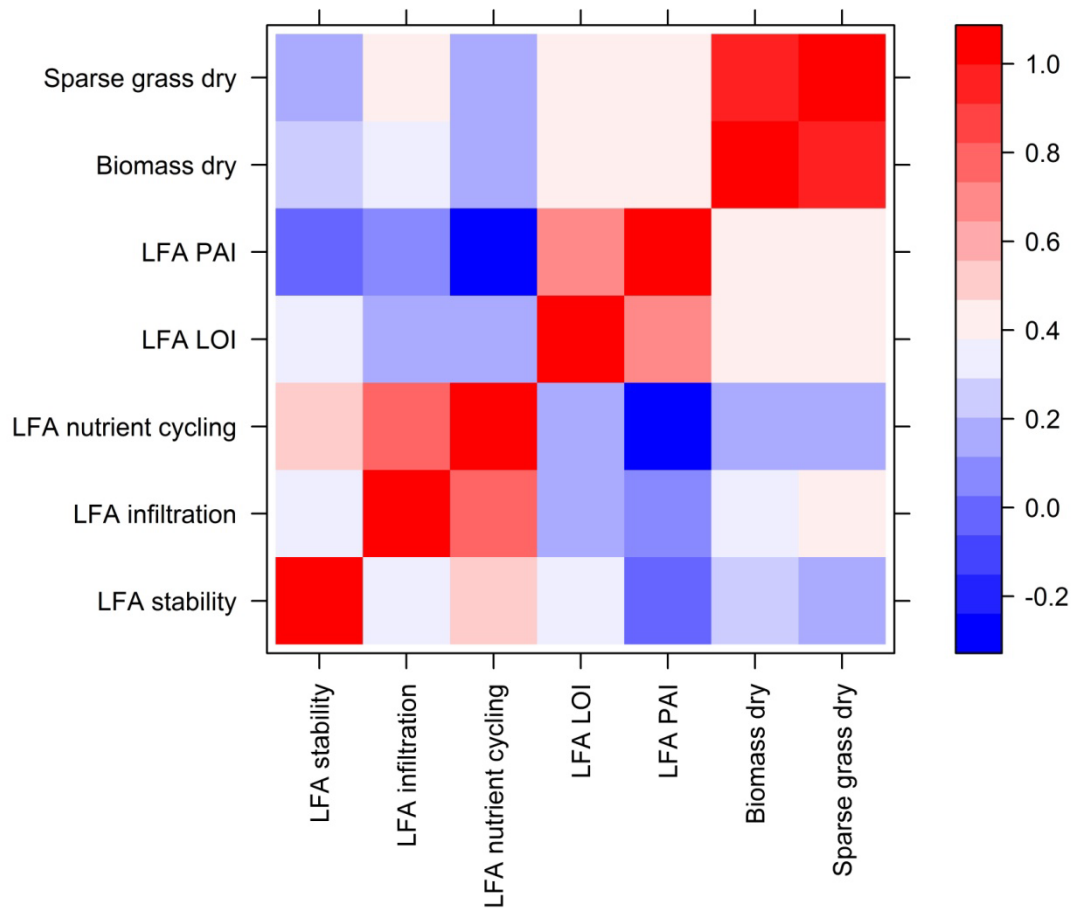


Figure 4.7 Correlation between the response variables used in the PLSR modelling.

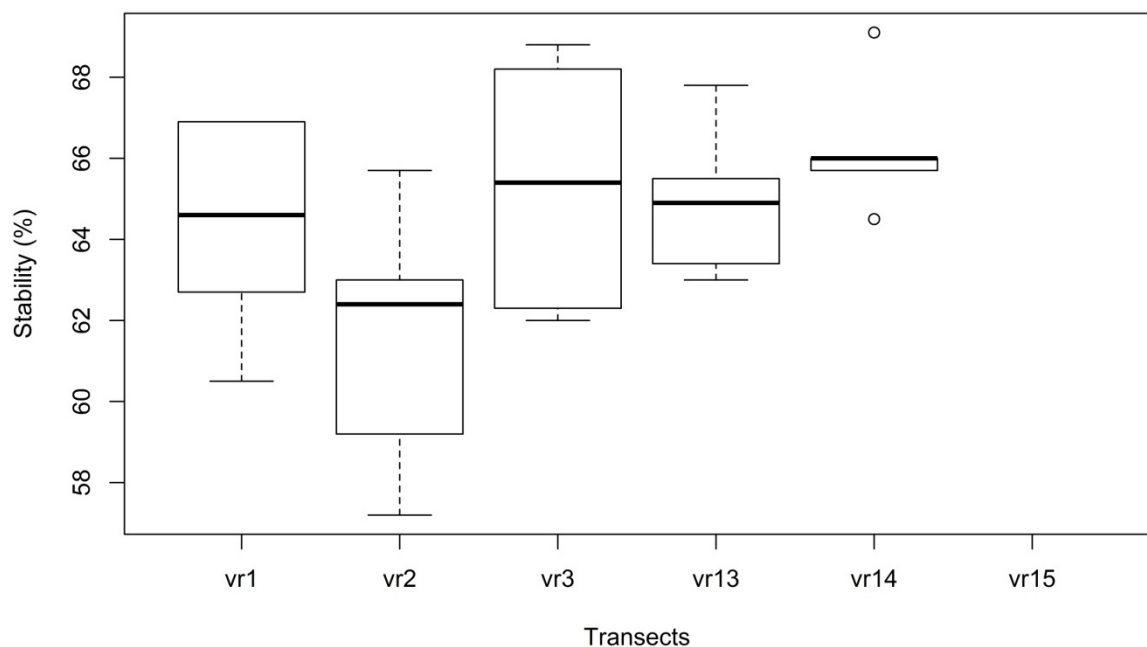
form a distinct group with high within-group correlation (Figure 4.6b). However, this group is very strongly, negatively correlated with the SWIR region from 2083 nm to 2355 nm, and weakly negatively correlated to not correlated with the large central block first described (Figure 4.6b).

Seven response variables, namely dry sparse grass biomass (kg/m^2), total dry biomass (kg/m^2), patch area index (PAI), landscape organization index (LOI), LFA nutrient cycling index, LFA infiltration index, and the LFA soil stability index calculated from data collected in 2011 and normalized to a per pixel basis, were examined for correlation between them (Figure 4.7). The results show that both biomass measures were highly correlated ($r = 0.96$, $P < 0.0001$), which was expected as the sparse grass values formed the matrix of these grasslands, and hence, makes up a large component of the overall biomass on the LFA transects. These biomass values were weakly correlated with the PAI ($r = 0.41$, $P = 0.04$ and $r = 0.38$, $P = 0.06$), the LOI ($r = 0.41$, $P = 0.04$ and $r = 0.43$, $P = 0.03$, respectively) and the LFA

4181 nutrient cycling index ($r = 0.40$, $P = 0.04$ and $r = 0.36$, $P = 0.08$, respectively). The PAI and LOI
 4182 had a fairly strong correlation ($r = 0.65$, $P < 0.0001$) with each other but had almost no or
 4183 even slightly weak correlation with the three LFA indices of soil stability, nutrient cycling and
 4184 infiltration. The LFA nutrient cycling index was strongly correlated with the LFA infiltration
 4185 index ($r = 0.75$, $P < 0.0001$) and less strongly with the LFA soil stability index ($r = 0.52$, $P =$
 4186 0.007) but the LFA infiltration and LFA soil stability index were weakly correlated ($r = 0.32$, P
 4187 $= 0.12$).

4188 4.3.2 Predicting the LFA Soil Stability Index

4189 The LFA soil stability index for pixels from Vaal River ranges between 57.2 and 69.1% with a
 4190 mean of $64.5 \pm 2.9\%$ (Figure 4.8) with lower values suggesting a site is more susceptible to
 4191 soil erosion processes. Generally transects vr1, vr3, vr13 and vr14 had very similar means
 4192 although vr14 had three pixels that had almost identical values and two outliers. Transect
 4193 vr2 had a much lower distribution of LFA stability values for its pixels than did the other
 4194 transects and vr15 was burnt before any LFA soil surface indices could be evaluated (Figure
 4195 4.8).



4196

4197 **Figure 4.8** The distribution of LFA soil stability index values for each transect at Vaal
 4198 River. Transect vr15 has no values for this index as herdsman burnt the
 4199 grassland half way through the LFA sampling.

4200 The PLSR models with 105 bands did not produce strong prediction statistics when
 4201 predicting the LFA stability index (Table 4.2). With unsmoothed data, the strongest model
 4202 was a 2-component model, which with validation data (9 pixels), was not statistically
 4203 significant ($r^2 = 0.25$, $P = 0.17$). Smoothing improved the model predictions such that the
 4204 fourth iteration produced a calibration $r^2 = 0.95$ and a validation $r^2 = 0.41$ ($P = 0.065$) but
 4205 also required more components to achieve the best model. Considering bands selected
 4206 across the whole spectral range did not produce moderately strong models with either
 4207 unsmoothed or smoothed spectra, the spectral range was broken into contiguous blocks
 4208 delineated by sections of bands deleted from the data cube as they contained atmospheric
 4209 noise. These blocks were given names based on their position in the spectral range such as
 4210 visible and near infrared spectrum between 478 and 885 nm (VNIR), near infrared (NIR) and
 4211 short-wave infrared (SWIR). The NIR and SWIR regions were further separated into two
 4212 blocks each and named the near NIR (1003 – 1084 nm) and far NIR (1175 – 1306 nm) and
 4213 near SWIR (1508 – 1790 nm) and far SWIR (2083 – 2325 nm) based on their spectral
 4214 distance from the visible portion of the Hyperion spectrum (Table 4.2 and Figure 4.5).

4215 Two blocks produced relatively strong PLSR models when predicting the LFA stability index.
 4216 These were the near SWIR and the VNIR. However, the near SWIR bands produced
 4217 prediction maps that showed serious striping (not shown) which is caused by the ubiquitous
 4218 striping in the individual bands. This striping was common throughout the Vaal River
 4219 Hyperion images but got progressively worse towards the SWIR region with the visible
 4220 portion of the spectrum the least striped. Of the different spectral blocks, the 40
 4221 unsmoothed bands in the VNIR produced the best prediction statistics using a 5-component
 4222 model with both calibration (16 pixels, $r^2 = 0.93$, $P < 0.0001$) and validation data (9 pixels, r^2
 4223 $= 0.72$, $P = 0.004$; Table 4.2 and Figure 4.9a). However, despite the strong prediction
 4224 statistics, this 5-component model produced a highly speckled prediction map (Figure 4.10)
 4225 when applied to the entire unsmoothed data cube. Furthermore, there is a blue line
 4226 apparent in the upper right section of the prediction map (Figure 4.10) that corresponds
 4227 with a line from a dead or poorly calibrated detector in the 620 nm band which was
 4228 removed from the data cube prior to modelling and did not show in any of the adjacent
 4229 bands when visually inspected under any stretched display tested. These pixels are also not
 4230 involved in either the calibration or validation modelling but only when generating the

Table 4.2 Exploratory models predicting the LFA stability index from the Vaal River Hyperion hyperspectral data cube. The terms “far” and “near” are used as relative terms describing portions of the Hyperion spectrum relative to the visible portion of the spectrum. Raw refers to the spectrum as they are post-FLAASH atmospheric correction and the terms sm1 – sm4 pertain to the number of iterations of the smoothing algorithm applied to the spectra.

Bands in Model	wavelengths	Calibration (n = 16 pixels)			Validation (n = 9 pixels)		
		RMSEP (%)	Spectral r ² (%)	Stability r ² (%)	comps	Stability r ² (%)	P
105 bands raw	478 - 610, 630 - 885, 1003 - 1084, 1185 -1306, 1518 - 1578, 1629 - 1780, 2093 - 2123, 2143 - 2163, 2194 - 2295, 2325 -2335	3.02	97.6	54.8	2	24.8	0.173
105 bands sm1		3.73	98.7	72.0	3	28.0	0.143
105 bands sm2		3.75	98.9	68.2	7	36.9	0.083
105 bands sm3		3.58	99.8	96.3	7	40.1	0.067
105 bands sm4		3.52	99.8	95.1	7	40.6	0.065
Far SWIR raw	2083 - 2123, 2143 - 2174,2194 - 2325	4.78	99.1	55.2	3	13.2	0.336
Far SWIR sm1		10.01	100.0	89.9	7	28.1	0.142
Far SWIR sm2		9.70	100.0	62.7	8	15.7	0.291
Far SWIR sm3		10.47	100.0	65.6	9	15.5	0.295
Far SWIR sm4		9.39	100.0	68.4	9	12.3	0.356
Near SWIR raw	1508 - 1578, 1629 - 1780	2.95	96.8	85.0	4	45.9	0.045
Near SWIR sm1		3.12	98.8	78.7	4	43.2	0.054
Near SWIR sm2		3.18	99.3	75.1	4	50.2	0.033
Near SWIR sm3		3.17	99.5	73.2	4	55.3	0.022
Near SWIR sm4		3.16	99.6	72.4	4	58.4	0.017
Far NIR raw	1175 - 1306	3.09	94.8	23.2	1	15.0	0.303
Far NIR sm1		3.09	97.6	23.3	1	14.9	0.304
Far NIR sm2		3.09	98.2	23.2	1	14.9	0.305
Far NIR sm3		3.98	99.9	40.4	4	22.4	0.198
Far NIR sm4		3.89	99.9	39.5	4	22.8	0.194
Near NIR raw	1003 - 1084	3.15	95.9	19.2	1	6.2	0.519
Near NIR sm1		3.15	98.3	19.1	1	6.1	0.522
Near NIR sm2		3.15	98.9	19.1	1	6.1	0.521
Near NIR sm3		3.15	99.2	19.1	1	6.2	0.519
Near NIR sm4		3.15	99.4	19.1	1	6.3	0.515
VNIR raw	478 - 610, 630 - 885	6.38	99.4	93.2	5	72.2	0.004
VNIR sm1		7.64	99.8	75.7	5	68.9	0.006
VNIR sm2		8.07	99.9	87.4	7	61.1	0.013
VNIR sm3		8.59	99.9	81.2	7	51.8	0.029
VNIR sm4		9.56	100.0	91.7	9	38.6	0.074

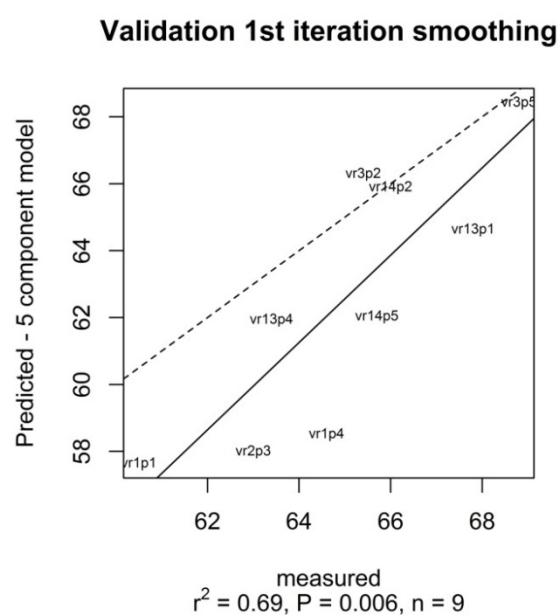
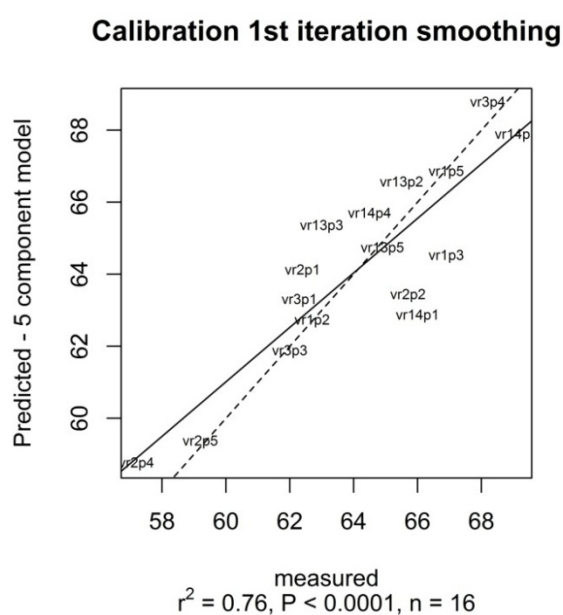
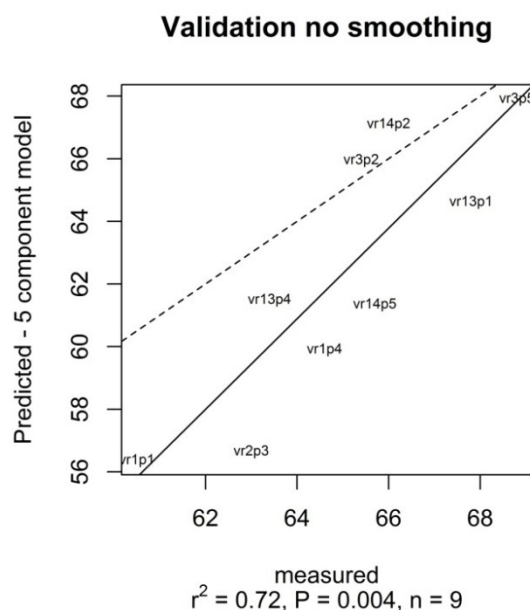
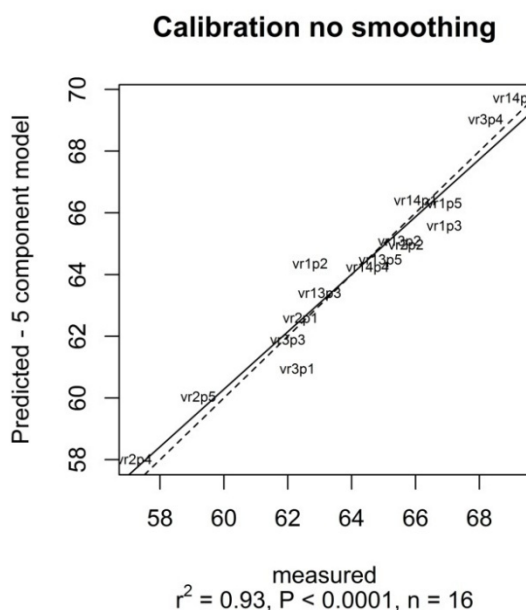
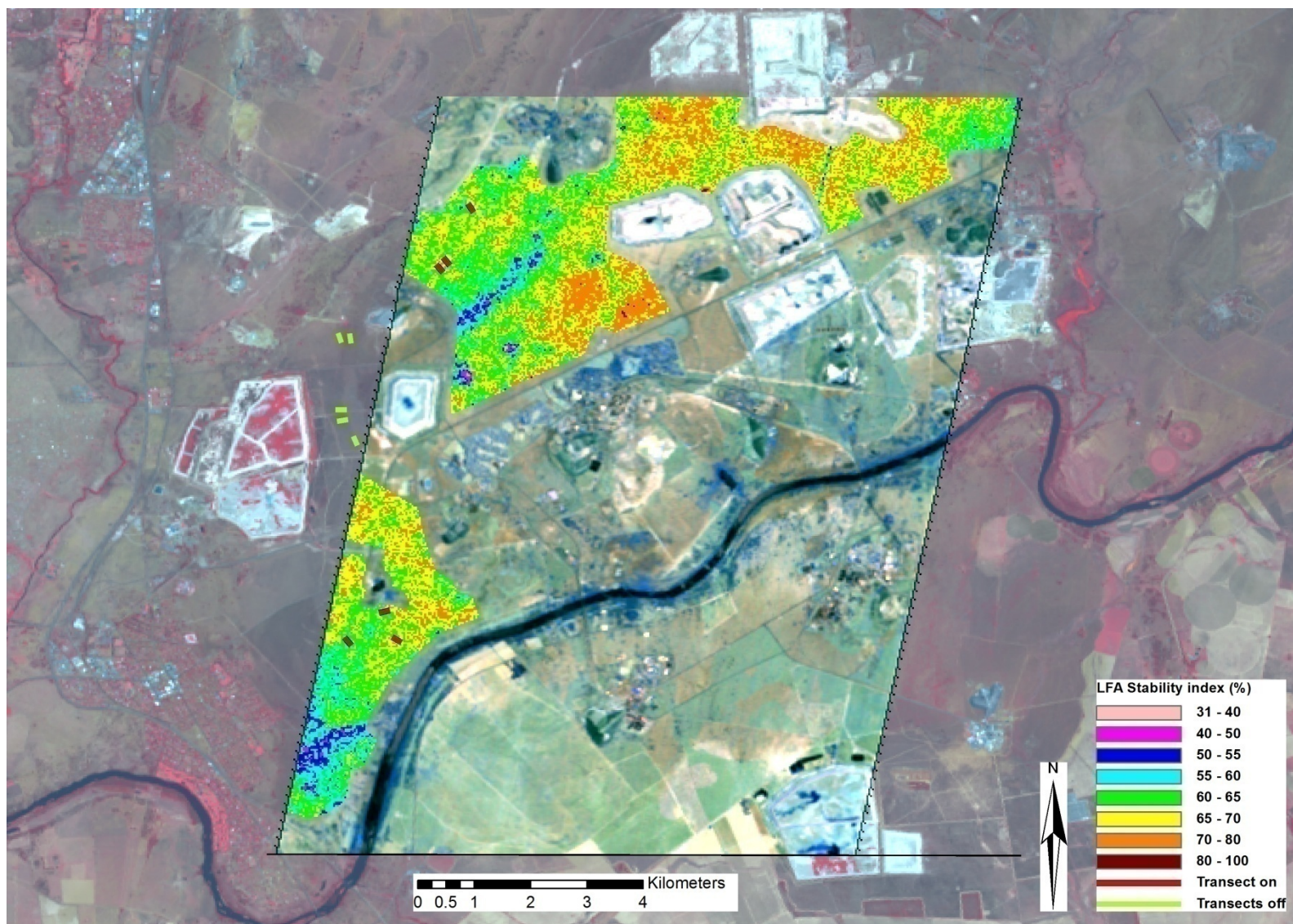


Figure 4.9

Measured LFA stability index values versus predicted stability values using a 5-component PLSR model using 40 VNIR bands. (a) Hyperion spectra with no smoothing and (b) one iteration of the smoothing algorithm applied to all spectra in the data cube prior to PLSR modelling. Solid lines represent the best fit and dotted lines represent the equality line.



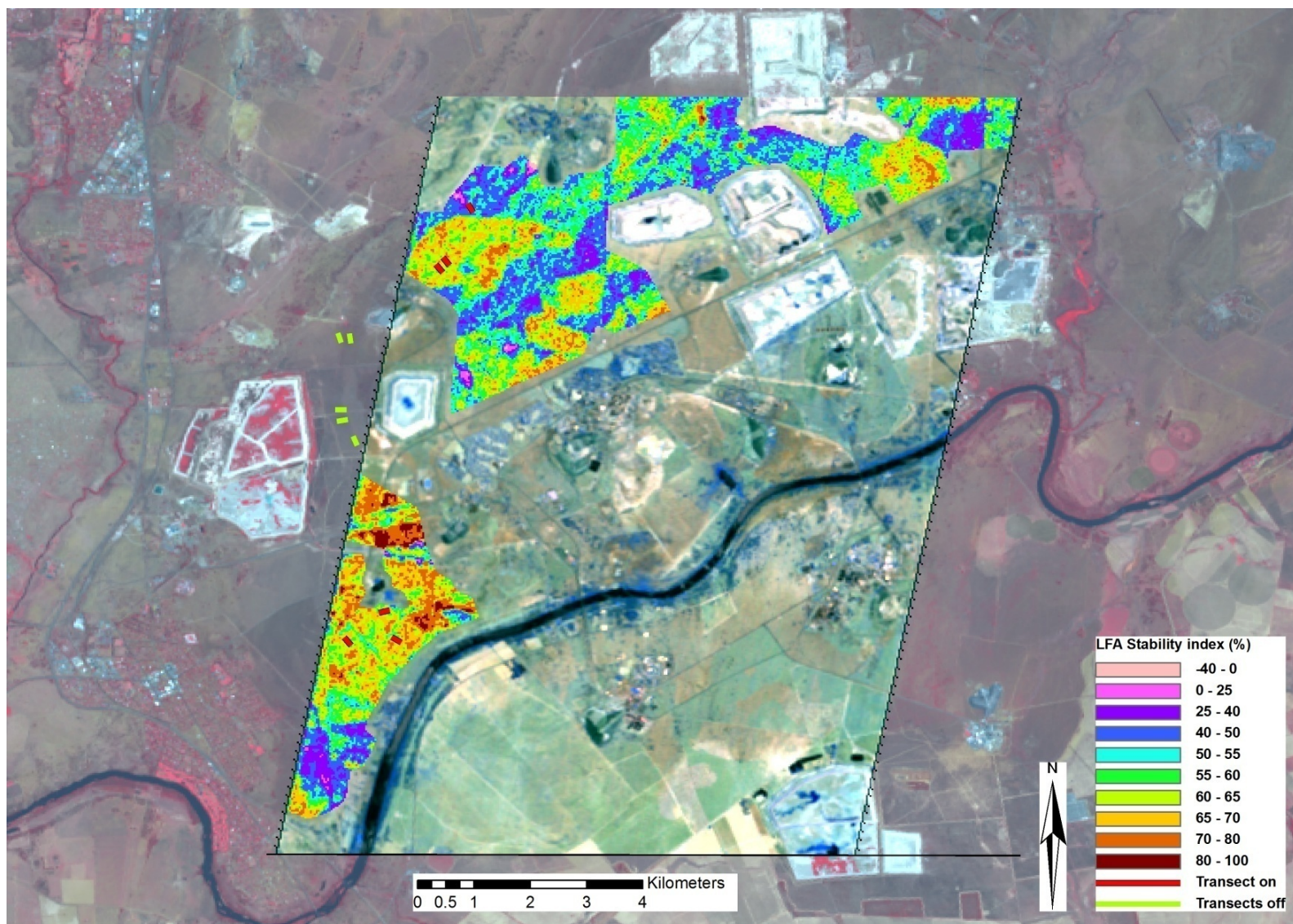
4245

4246 **Figure 4.10** LFA stability index predicted by a 5-component PLSR model using 40 unsmoothed bands from the VNIR Hyperion spectrum.

4247 prediction map. This suggests that the PLSR modelling is amplifying a very subtle influence
4248 of this 620 nm band on the adjacent bands and is also apparent in some other bands where
4249 a particular fault with a specific detector showed as a strong line of pixels parallel with the
4250 flight path.

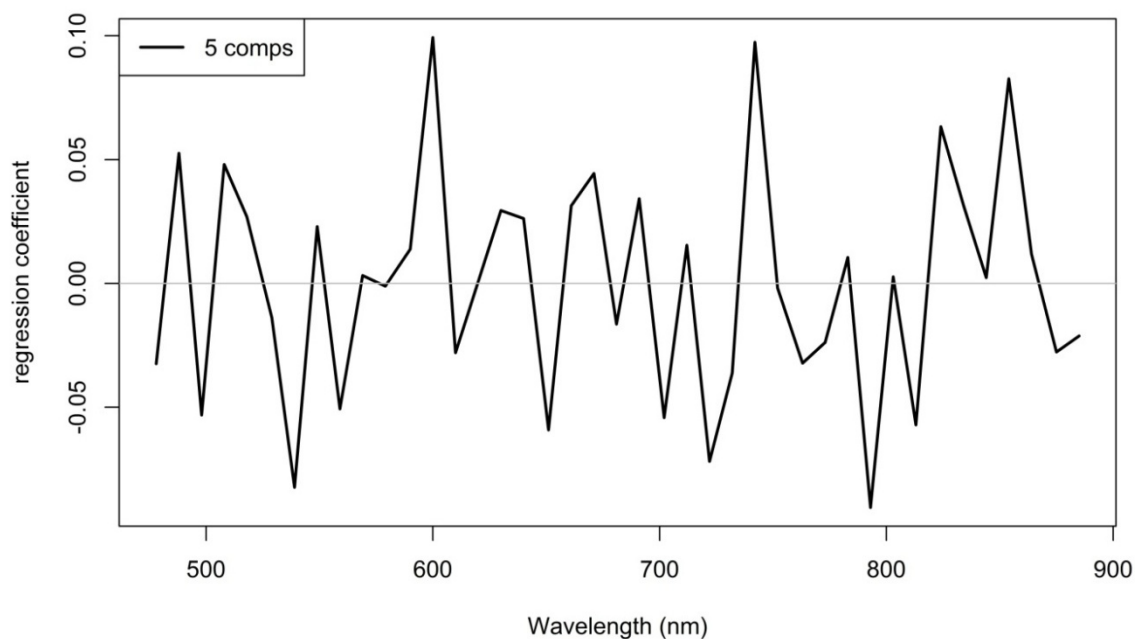
4251 Smoothing the spectra of each pixel did reduce the speckling effect produced in the
4252 prediction map for the LFA soil stability index and also improved the delineation of features,
4253 such as the burn scar forming an arc around transects vr13 and vr14 in the western portion
4254 of the northern block of the prediction map (Figure 4.11). However, each iteration of the
4255 smoothing algorithm also resulted in more components being added to the model to
4256 achieve the best prediction statistics, and a reduction in the strength of these prediction
4257 statistics with validation data (Table 4.2). Of the smoothed spectral data cubes using the 40
4258 bands from the VNIR, the first iteration gave the strongest prediction statistics when
4259 predicting the LFA soil stability index (validation: 9 pixels, $r^2 = 0.69$, $P = 0.006$; Table 4.2 and
4260 Figure 4.9b). This first iteration required the same number of components (five) to achieve
4261 the best prediction statistics as the unsmoothed data required. However, the fourth
4262 iteration smoothing required 9 components and the coefficient of determination with
4263 validation data had fallen to 0.39 ($P = 0.07$; Table 4.2).

4264 The speckling in the unsmoothed prediction map predicting the LFA soil stability index
4265 points to one form of noise in the Hyperion imagery. This is in the spectra themselves, i.e.
4266 the individual band measurements for each pixel, or the z-axis in the data cube which
4267 corresponds to each pixel. An examination of the regression coefficients (Figure 4.12a) and
4268 loadings (Figure 4.13a) for the model derived from unsmoothed spectra shows a very jagged
4269 appearance where the spectrum alternates between local maximum and minimum values in
4270 most successive bands. Figure 4.12b and Figure 4.13b shows the regression coefficients and
4271 loadings for the model derived from the data cube subjected to the first iteration of the
4272 smoothing algorithm and it can clearly be seen that some of this noise has been smoothed
4273 out. An examination of the individual loadings (Figure 4.13a and b) shows that a single
4274 application of the smoothing algorithm has the most effect on the first component and this
4275 effect reduces with each additional component added to the PLSR model. Increasing the
4276 smoothing iterations increases the effect through the components but also increasingly
4277 removes the smaller spectral features in the spectrum.

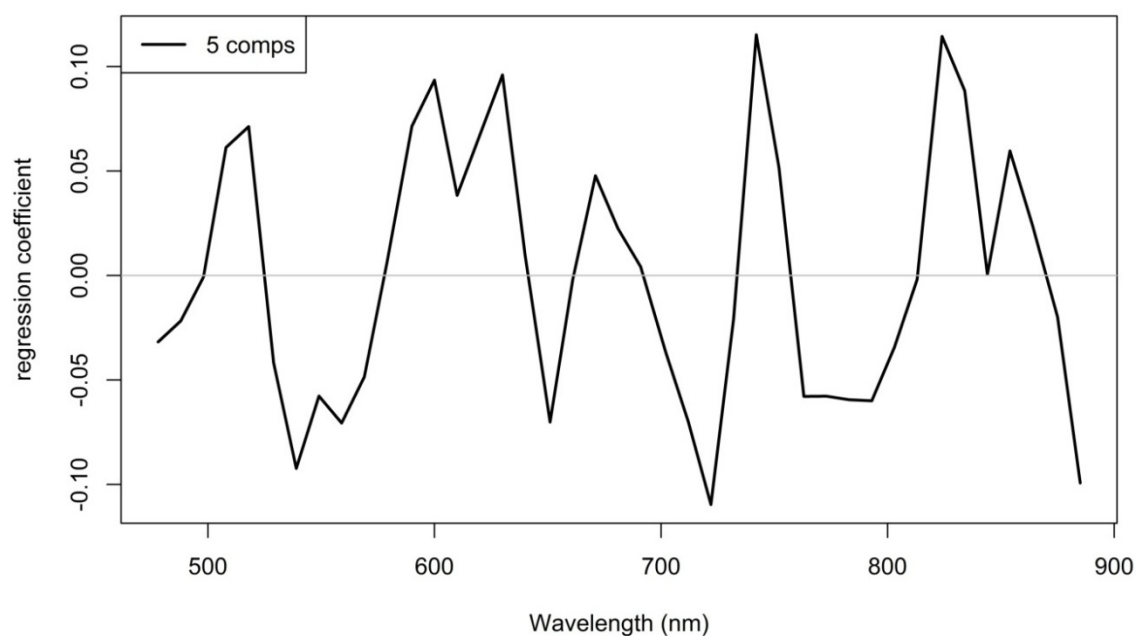


4278

4279 **Figure 4.11** LFA stability index predicted by a 5-component PLSR model using 40 VNIR bands after one iteration of the smoothing algorithm.

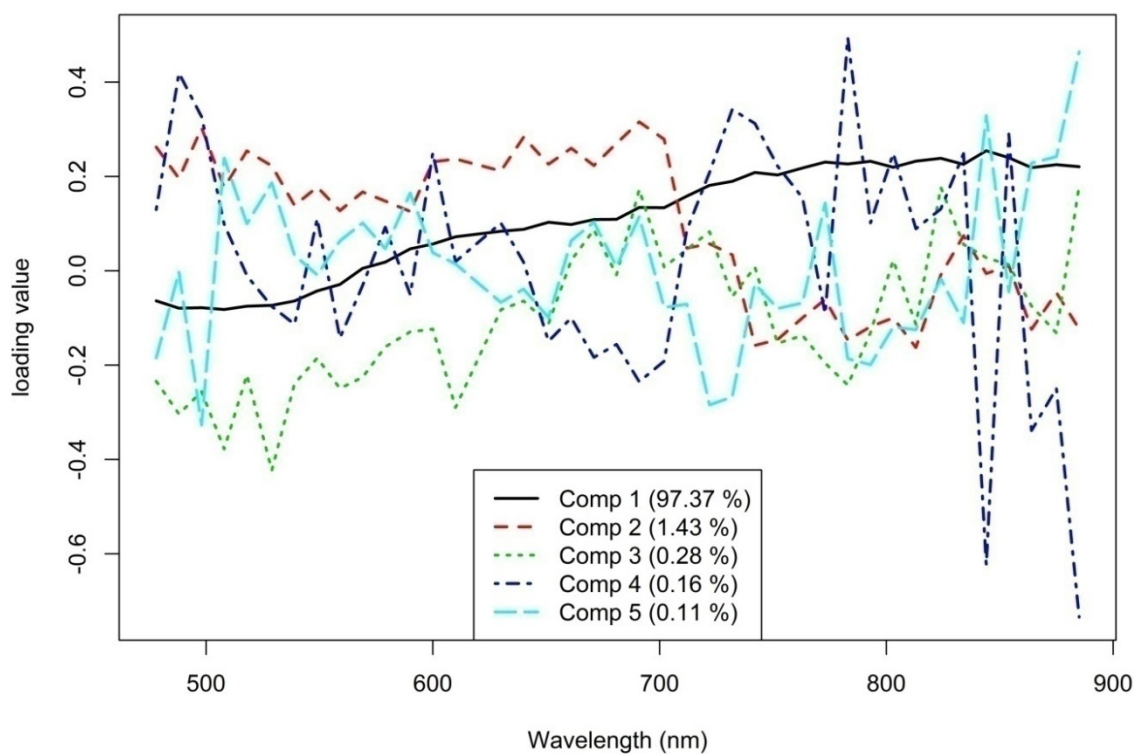


4280 a)

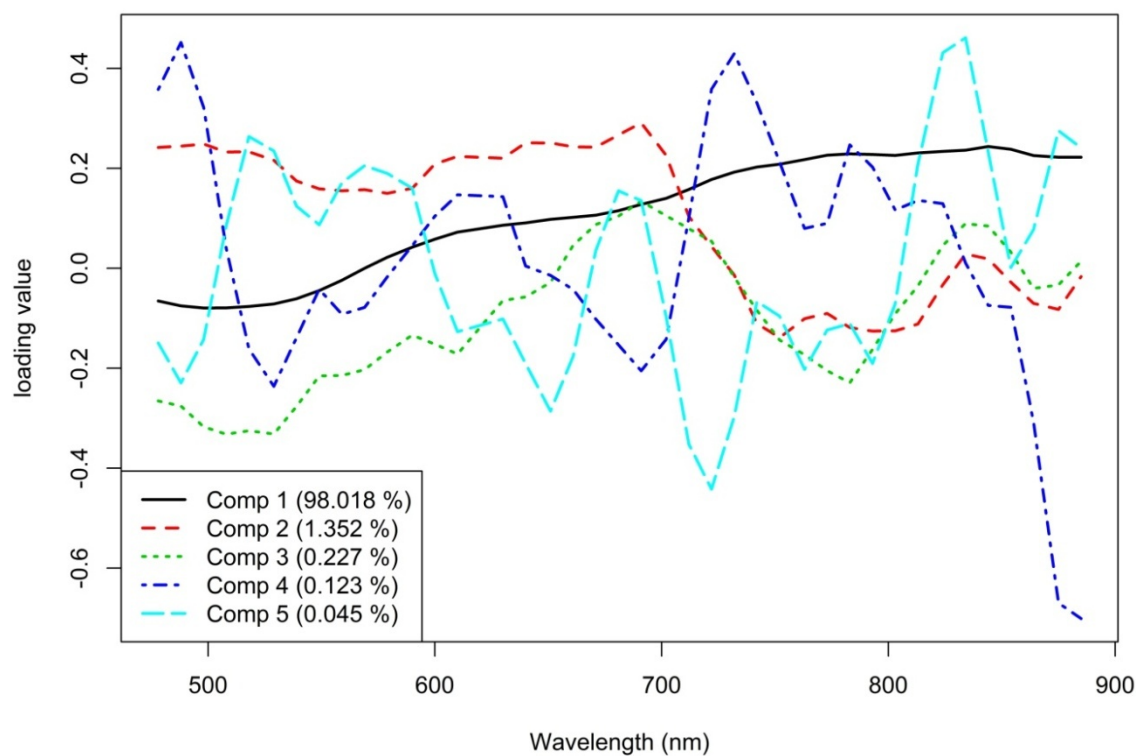


4281 b)

4282 **Figure 4.12** Regression coefficients for (a) a five-component PLSR model with no spectral
 4283 smoothing, and (b) with one iteration of the smoothing algorithm prior to
 4284 PLSR modelling.



4285 a)



4286 b)

4287 **Figure 4.13** Loading values for the VNIR PLSR models with (a) unsmoothed Hyperion
 4288 spectra, and (b) Hyperion spectra with one iteration of the smoothing
 4289 algorithm applied to the whole data cube prior to PLSR modelling.

As the VNIR and near SWIR blocks produced fairly strong model predictions for the LFA stability index (Table 4.2), a combination of the two blocks was tested but these models were dominated by the near SWIR and the prediction maps showed strong striping (results not shown). Therefore, band combinations within the VNIR were tested to see if a better model could be derived using a subset of the 40 bands from the VNIR block. None of these models produced significantly better results than the unsmoothed 40 band VNIR model already explored, but they do show the difficulty in selecting the best model. Using the raw unsmoothed data for 31 VNIR bands and a 6-component model using 31 bands produced slightly stronger coefficients of determination ($r^2 = 0.76$, $P = 0.002$; Table 4.3) with validation data (9 pixels) than did the 40 band VNIR model. An examination of the RMSEP shows it generally increasing with the addition of successive components to the model (Table 4.3) thus undermining Mevik and Wehrens (2007) suggestion of selecting the number of components in the model based on the lowest RMSEP. In this example it would be a one-component model (RMSEP = 3.19%) but the calibration ($r^2 = 0.16$) and validation ($r^2 = 0.045$, $P = 0.58$) statistics are very weak (Table 4.3). An examination of the percentage change in RMSEP with the addition of each component (Table 4.3) shows that the addition of none of the components in the possible model results in a 2% decrease in the RMSEP (Cho *et al.*, 2007, Kooistra *et al.*, 2004). Thus, neither of these two methods of deciding the number of components in the model is feasible with the Hyperion spectra and this trend is common to all the Hyperion models presented.

Therefore, the method used to select the best models in this study was the performance of the model with the validation data. The performance of the model with the calibration data is also taken into account but as this essentially “knows” the prediction values as these are part of the compression process (Mevik and Wehrens, 2007, Geladi and Kowalski, 1986a, Geladi and Kowalski, 1986b) the performance of the model with the unseen validation data takes priority. The output of the chosen model was further evaluated by comparing the distribution of the predicted LFA stability values for the entire grassland area with the distribution of the field values derived for each pixel from the transects (25 pixels) using a t-test and a Wilcoxon Rank-sum test on the assumption that the two distributions should be the same. These tests suggested that the best model was actually a 5-component model (validation: $r^2 = 0.68$, $P = 0.006$; Table 4.3) as this had the least difference between the

4321 **Table 4.3** PLSR model results using Hyperion spectra to predict the LFA stability index using 31 bands from the VNIR region of the spectra
4322 (488 – 610, 630 – 722, 773 – 834, 854 nm).
4323

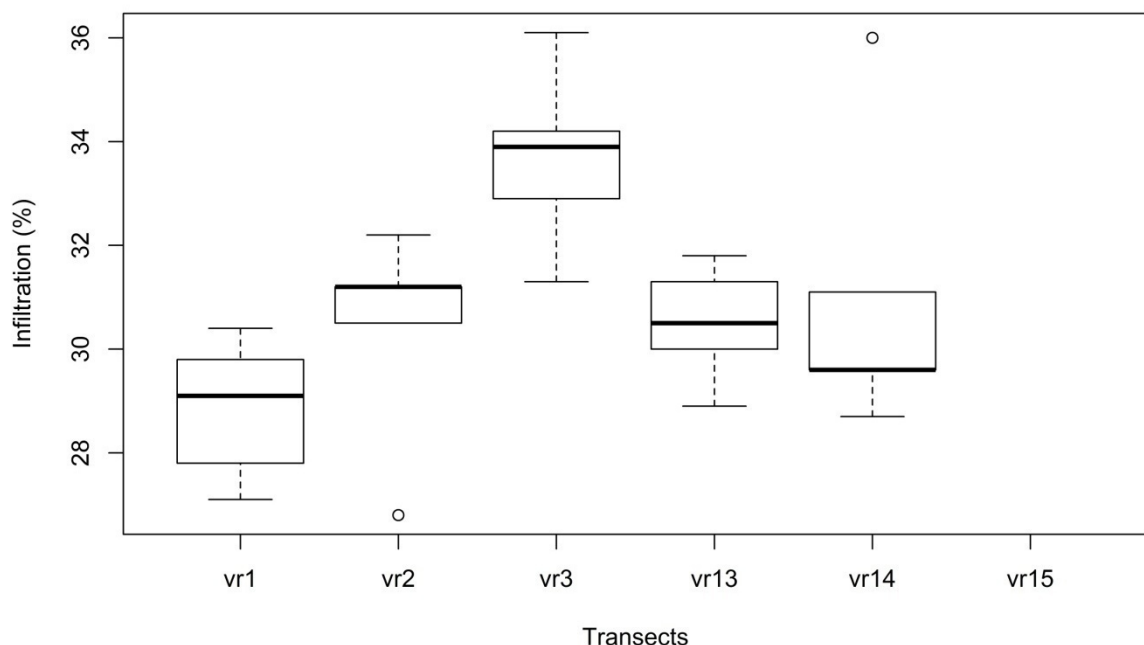
Component	PLSR calibration (n = 16) % RMSEP change		PLSR Calibration Spectral Stability r2 r2		PLSR Validation (n = 9) Stability r2 Prob.		LFA stability index measured values Mean SD		Predicted Hyperion values Mean SD		Measured values used in model versus predicted values for whole grassland T.test DF Prob		
1	3.193		96.73	16.09	4.5	0.584	64.5	2.94	66.5	2.28	-3.54	24.025	0.0016
2	3.384	-6.0	98.78	29.64	17.4	0.264			68.9	7.61	-7.60	24.283	7.23E-08
3	3.663	-8.2	99.17	51.03	7.4	0.479			61.0	7.12	5.91	24.248	4.13E-06
4	5.256	-43.5	99.25	80.56	43.0	0.055			66.3	4.99	-3.15	24.122	0.0043
5	5.165	1.7	99.35	89.75	68.1	0.006	-	-	64.8	6.68	-0.63	24.218	0.5339
6	5.274	-2.1	99.44	94.59	75.9	0.002			61.8	9.41	4.48	24.434	0.00015
7	5.264	0.2	99.52	97.61	75.5	0.002			61.4	10.81	5.16	24.574	2.62E-05
8	5.376	-2.1	99.62	98.78	70.9	0.004			60.7	11.07	6.42	24.602	1.09E-06
9	5.462	-1.6	99.74	99.25	70.1	0.005			61.2	10.62	5.61	24.554	8.27E-06
10	5.36	1.9	99.77	99.78									
11	5.328	0.6	99.83	99.91									
12	5.329	0.0	99.89	99.95									
13	5.346	-0.3	99.92	99.99									
14	5.346	0.0	99.96	100									

4324

two distributions ($T = -0.63$, $P = 0.53$). The prediction map is not shown but has the same problems with speckling as the previous prediction map for unsmoothed spectra and forty bands in the model (Figure 4.10) and likewise smoothing of the 31 band prior to modelling leads to a major drop in the strength of the model predictions with validation data (not shown).

4.3.3 Predicting the LFA Infiltration Index

The LFA infiltration index is an indicator of the ability of water to infiltrate the soil or run off the surface and be lost from the system. The distribution of the LFA infiltration values for pixels from the different transects at Vaal River show quite different patterns between transects in the R.G. Williams nature reserve where transect vr1 is low, mainly below 30% and transect vr3 is high, between 31% and 36%, while transect vr2 lies between the two extreme transects. The two transects in the northern grasslands have very similar infiltration values between 29.5 and 32 %, although one pixel in transect vr14 is an outlier with a value of 36 % that more closely fits the distribution for vr3. Transect vr15 was burnt before any soil surface condition indices could be collected and therefore no values are presented.



4340

4341 **Figure 4.14** Distribution of LFA infiltration index values for pixels from each transect at
4342 Vaal River

4343 The PLSR models predicting the LFA infiltration index performed much like those for the LFA
4344 stability index albeit slightly weaker. The best full spectrum model using 105 bands from
4345 across the spectrum had weak validation statistics ($r^2 = 0.32$, $P = 0.11$) for a 2-component
4346 model with unsmoothed spectra (Table 4.4). Applying the smoothing iteration did not
4347 significantly improve the predictions for the full spectrum models but did require more
4348 components as the number of iterations of the smoothing algorithm were applied (Table
4349 4.4). Like the LFA stability index, both the near SWIR and the VNIR regions produced the
4350 best, albeit weak models with validation data. Of the near-SWIR models, a 3 component
4351 model derived from spectra pre-processed with four iterations of the smoothing algorithm
4352 produced the best validation statistics ($r^2 = 0.51$, $P = 0.03$; Table 4.4). However, the
4353 prediction map derived using this model applied to the entire data cube shows the severe
4354 striping that occurs in most of the SWIR region (Figure 4.15).

4355 The best models for predicting the LFA infiltration index were again found in the VNIR
4356 region where the 40-band 5-component model produced a validation r^2 of 0.43 ($P = 0.055$)
4357 with no spectral smoothing (Table 4.4). Four iterations of the smoothing algorithm prior to
4358 modelling reduced the number of components in the best 40-band model to 3 components
4359 and slightly improved the validation coefficient of determination ($r^2 = 0.46$, $P = 0.044$; Table
4360 4.4). A selection of 35 bands in the VNIR region of the spectrum produced significantly
4361 better results in predicting the LFA infiltration index with an 8-component model using
4362 unsmoothed spectra which gave a validation coefficient of determination of 0.73 ($P = 0.003$;
4363 Table 4.5). However, the last three components in the model only accounted for 8.5% of the
4364 variability in the validation predictions (Table 4.5). Therefore the 5-component model ($r^2 =$
4365 0.65 , $P = 0.009$) was preferred (Table 4.5, Figure 4.16a and Figure 4.17a). This 35-band
4366 model also produced strong prediction statistics when derived from the smoothed Hyperion
4367 hyperspectral data with a 4-component model using the same bands from a Hyperion data
4368 cube subjected to four iterations of the smoothing algorithm producing a coefficient of
4369 determination of 0.66 ($P = 0.008$; Table 4.5, Figures 4.16b and Figure 4.17b). It should be
4370 noted though that the values for predictions from both these 35-band models over the
4371 entire grassland were significantly different ($T = -16.22$, $P < 0.0001$, 8 component
4372 unsmoothed model) from the distribution of field calculated values for each pixel on the

Table 4.4 Exploratory models predicting the LFA infiltration index from the Vaal River Hyperion hyperspectral data cube. The terms “far” and “near” are used as relative terms describing portions of the Hyperion spectrum relative to the visible portion of the spectrum. Raw refers to the spectra as they are, post-FLAASH atmospheric correction, and the terms sm1 – sm4 pertain to the number of iterations of the smoothing algorithm applied to the spectra.

Bands and smoothing	wavelengths (nm)	Calibration (n = 16 pixels)			Validation (n = 9 pixels)		
		RMSEP (%)	Spectral r ² (%)	Infiltration r ² (%)	comps	Infiltration r ² (%)	P
105 raw	478 - 610, 630 - 885, 1003 - 1084, 1185 - 1306, 1518 - 1578, 1629 - 1780, 2093 - 2123, 2143 - 2163, 2194 - 2295, 2325 - 2335	2.437	97.47	61.01	2	31.97	0.113
105 sm1		2.55	98.36	51.04	2	30.34	0.124
105 sm2		2.377	99.26	79.31	4	33.69	0.101
105 sm3		2.265	99.44	75.58	4	32.88	0.107
105 sm4		2.248	99.55	73.60	4	29.29	0.132
far SWIR raw	2093 - 2113, 2133 - 2264, 2295, 2325 – 2335	2.324	99.11	51.76	2	14.87	0.305
far SWIR sm1		2.38	99.56	45.20	2	6.91	0.494
far SWIR sm2		4.693	99.98	92.42	9	7.07	0.489
far SWIR sm3		5.514	99.99	91.12	9	9.78	0.413
far SWIR sm4		5.716	99.99	86.74	9	6.98	0.492
near SWIR raw	1508 - 1578, 1629 - 1790	2.342	95.16	63.85	2	29.21	0.133
near SWIR sm1		2.3	97.80	49.76	2	39.40	0.070
near SWIR sm2		2.33	98.54	42.88	2	41.58	0.061
near SWIR sm3		2.389	98.90	37.79	2	42.25	0.058
near SWIR sm4		2.48	99.45	50.19	3	50.59	0.032
Far NIR raw	1175 – 1306	2.974	72.71	14.17	1	7.90	0.464
Far NIR sm1		4.752	99.86	77.78	7	31.34	0.117
Far NIR sm2		4.157	100.00	69.15	8	31.52	0.116
Far NIR sm3		3.406	99.92	52.21	5	16.63	0.276
Far NIR sm4		4.204	100.00	55.21	7	13.43	0.332
Near NIR raw	1003 – 1084	2.544	95.39	1.68	1	4.04	0.604
Near NIR sm1		5.501	100.00	71.69	9	6.50	0.508
Near NIR sm2		4.358	100.00	65.16	9	11.81	0.365
Near NIR sm3		3.356	100.00	55.18	5	7.91	0.463
Near NIR sm4		3.214	100.00	31.24	5	13.23	0.336
VNIR raw	478 - 610, 630 – 885	2.78	99.31	95.07	5	43.13	0.055
VNIR sm1		2.815	99.51	64.24	3	38.03	0.077
VNIR sm2		2.866	99.60	63.24	3	39.67	0.069
VNIR sm3		2.911	99.64	63.77	3	43.32	0.054
VNIR sm4		2.955	99.67	64.72	3	46.21	0.044

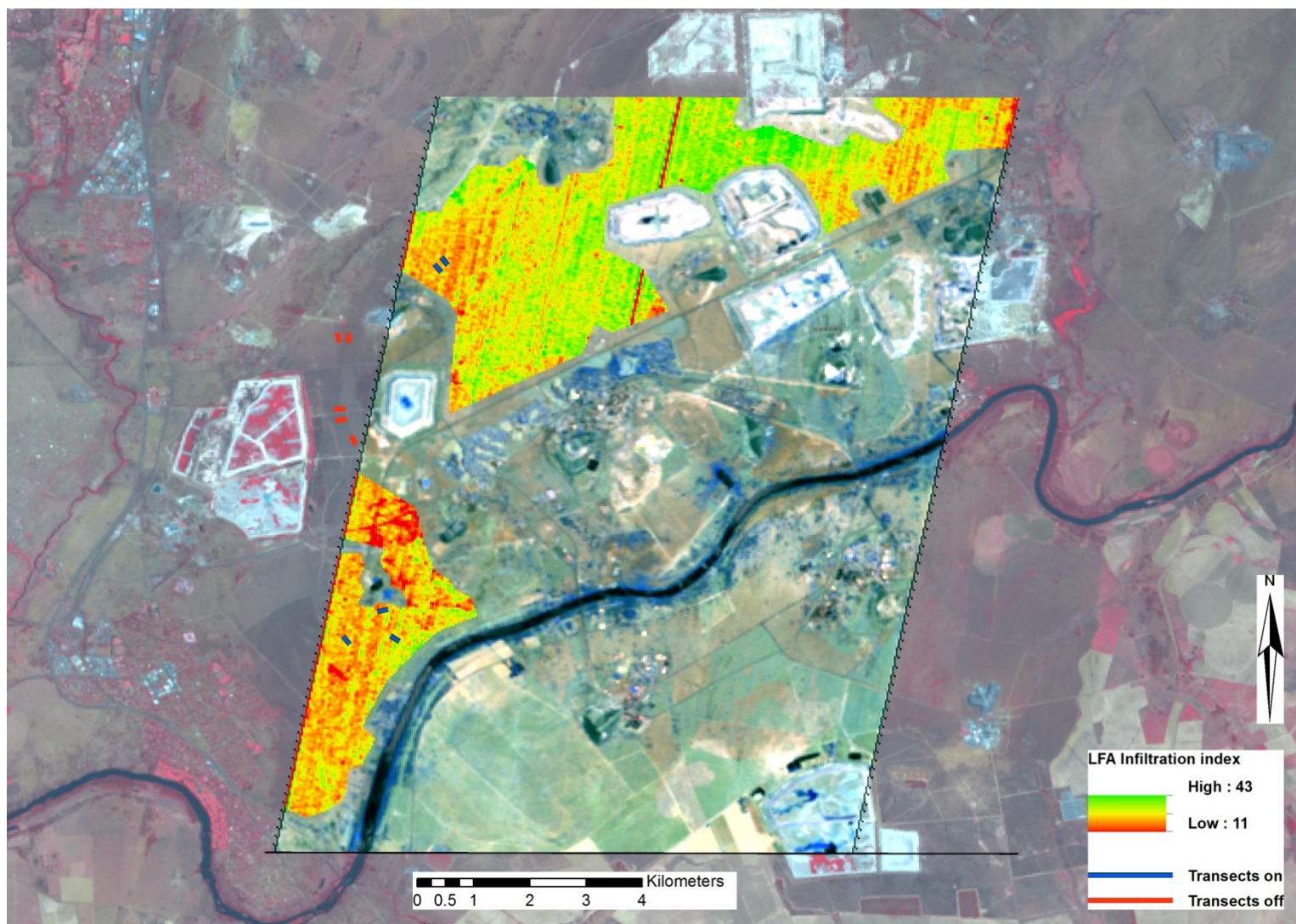


Figure 4.15 LFA infiltration index predicted by a 3-component PLSR model using the near SWIR block and a 4th iteration smoothing of the Hyperion data cube prior to PLSR modelling.

4383 **Table 4.5** PLSR Model predicting the LFA infiltration index using 35 bands (508 - 610, 630 – 864 nm) from the VNIR portion of the
 4384 Hyperion spectrum using unsmoothed Hyperion spectra, and the fourth component from a model using the same bands but
 4385 subjected to four iterations of the smoothing algorithm prior to modelling.
 4386

Component	PLSR calibration (n = 16)		PLSR Calibration		PLSR Validation (n = 9)		LFA infiltration index measured values		Predicted Hyperion values		Measured values from field versus predicted values for whole grassland		
	RMSEP (%)	% change	Spectral r ²	Infiltration r ²	Infiltration r ²	P	Mean	SD	Mean	SD	T.test	DF	P
Unsmoothed													
1	2.82		96.26	1.27	2.17	0.705	30.88	2.41	30.74	0.37	0.30	24.00	0.77
2	2.58	8.34	99.00	27.60	16.37	0.280			34.91	7.35	-8.35	24.39	< 0.0001
3	2.77	-7.24	99.17	70.97	30.36	0.124			40.13	10.62	-19.07	24.83	< 0.0001
4	2.73	1.30	99.35	82.69	53.38	0.025			37.44	6.89	-13.59	24.35	< 0.0001
5	<u>2.75</u>	<u>-0.51</u>	<u>99.43</u>	<u>92.61</u>	<u>64.82</u>	<u>0.009</u>	-	-	<u>37.82</u>	<u>7.44</u>	<u>-14.36</u>	<u>24.40</u>	<u>< 0.0001</u>
6	2.80	-1.86	99.51	97.12	68.05	0.006			38.32	7.79	-15.40	24.44	< 0.0001
7	2.76	1.46	99.64	98.51	70.64	0.005			38.31	7.38	-15.39	24.40	< 0.0001
8	<u>2.79</u>	<u>-0.98</u>	<u>99.73</u>	<u>99.16</u>	<u>73.37</u>	<u>0.003</u>	-	-	<u>38.72</u>	<u>7.62</u>	<u>-16.22</u>	<u>24.42</u>	<u>< 0.0001</u>
9	2.80	-0.50	99.81	99.53	71.81	0.004			37.84	6.69	-14.42	24.33	< 0.0001
4 th iteration													
4	2.90	-0.07	99.89	62.05	66.07	0.008			40.39	9.98	-19.62	24.73	< 0.0001

4387

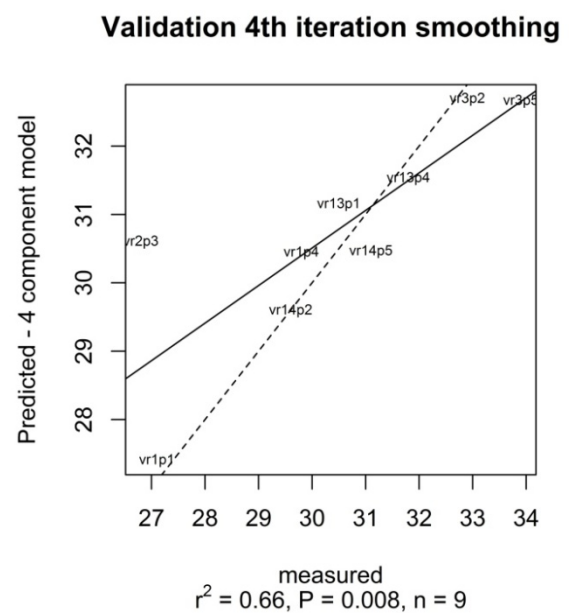
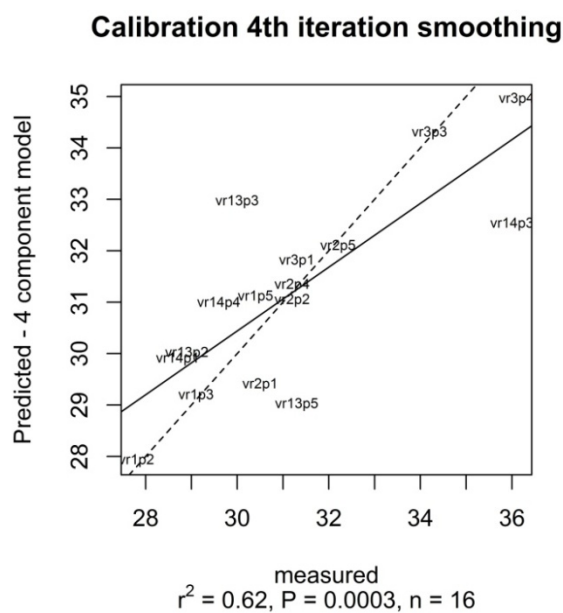
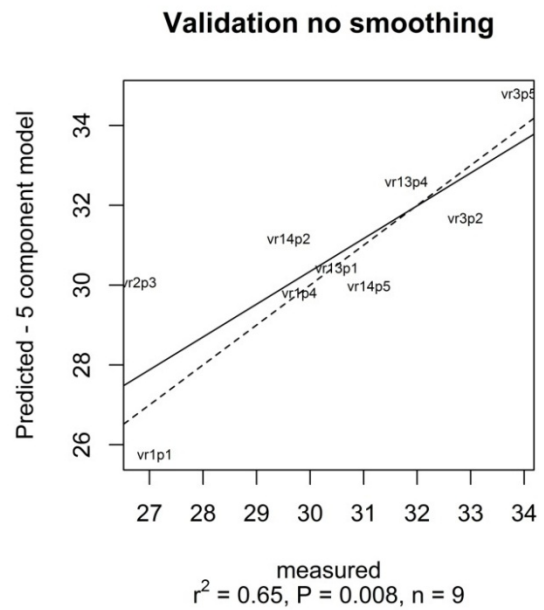
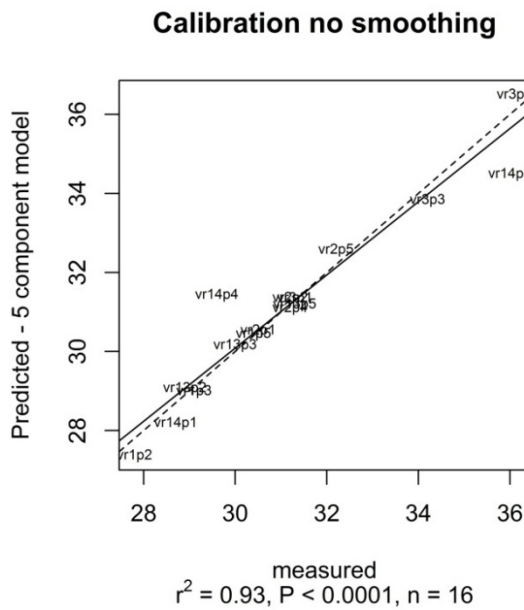
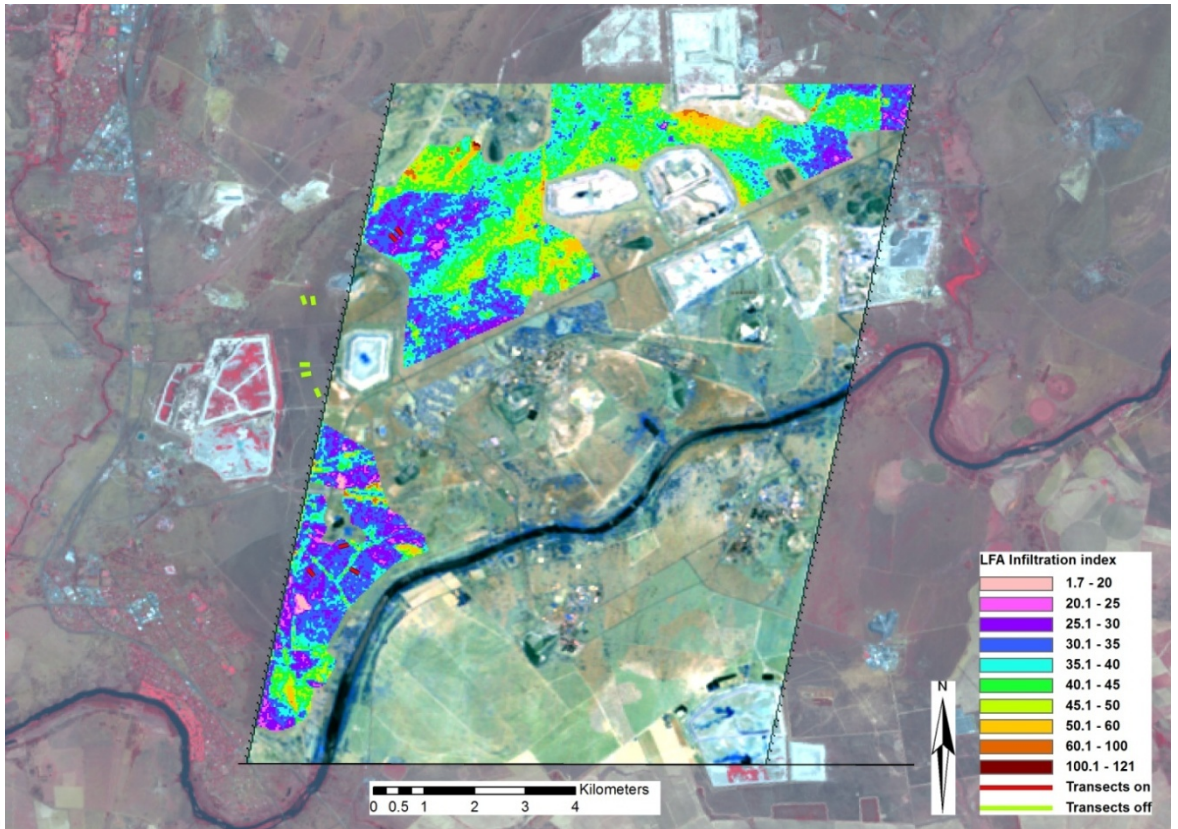
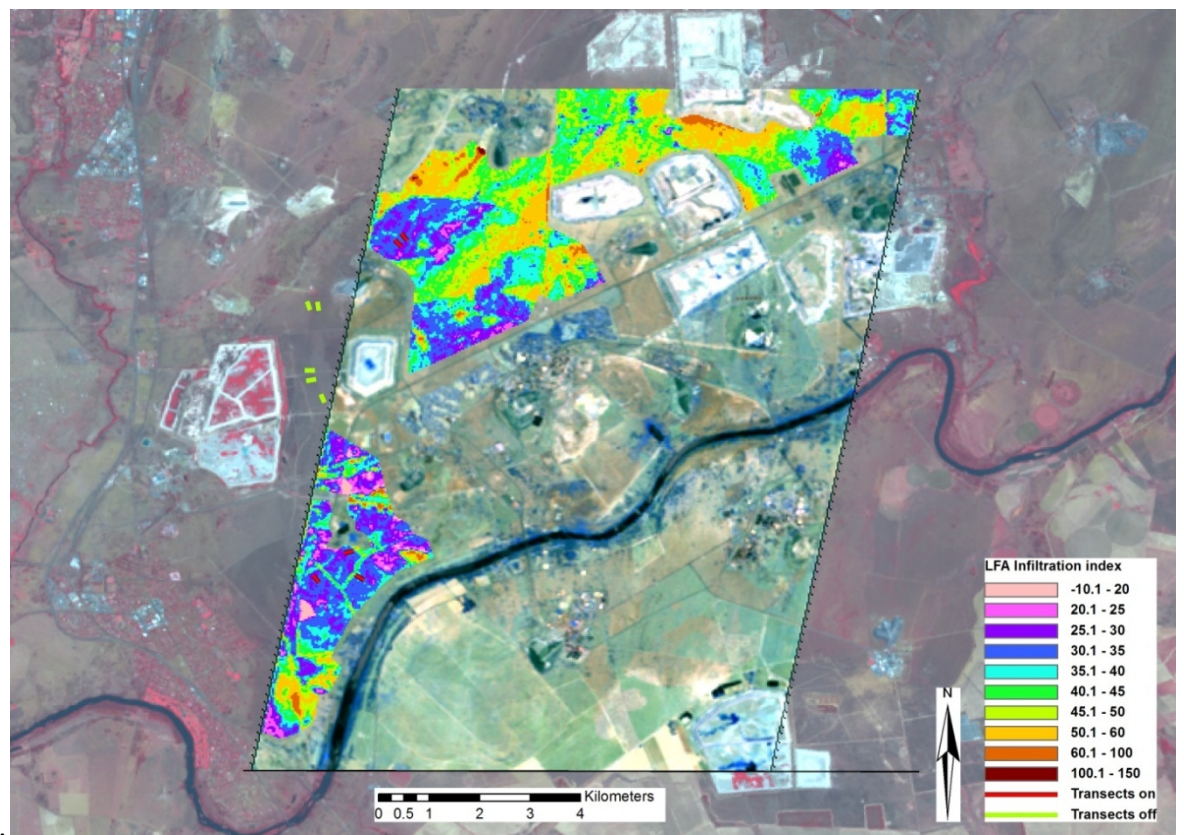


Figure 4.16 Measured versus PLSR predicted LFA infiltration index values using 35 bands from the VNIR block for (a) a 5-component PLSR model and unsmoothed spectra, and (b) a 4-component PLSR model with spectra after four iterations of the smoothing algorithm. Solid lines represent the best fit and dotted lines are the equality line.



4395 a)



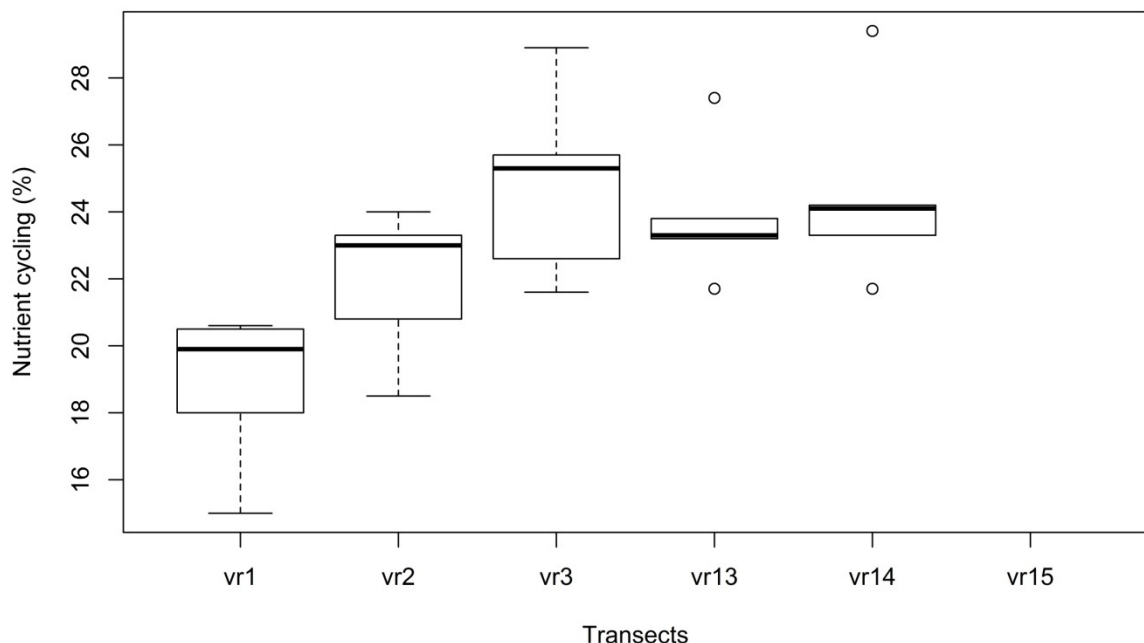
4396 b).

Figure 4.17 LFA infiltration prediction maps for (a) a 35 band, 5-component model with no spectral smoothing, and (b) the same bands with a 4-component model and four iterations of the smoothing algorithm.

transects. However, this assumption of similar distributions is not a very rigorous assumption as the study was not designed to predict the eastern edges of the northern grassland and features such as roads and tailings spills did not have their spectra incorporated into the PLSR modelling. Generally, the two prediction maps (Figure 4.17a and b) are very similar with the major difference between the two models being that the range in the 4th iteration model is much greater (-10 to 150%) than that in the unsmoothed 5-component model (1.7 – 120%). Notwithstanding this, the values predicted for the LFA infiltration index (25.1 – 35%) in the areas where transects were run correspond quite well with the range of values measured in the field (26.8 – 36.1%).

4.3.4 Predicting the LFA Nutrient Cycling Index

The distribution of the LFA nutrient cycling index for pixels associated with transects shows strong similarities with the LFA infiltration index. Transect vr1 has the lowest values followed by vr2 and vr3 has the highest values with vr13 and vr14 having a distribution that generally lies between vr2 and vr3 values (Figure 4.18). There is far less variability in the LFA nutrient cycling index for vr13 and vr14 compared to the southern grassland transects (Figure 4.18).



4416

4417 **Figure 4.18** Distribution of LFA nutrient cycling index values for pixels from each transect
4418 at Vaal River

The PLSR modelling followed a similar pattern as that for the LFA stability and infiltration indices with the only portions of the Hyperion spectrum providing relatively strong PLSR models being the near SWIR and the VNIR blocks. The full spectra models using 105 bands selected from across the Hyperion spectrum were slightly stronger than they were for the previous two LFA indices with the best model (validation $r^2 = 0.47$, $P = 0.04$; Table 4.6) being a 4-component model derived using a Hyperion data cube subjected to two iterations of the smoothing algorithm. However, the prediction map for this full spectrum model has a fair amount of striping running parallel to the direction of the EO-1 satellite track. For the near SWIR block, the unsmoothed spectra had fairly strong prediction statistics with calibration ($r^2 = 0.6054$) and validation data ($r^2 = 0.61$, $P = 0.013$). Smoothing reduced the calibration statistics and improved the validation statistics with each application of the smoothing algorithm, but also increased the number of components in the validation model. Thus, a 3-component PLSR model derived using a spectral data cube subjected to four iterations of the smoothing algorithm produced a validation coefficient of determination of 0.82 ($P = 0.001$) but this dropped to 0.47 ($P = 0.003$) for the calibration data (Table 4.5). Nevertheless, despite these strong calibration statistics, the LFA nutrient cycling index prediction map derived from this 3-component, 4th iteration, near-SWIR PLSR model suffered from the striping problem generally typical of these SWIR blocks.

Therefore, the VNIR block was preferred for predicting the LFA nutrient cycling index. A 3-component PLSR model derived from 40 unsmoothed bands in the VNIR had fairly strong calibration ($r^2 = 0.79$, $P < 0.0001$, $n = 16$ pixels) and validation ($r^2 = 0.69$, $P = 0.006$, $n = 9$ pixels) statistics when predicting the LFA nutrient cycling index (Table 4.5 and Figure 4.19a). The prediction map created from this model (Figure 4.20a) shows good delineation of features such as roads and burn scars and the areas associated with the transects lie within the field measured range of 15.0 – 29.4% for the nutrient cycling. However, it is not clear why certain areas fall into the -8 – 10% range although one of these, the triangular area in the southern prediction block close to the western edge of the Hyperion image is a very recent fire scar and one would expect this to have very low biomass and therefore likely very low nutrient cycling, certainly above and on the soil surface. VNIR models derived from smoothed spectra did weaken in their calibration statistics but not to the extent that occurred in the near SWIR models, and the validation statistics remained more or less

Table 4.6 Exploratory models predicting the LFA nutrient cycling index from the Vaal River Hyperion hyperspectral data cube. The terms “far” and “near” are used as relative terms describing portions of the Hyperion spectrum relative to the visible portion of the spectrum. Raw refers to the spectra as they are, post-FLAASH atmospheric correction, and the terms sm1 – sm4 pertain to the number of iterations of the smoothing algorithm applied to the spectra.

Bands and smoothing model	wavelengths (nm)	Calibration (n = 16 pixels)			Validation (n = 9 pixels)		
		RMSEP (%)	Spectral r ² (%)	Nutrient cycling r ² (%)	comp	Nutrient cycling r ² (%)	P
105 raw	478 - 610, 630 - 885, 1003 - 1084, 1185 - 1306, 1518 - 1578, 1629 - 1780, 2093 - 2123, 2143 - 2163, 2194 - 2295, 2325 - 2335	3.396	98.22	92.20	4	36.37	0.086
105 sm1		3.504	99.00	82.95	4	45.72	0.046
105 sm2		3.272	99.30	78.62	4	47.34	0.040
105 sm3		3.278	99.47	76.24	5	44.76	0.049
105 sm4		3.191	99.70	84.33	5	44.68	0.049
far SWIR raw	2093 - 2113, 2133 - 2264, 2295, 2325 - 2335	3.161	98.69	12.54	1	19.03	0.241
far SWIR sm1		3.167	99.25	12.19	1	19.00	0.241
far SWIR sm2		4.975	99.96	88.82	8	26.12	0.160
far SWIR sm3		3.173	99.48	11.90	1	19.05	0.240
far SWIR sm4		3.174	99.54	11.81	1	19.08	0.240
near SWIR raw	1518 - 1578, 1629 - 1780	3.275	94.94	60.54	2	60.72	0.013
near SWIR sm1		3.408	97.55	50.94	2	71.30	0.004
near SWIR sm2		3.511	98.32	44.04	2	72.15	0.004
near SWIR sm3		3.261	99.30	52.41	3	74.65	0.003
near SWIR sm4		3.407	99.46	46.90	3	82.40	0.001
Far NIR raw	1185 - 1306	2.666	96.07	68.75	2	25.86	0.162
Far NIR sm1		3.272	97.96	65.23	2	25.39	0.167
Far NIR sm2		3.689	98.55	50.20	2	39.82	0.068
Far NIR sm3		3.677	98.93	34.77	2	47.82	0.039
Far NIR sm4		3.604	99.13	28.90	2	48.08	0.038
Near NIR raw	1003 - 1084	3.064	96.97	45.57	2	8.81	0.438
Near NIR sm1		6.198	100.00	57.41	9	32.29	0.110
Near NIR sm2		3.162	98.93	13.71	1	15.88	0.288
Near NIR sm3		3.470	99.99	48.99	5	24.83	0.172
Near NIR sm4		3.553	100.00	46.30	5	40.59	0.065
VNIR raw	478 - 610, 630 - 885	3.540	99.01	78.76	3	68.22	0.006
VNIR sm1		3.757	99.51	60.97	3	64.25	0.009
VNIR sm2		3.799	99.60	58.25	3	62.67	0.011
VNIR sm3		3.691	99.78	64.74	4	67.85	0.006
VNIR sm4		3.670	99.82	63.77	4	69.72	0.005

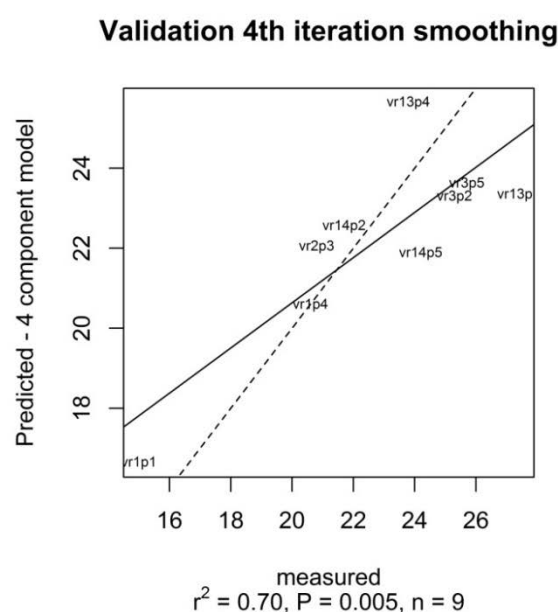
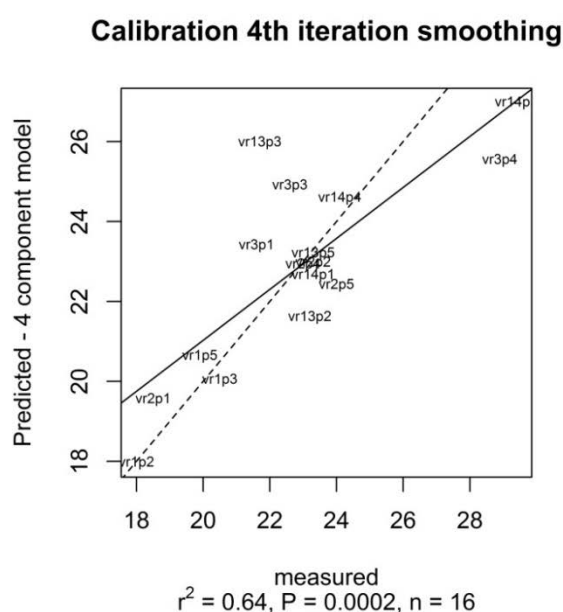
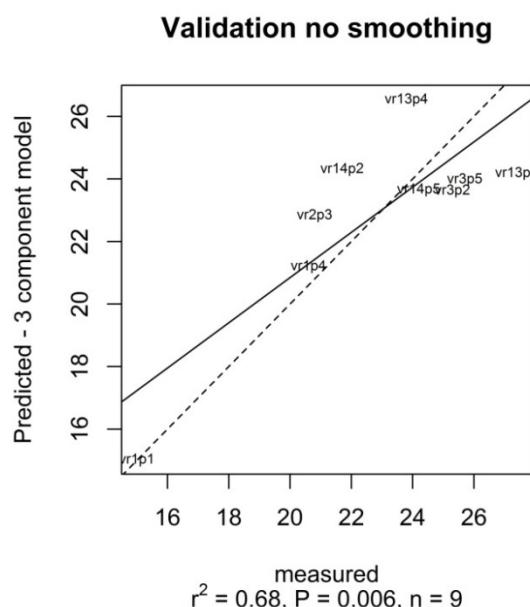
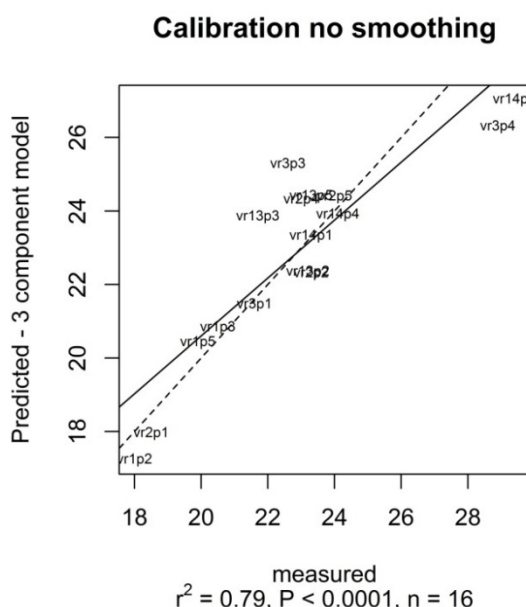
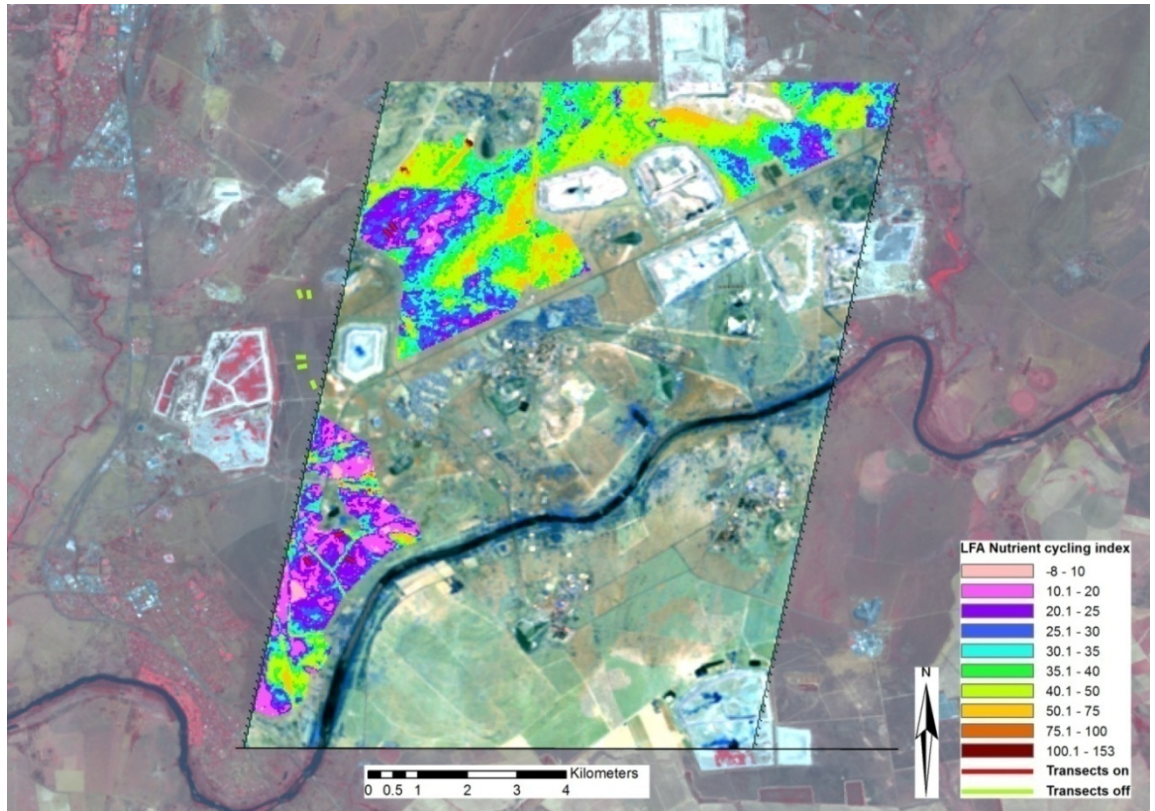
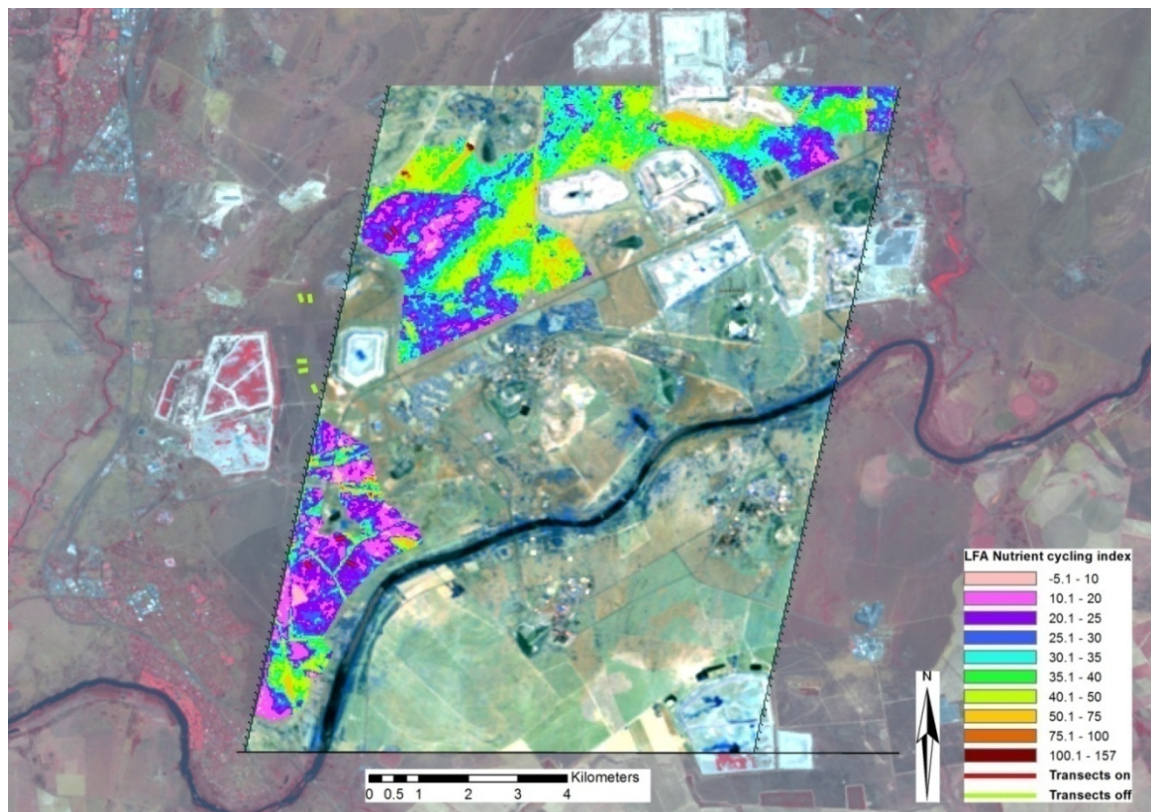


Figure 4.19 Measured versus PLSR predicted LFA nutrient cycling index values using 40 bands from the VNIR block for (a) a 3-component PLSR model with unsmoothed spectra, and (b) a 4-component PLSR model using spectra after four iterations of the smoothing algorithm. Solid lines represent the best fit and dotted lines the equality line.

constant. A 4-component PLSR model using a four iteration smoothing algorithm on the data cube prior to modelling produced a validation coefficient of determination of 0.70 ($P = 0.005$; Table 4.5, Figure 4.19b and Figure 4.20b). The prediction map derived from this model (Figure 4.20b) is very similar to the prediction map derived from the unsmoothed



a)



b)

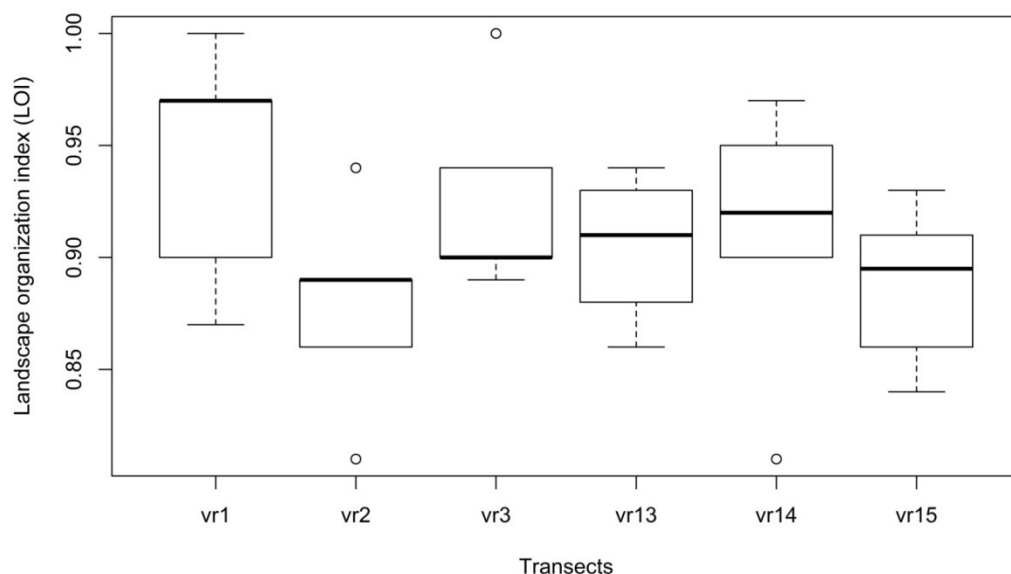
Figure 4.20 LFA nutrient cycling index prediction maps derived from (a) a VNIR, 40 band, 3-component PLSR model with no spectral smoothing, and (b) the same bands but subjected to four iterations of the smoothing algorithm prior to deriving a 4-component PLSR model.

4474 data cube (Figure 4.20a) with the main difference being in the boundary positions of the
4475 various categories.

4476 4.3.5 Predicting the LFA landscape Organization Index

4477 The landscape organization index (LOI) is an indication of the degree to which the transect is
4478 covered in patches. A value of 1 would indicate the whole transect is a single patch and
4479 values less than this would indicate that a portion of the transect is patch and the rest is
4480 exposed soil. The range of LOI values obtained for the transect pixels was between 0.85 and
4481 1 with two pixels having values of 1. The medians for transects vr2 to vr15 are very similar,
4482 just above or below 0.9, but vr1 is considerably higher 0.97. Transects vr2 and vr14 each
4483 have a low outlier (both 0.81) considerably outside the distribution of any transect, while
4484 vr2 and vr3 each have a high outlier which are still within the distribution of the LOI values
4485 for other transects.

4486 The PLSR models were generally very weak regardless of the spectral block considered in
4487 the model. For models with the 105 band full spectrum, the far SWIR, the near NIR and the
4488 VNIR all had a 1-component PLSR model with validation data as their best model regardless
4489 of whether smoothing was applied or not (Table 4.7). Thereafter, the validation results
4490 tended to collapse for these blocks. The block that produced the best, albeit still weak,
4491 validation statistics was the near SWIR block where 23 bands produced a validation



4492
4493 **Figure 4.21** Distribution of LFA landscape organization index (LOI) values for pixels from
4494 each transect at Vaal River

Table 4.7 Exploratory models predicting the LFA landscape organization index from the Vaal River Hyperion hyperspectral data cube. The terms “far” and “near” are used as relative terms describing portions of the Hyperion spectrum relative to the visible portion of the spectrum. Raw refers to the spectra as they are, post-FLAASH atmospheric correction, and the terms sm1 – sm4 pertain to the number of iterations of the smoothing algorithm applied to the spectra.

Bands and smoothing Model	wavelengths (nm)	Calibration (n = 20 pixels)			Validation (n = 11 pixels)		
		RMSEP	Spectral r ² (%)	LOI r ² (%)	comps	LOI r ² (%)	P
105 raw	478 - 610, 630 - 885, 1003 - 1084, 1185 - 1306, 1518 - 1780, 2093 - 2113, 2133 - 2264, 2295, 2325 - 2335	0.054	87.22	0.72	1	27.10	0.101
105 sm1		0.054	88.13	0.36	1	28.13	0.093
105 sm2		0.054	88.33	0.30	1	28.25	0.092
105 sm3		0.053	88.45	0.27	1	28.22	0.093
105 sm4		0.053	88.54	0.25	1	28.18	0.093
far SWIR raw	2093 - 2113, 2133 - 2264, 2295, 2325 – 2335	0.053	97.26	0.44	1	19.08	0.179
far SWIR sm1		0.053	98.13	0.28	1	19.23	0.177
far SWIR sm2		0.053	98.37	0.25	1	19.19	0.178
far SWIR sm3		0.053	98.50	0.24	1	19.17	0.178
far SWIR sm4		0.053	98.58	0.23	1	19.17	0.178
near SWIR raw	1518 - 1578, 1629 - 1780	0.046	98.4	61.70	3	47.40	0.019
near SWIR sm1		0.050	99.29	51.03	3	41.00	0.034
near SWIR sm2		0.052	96.89	0.33	1	33.81	0.061
near SWIR sm3		0.052	97.14	0.28	1	34.11	0.059
near SWIR sm4		0.052	97.28	0.25	1	34.19	0.059
far NIR raw	1185 - 1306	0.054	96.22	0.92	1	28.82	0.089
far NIR sm1		0.053	98.57	0.14	1	31.11	0.075
far NIR sm2		0.086	99.98	66.72	7	37.90	0.044
far NIR sm3		0.070	100.00	62.94	9	34.53	0.057
far NIR sm4		0.052	99.13	0.05	1	31.27	0.074
near NIR raw	1003 - 1084	0.052	81.94	1.86	1	28.40	0.091
near NIR sm1		0.052	98.59	0.08	1	33.15	0.064
near NIR sm2		0.052	99.16	0.05	1	33.21	0.064
near NIR sm3		0.052	99.39	0.05	1	33.39	0.063
near NIR sm4		0.052	99.53	0.04	1	33.55	0.062
VNIR raw	478 - 610, 630 - 885	0.053	94.54	0.69	1	36.06	0.051
VNIR sm1		0.052	95.29	0.47	1	35.73	0.052
VNIR sm2		0.053	95.56	0.41	1	35.57	0.053
VNIR sm3		0.053	95.76	0.38	1	35.46	0.053
VNIR sm4		0.053	95.91	0.36	1	35.37	0.054

coefficient of determination of 0.47 ($P = 0.019$; Table 4.7). A subset of the near SWIR using 13 bands produced even stronger models with the unsmoothed Hyperion spectra having a validation coefficient of determination of 0.71 ($P = 0.001$; Table 4.8 and Figure 4.22). Smoothing the spectra for the 13-band near-SWIR models initially caused an increase in the strength of the validation statistics but the calibration statistics showed a continuous decline with iteration of the smoothing algorithm and after the first iteration, the validation statistics also fell with each additional iteration. However, despite the good prediction statistics, the prediction map using either the unsmoothed or smoothed Hyperion spectra was of poor quality with major striping obscuring any real patterns in the prediction maps (Figure 4.23). The block with the least striping is the VNIR block. However, this block produced very weak model statistics and subsets of the VNIR showed no improvement in the models. Only a 1-component model produced any validation statistics of any strength ($r^2 = 0.36$, $P = 0.05$; Table 4.7) but plotting the predicted versus measured values showed a

Table 4.8 PLSR models predicting the LFA landscape organization index using 13 select bands from the near-SWIR region (1518 – 1679 nm) of the Hyperion spectra.

Smoothing iteration		raw	sm1	sm2	sm3	sm4
Model components		5	6	5	6	6
Calibration	RMSEP	0.060	0.055	0.058	0.059	0.059
	Spectral r^2	99.37	99.89	99.88	99.98	99.99
	Stability r^2	66.64	63.24	52.16	48.39	43.76
Validation	Stability r^2	70.57	75.36	73.09	65.44	56.25
	Prob	0.001	0.0005	0.0008	0.003	0.008
LFA landscape organization index field measure	Mean	0.907				
	Standard deviation	0.048				
Model prediction	Mean	0.903	0.897	0.892	0.896	0.896
	Standard deviation	0.047	0.047	0.044	0.043	0.049
Field distribution vs Hyperion predicted distribution	T-test	0.48	1.26	1.76	1.27	1.23
	Degrees of freedom	30.06	30.06	30.05	30.05	30.07
	Probability	0.63	0.22	0.09	0.21	0.23
	Wilcoxon rank-sum	470888	503419	525404	504961	503878
	Probability	0.60	0.22	0.09	0.21	0.22

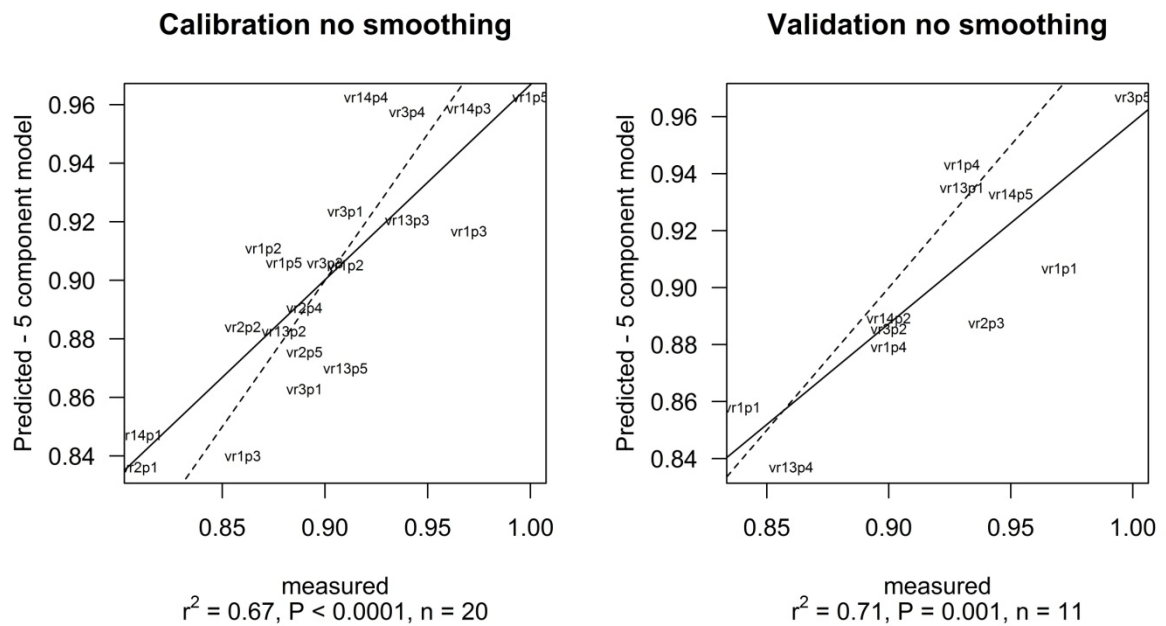


Figure 4.22 Measured versus PLSR predicted LFA landscape organization index using 13 bands from the near SWIR block with no spectral smoothing.

weak inverse relationship for the validation predictions (Figure 4.24a). Dropping either or both of the two very low outliers in transects vr2 and vr14 (Figure 4.21) from the analysis did not improve the models, with the best model when both were dropped being a 1-component model (calibration $r^2 = 0.04$, $P = 0.39$, $n = 19$ pixels, validation $r^2 = 0.22$, $P = 0.17$, $n = 11$). Although this model was weaker than that with all pixels, the best fit line was at least indicating a direct relationship between measured and predicted pixel values (Figure 4.24b). These two models were used to derive the prediction maps shown in Figure 4.25 and it can be seen how the first map (a) is an inversion of the second map (b) despite the model statistics being very weak. Furthermore, both prediction maps for the LOI show inverted predictions for burnt and unburnt areas between the southern and northern prediction blocks (Figure 4.25a and b). In the northern block, burnt areas ranged from 0.91 – 0.93 while in the southern area they ranged from 0.88 – 0.91 and unburnt grasslands in the northern block ranged from 0.88 – 0.91 while in the southern block unburnt grasslands ranged from 0.91 – 0.93 for the LOI index (Figure 4.25b) suggesting that the mixed presence of fire on the transects is having a major impact on the models. There is also major compression in the prediction values with field-measured values for LOI between 0.81 – 1.00 for 31 pixels on transects but predictions for the grasslands were in a much smaller range between 0.91 and



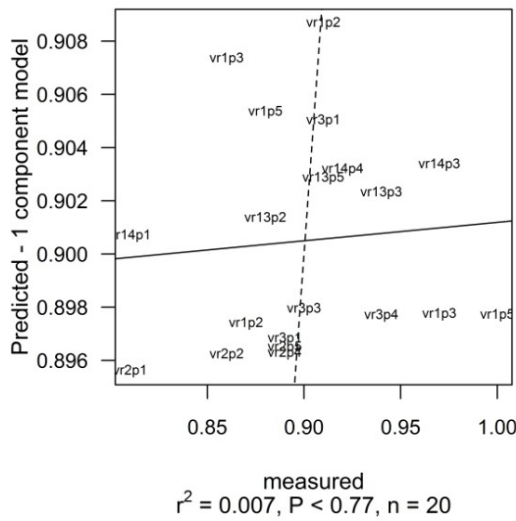
4539

4540

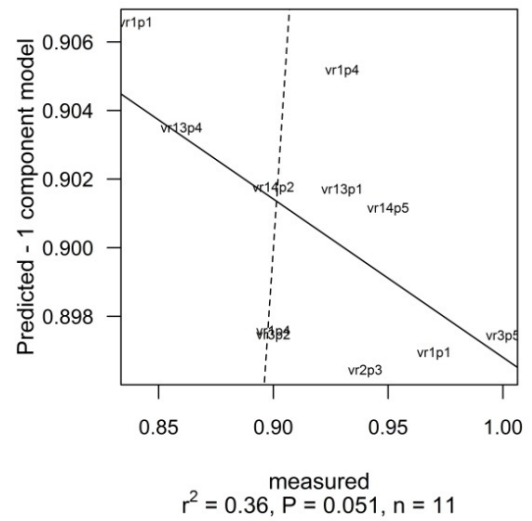
4541

Figure 4.23 LFA landscape organization index prediction map derived with a 6-component PLSR model using 13 bands in the near SWIR block with one application of the smoothing algorithm prior to PLSR modelling.

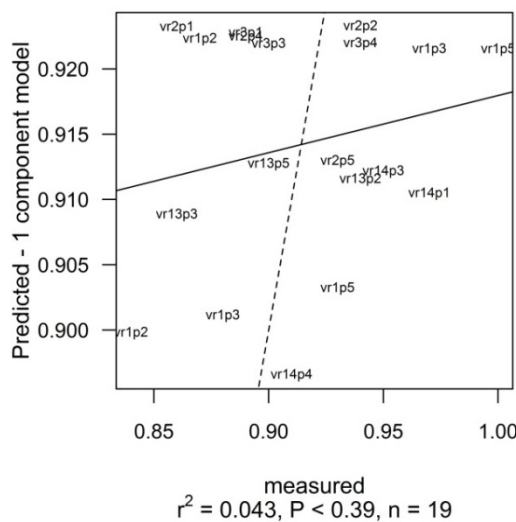
Calibration no smoothing



Validation no smoothing



Calibration no smoothing



Validation no smoothing

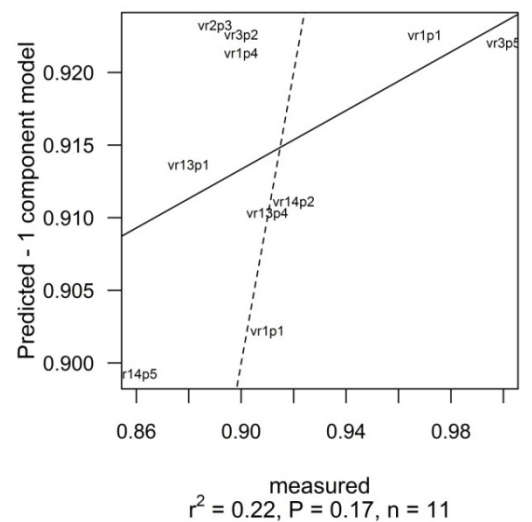


Figure 4.24 Measured versus predicted LFA LOI values for (a) a 1-component PLSR model using 40 unsmoothed VNIR bands, and (b) a 1-component PLSR model using 40 unsmoothed bands from the VNIR with two outlier pixels removed from the analysis.

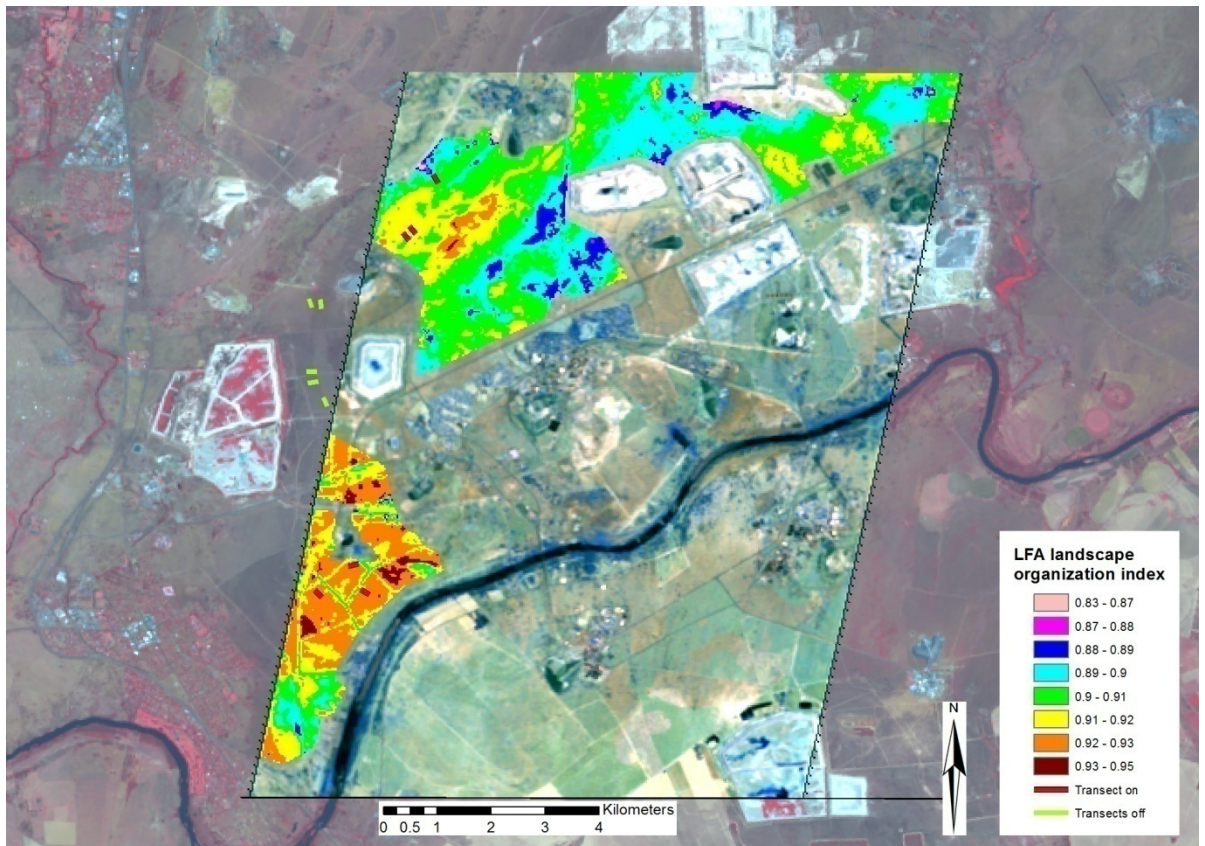
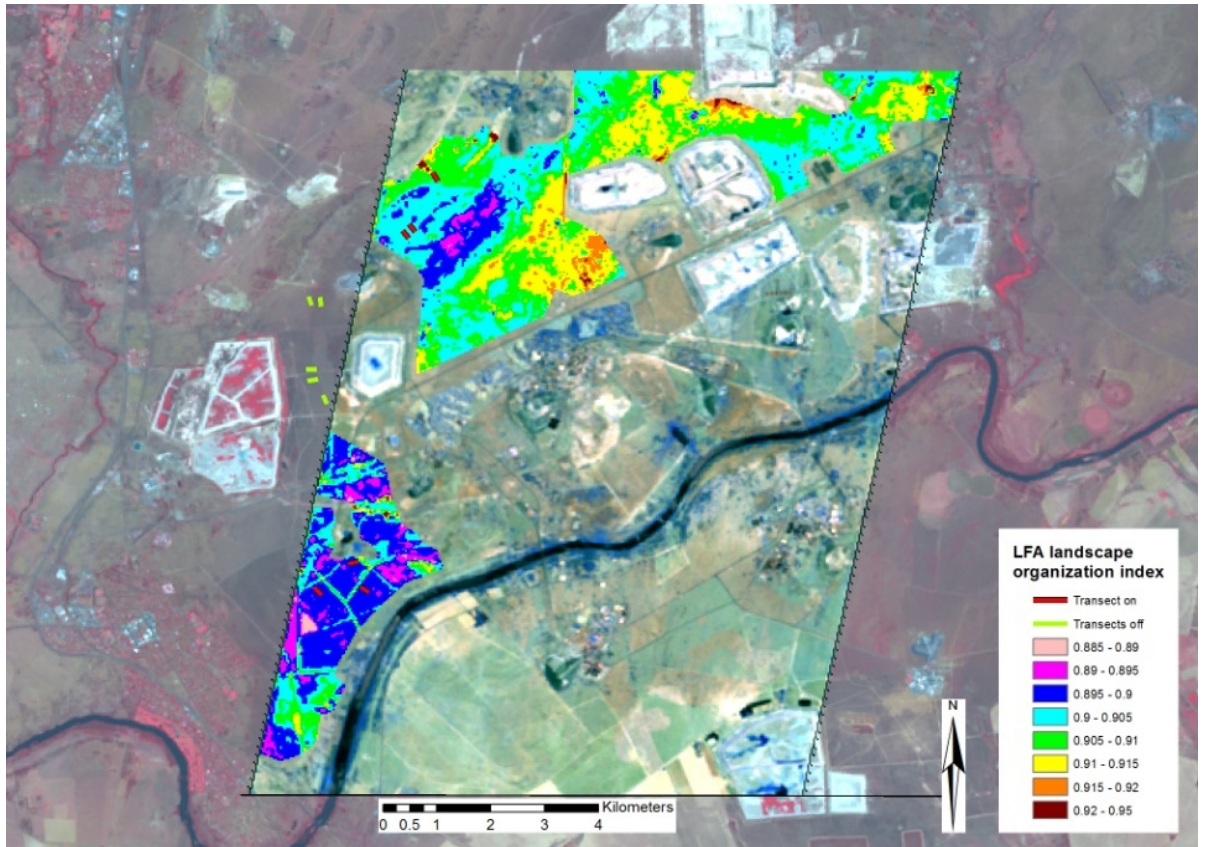


Figure 4.25 LFA LOI index prediction map derived with (a) a 1-component PLSR model using 40 unsmoothed bands from the VNIR block, and (b) a 1-component model using the same blocks but with two outlier pixels removed.

4.3.6 Predicting the LFA Patch Area Index

The LFA patch area index (PAI) is the ratio of patch area to total transect area where a value of 1 would indicate that the whole transect is patch with no inter-patch zones. The highest PAI's were recorded for vr1 closely followed by vr2, vr3, vr13 and vr14 (Figure 4.26). Transect vr15 had a considerably lower distribution of PAI values compared to the other five transects probably due to severe grazing by cattle as it was observed while sampling vr13 and vr14 that cattle grazed this area every afternoon (Figure 4.26). There was one outlier on vr14 that was very low (Figure 4.26) although the reason for this is not clear from the raw data. Removing this pixel from the analysis significantly altered the Hyperion PLSR model output.

Models with all 31 pixels were very weak (results not shown). Dropping the first pixel from vr14 produced PLSR models predicting the PAI with validation data that were stronger than any model predicting any other index or variable in this study. This strength in the models was common whether the model was derived from the full spectrum or from specific blocks within the spectrum (Table 4.9). The coefficients of determination for the models with calibration data were lower than they were for the validation data but as these models used

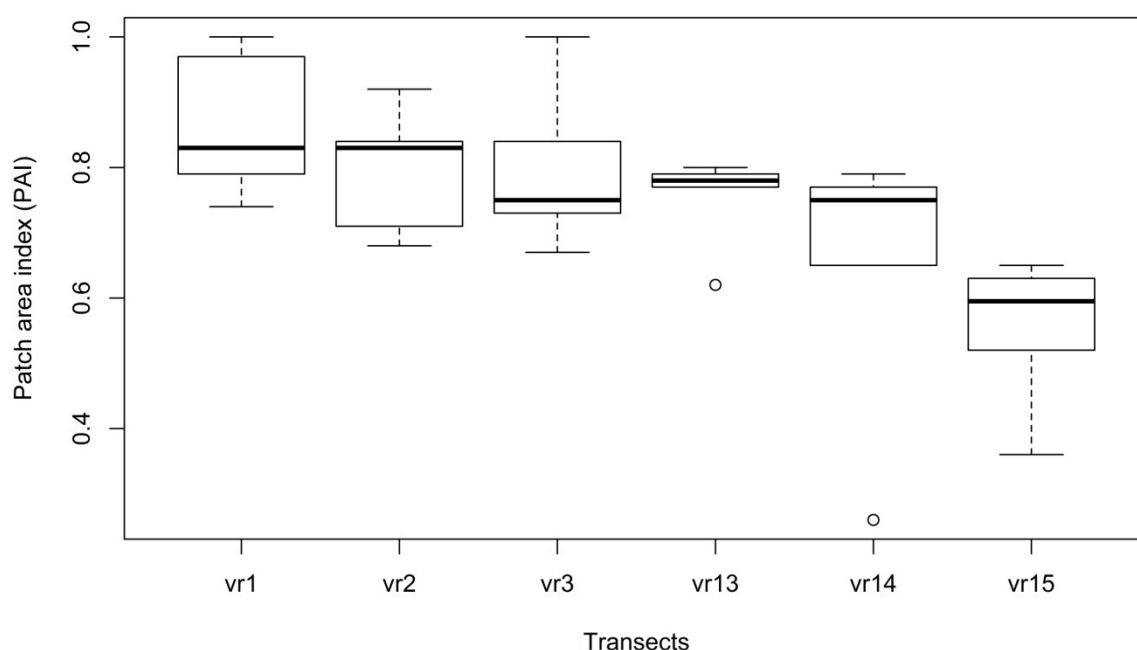


Figure 4.26 Distribution of LFA patch area index (PAI) values for pixels from each transect at Vaal River

Table 4.9 Exploratory models predicting the LFA patch area index (PAI) from the Vaal River Hyperion hyperspectral data cube. The terms “far” and “near” are used as relative terms describing portions of the Hyperion spectrum relative to the visible portion of the spectrum. Raw refers to the spectra as they are, post-FLAASH atmospheric correction, and the terms sm1 – sm4 pertain to the number of iterations of the smoothing algorithm applied to the spectra.

Bands and smoothing model	wavelengths (nm)	Calibration (n = 19 pixels) Spectral r ²			Validation (n = 11 pixels) PAI r ²		
		RMSEP	(%)	PAI r ² (%)	comps	(%)	P
105 raw	478 - 610, 630 - 885, 1003 - 1084, 1185 - 1306, 1518 - 1780, 2093 - 2113, 2133 - 2264, 2295, 2325 – 2335	0.100	98.29	39.19	2	77.84	0.0003
105 sm1		0.100	98.92	38.61	2	77.47	0.0004
105 sm2		0.100	99.08	38.57	2	77.53	0.0003
105 sm3		0.100	99.17	38.58	2	77.58	0.0003
105 sm4		0.100	99.23	38.59	2	77.60	0.0003
far SWIR raw	2093 - 2113, 2133 - 2264, 2295, 2325 – 2335	0.099	98.62	52.27	2	71.01	0.0011
far SWIR sm1		0.098	99.24	49.48	2	72.82	0.0008
far SWIR sm2		0.097	99.41	50.72	2	71.96	0.0010
far SWIR sm3		0.095	99.50	51.93	2	70.76	0.0012
far SWIR sm4		0.094	99.56	52.84	2	69.60	0.0014
near SWIR raw	1518 - 1578, 1629 - 1780	0.126	96.99	49.85	2	83.75	0.0001
near SWIR sm1		0.128	98.52	41.01	2	78.46	0.0003
near SWIR sm2		0.122	99.04	37.95	2	79.57	0.0002
near SWIR sm3		0.118	99.31	36.97	2	79.77	0.0002
near SWIR sm4		0.115	99.47	36.58	2	79.44	0.0002
Far NIR raw	1185 - 1306	0.095	98.29	38.13	1	79.46	0.0002
Far NIR sm1		0.095	98.98	37.94	1	79.46	0.0002
Far NIR sm2		0.095	99.12	37.90	1	79.41	0.0002
Far NIR sm3		0.095	99.21	37.90	1	79.35	0.0002
Far NIR sm4		0.095	99.28	37.90	1	79.29	0.0002
Near NIR raw	1003 - 1084	0.095	98.79	38.52	1	72.19	0.0009
Near NIR sm1		0.095	99.51	38.37	1	72.74	0.0008
Near NIR sm2		0.095	99.68	38.33	1	72.93	0.0008
Near NIR sm3		0.095	99.77	38.33	1	73.01	0.0008
Near NIR sm4		0.095	99.82	38.33	1	73.04	0.0008
VNIR raw	478 - 610, 630 - 885	0.098	98.78	42.62	2	78.57	0.0003
VNIR sm1		0.098	99.03	41.80	2	78.37	0.0003
VNIR sm2		0.098	99.10	41.57	2	78.50	0.0003
VNIR sm3		0.098	99.15	41.41	2	78.61	0.0003
VNIR sm4		0.098	99.18	41.29	2	78.70	0.0003

very few components, they were generally fairly strong for models with one or two components in them (Table 4.9). The full spectrum model using 105 bands from across the Hyperion spectrum had a validation r^2 value of 0.78 ($P = 0.0003$) with a 2-component model and unsmoothed spectra (Table 4.9). This coefficient of determination barely changed for the validation data with any iteration of the smoothing algorithm, and this pattern was also true for the calibration data although the coefficients of determination were very slightly lower (Table 4.9).

The near SWIR with no spectral smoothing produced the strongest 2-component model with calibration ($r^2 = 0.50$, $P = 0.0007$) and validation data ($r^2 = 0.84$, $P = 0.0001$; Table 4.9) but the prediction map suffered the usual problem with the near SWIR block and was very stripey. Both the far NIR and near NIR blocks required only one component to produce strong PLSR models predicting the PAI with coefficients of determination in the seventies for validation data whether the spectra were smoothed or not (Table 4.9) but both these blocks produce stripey prediction maps. The VNIR block performed very slightly better than did the full spectrum block with calibration and validation data and there was little change whether the spectra were smoothed or not. The best VNIR model with validation data ($r^2 = 0.79$, $P = 0.0003$) was a 2-component model with spectra smoothed with four iterations of the smoothing algorithm. Plotting the predicted versus measured values for the 2-component, 4th iteration smoothed 40 band VNIR PLSR model (Figure 4.26b) surprisingly showed only subtle differences from the same plot of the full spectrum 105 bands, 2-component, no smoothing PLSR model (Figure 4.26a). Likewise the prediction maps from these two models show only subtle difference from each other with the 40-band 4th iteration smoothing model (Figure 4.27b) having a lower minimum value than the 105 band unsmoothed model (Figure 4.27a) and subtle differences in the position of boundaries of the different classes. In the VNIR PAI prediction map (Figure 4.28b), in the burnt area around transects vr13 and vr14 in the northern prediction block, prediction values range between 0.7 and 0.8. This is consistent with other burnt areas in this northern prediction area which tend to range from 0.65 to 0.8. The unburnt areas in this northern block have values mainly between 0.5 and 0.7. However the southern prediction block has unburnt grassland values between 0.75 to 0.9 and values for burnt areas that range from 0.5 to 1.1. Despite the burn, these values are consistent with the distribution of transect values

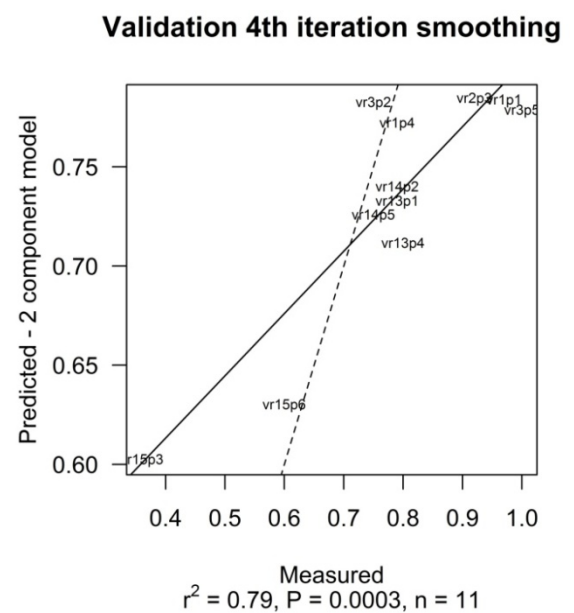
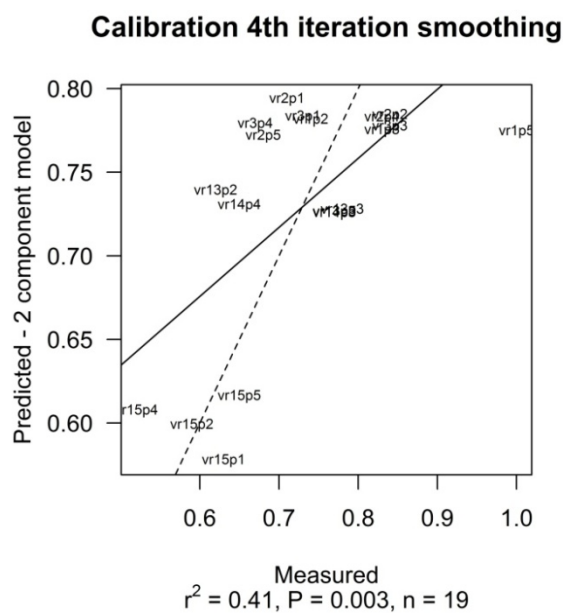
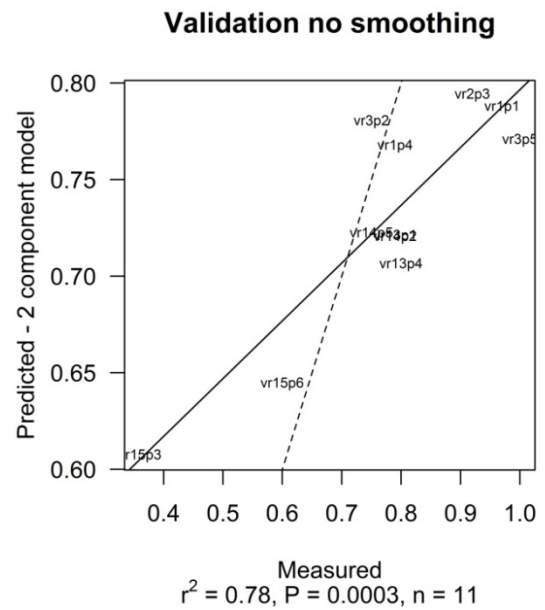
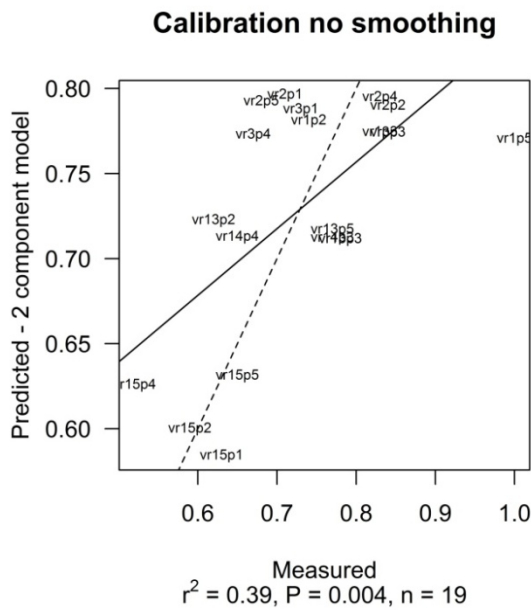
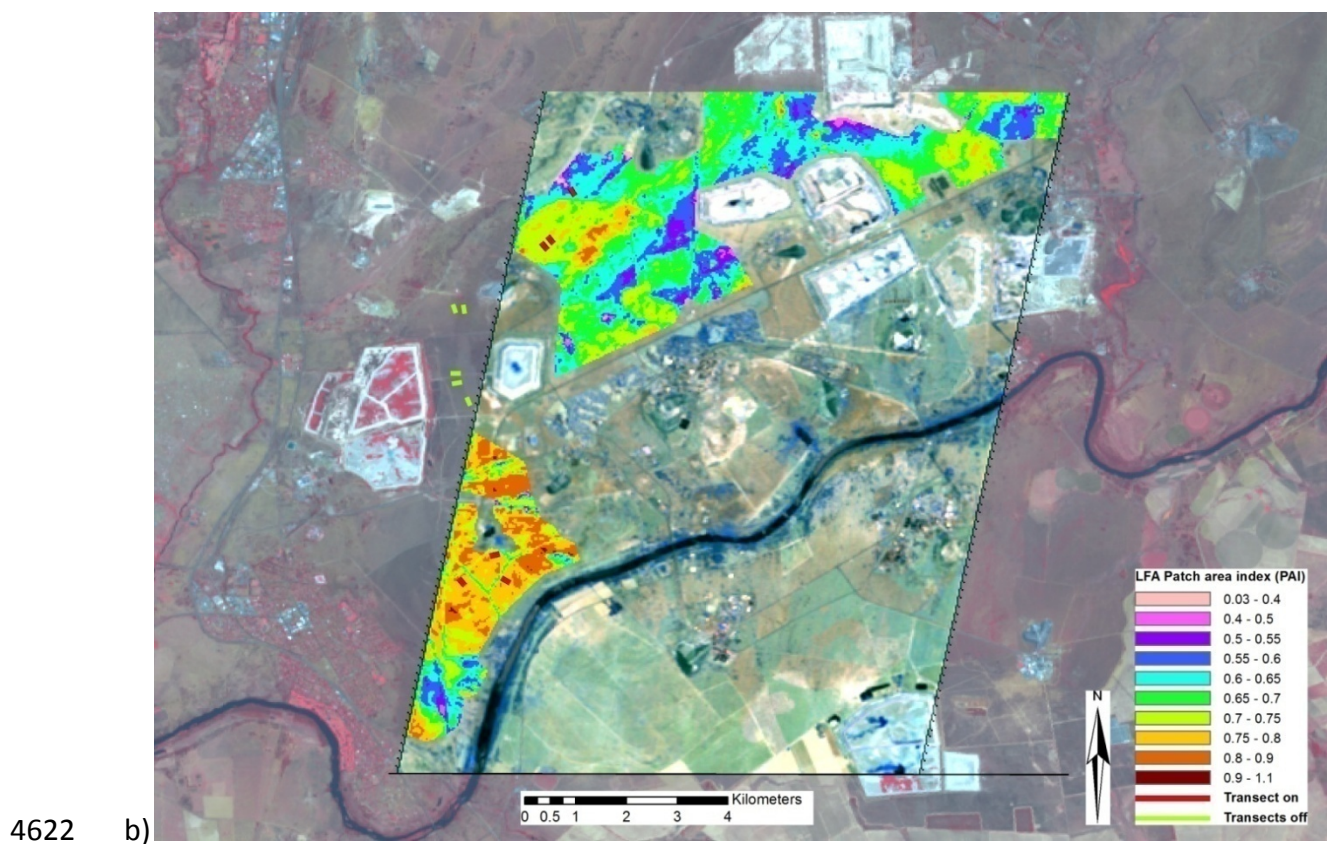
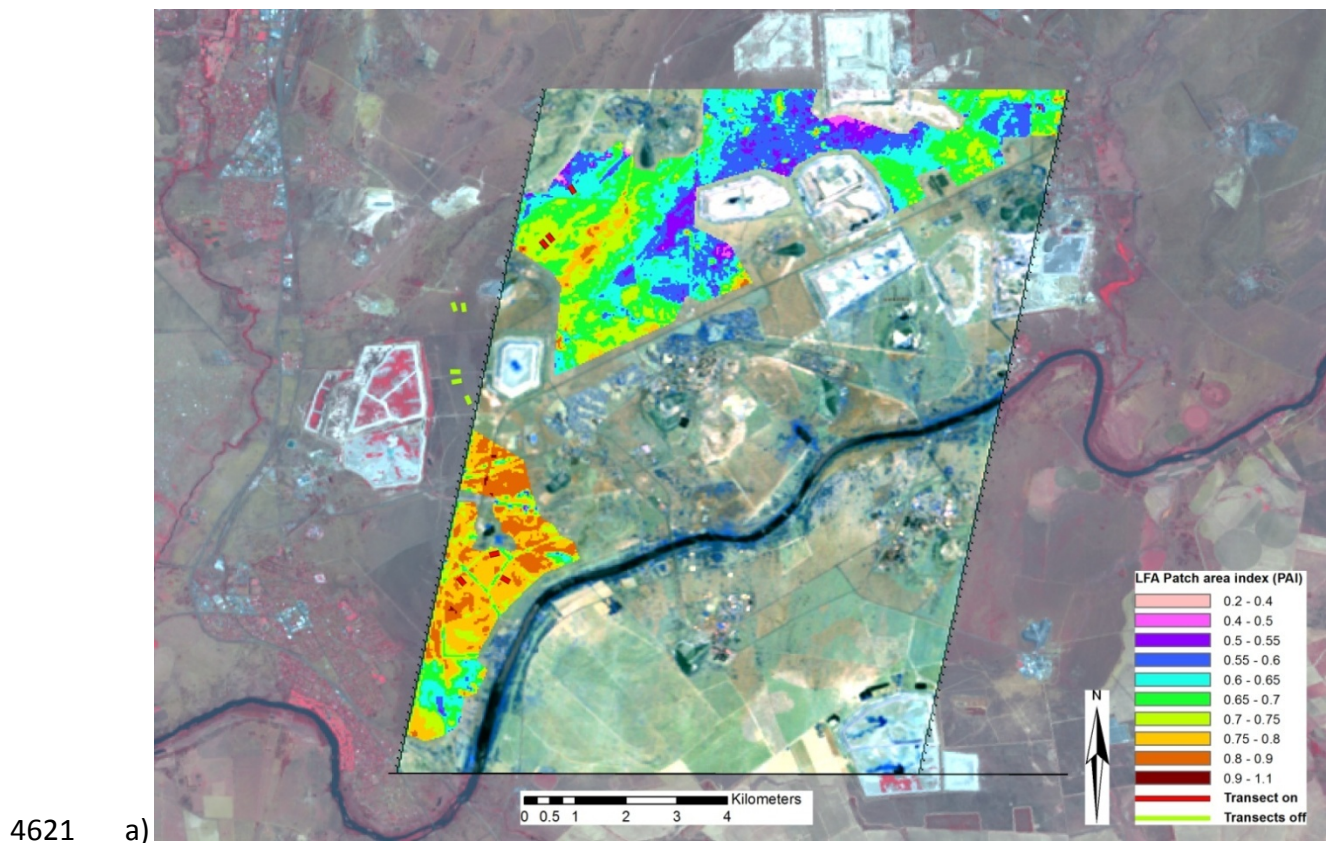


Figure 4.27 Measured versus PLSR predicted LFA patch area index (PAI) values using (a) 105 bands across the whole Hyperion spectrum and a 2-component model with no spectral smoothing, and (b) a PLSR model derived from 40 bands in the VNIR where the spectra were subjected to four iterations of the smoothing algorithm prior to PLSR modelling.



4623 **Figure 4.28** LFA patch area index (PAI) prediction maps for a (a) 105 band, full spectrum,
 4624 2-component model with no spectral smoothing, and (b) 40 VNIR bands with
 4625 4 iterations of the smoothing algorithm in a 2-component PLSR model.

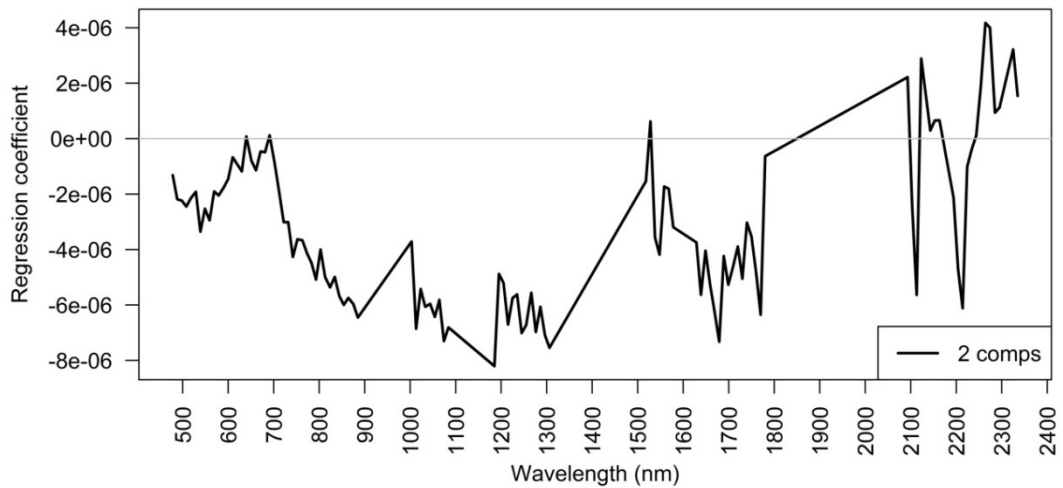
presented in the box plots in Figure 4.24. The regression coefficients for the 105-band full spectrum model with no smoothing (Figure 4.29a) shows a very jagged profile. The effect of smoothing on the spectra can be clearly seen in Figure 4.29b where a much smoother profile is apparent. Now the effect of the clay signal around 2200 nm is clear and a weaker chlorophyll red-edge effect around 670 – 700 nm. The chlorophyll feature is more clearly seen in Figure 4.29c showing the regression coefficient for the 2-component VNIR PLSR model where the spectra were subjected to four iterations of the smoothing algorithm.

4.3.7 Predicting Above-Ground Biomass

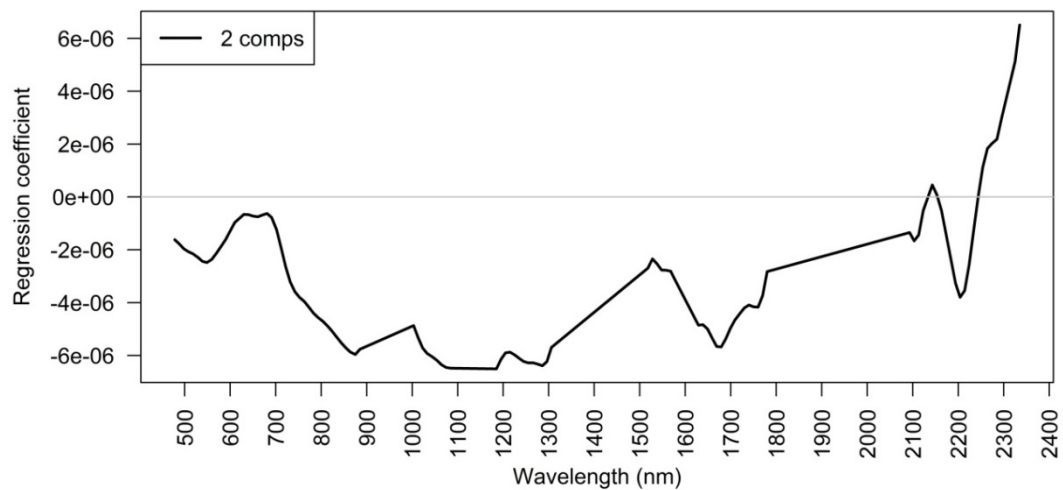
Biomass values are only available for five of the six transects that were covered by the Hyperion image as transect vr15 was burnt by local herdsman after collecting transect measurements. The two transects in the northern grassland near the Black Reef ridge, vr13 and vr14, had a wider range in dry biomass values compared to the three transects (vr1, vr2 and vr3) in the southwestern grassland near the Vaal River (Figure 4.30). Transect vr3 had the highest mean dry biomass ($0.350 \pm 0.141 \text{ kg.m}^{-2}$) and vr14 had the lowest mean dry biomass ($0.292 \pm 0.093 \text{ kg.m}^{-2}$) which was very slightly lower than the mean for vr2 ($0.295 \pm 0.040 \text{ kg.m}^{-2}$; Figure 4.30).

The PLSR models predicting dry biomass were weak regardless of the spectral block used in the modelling or any smoothing applied to the spectra prior to modelling (Table 4.10). This applied whether biomass was modelled on a kilogram per square metre basis, kilogram per pixel, or fresh or dry weights. As the sparse grass patch type formed the matrix in which the other patch types essentially formed islands, and sparse grass overwhelmingly dominated the transects, biomass for this patch type was also tested in models and was also found to be weak. Therefore, only results for dry biomass on kg.m^{-2} basis are presented here.

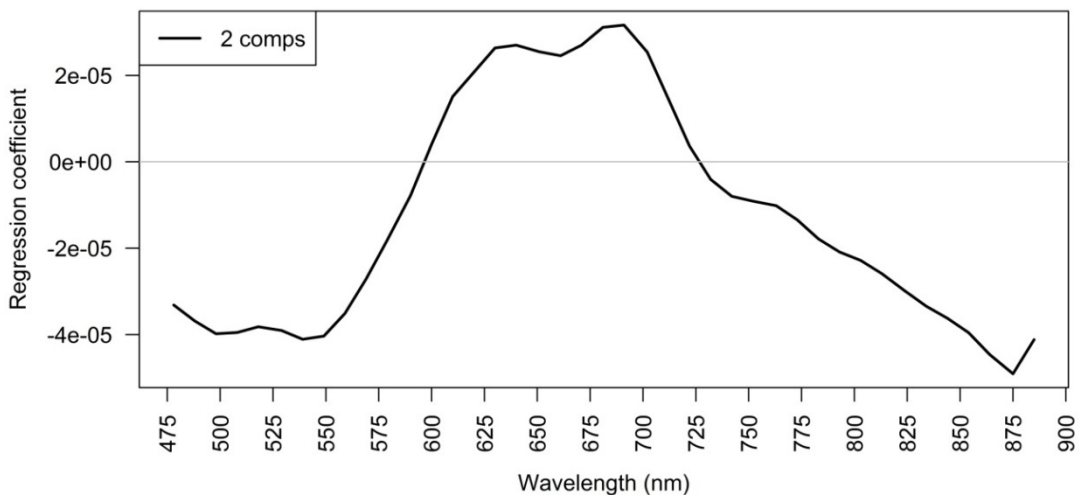
Full spectrum models using 105 select bands from across the whole Hyperion spectra were strongest in predicting dry biomass using 2-component models regardless of whether smoothing was applied or not. However, these were fairly weak with unsmoothed spectra giving a validation coefficient of determination of 0.25 ($P = 0.169$, $n = 9$ pixels) although the calibration statistics ($r^2 = 0.55$, $P = 0.001$, $n = 16$ pixels) for this model were fairly strong for a 2-component model in comparison to the results of other variables in this study (Table 4.10 and Figure 4.31a). Results of full spectrum PLSR models using smoothed spectra were very



4656 a)



4657 b)



4658 c)

4659 **Figure 4.29** Regression coefficients for (a) a 2-component unsmoothed full spectrum
 4660 PLSR model, and (b) the same bands but with four iterations of the
 4661 smoothing algorithm, and (c) a 2-component 40 band VNIR model with four
 4662 iterations of the smoothing algorithm

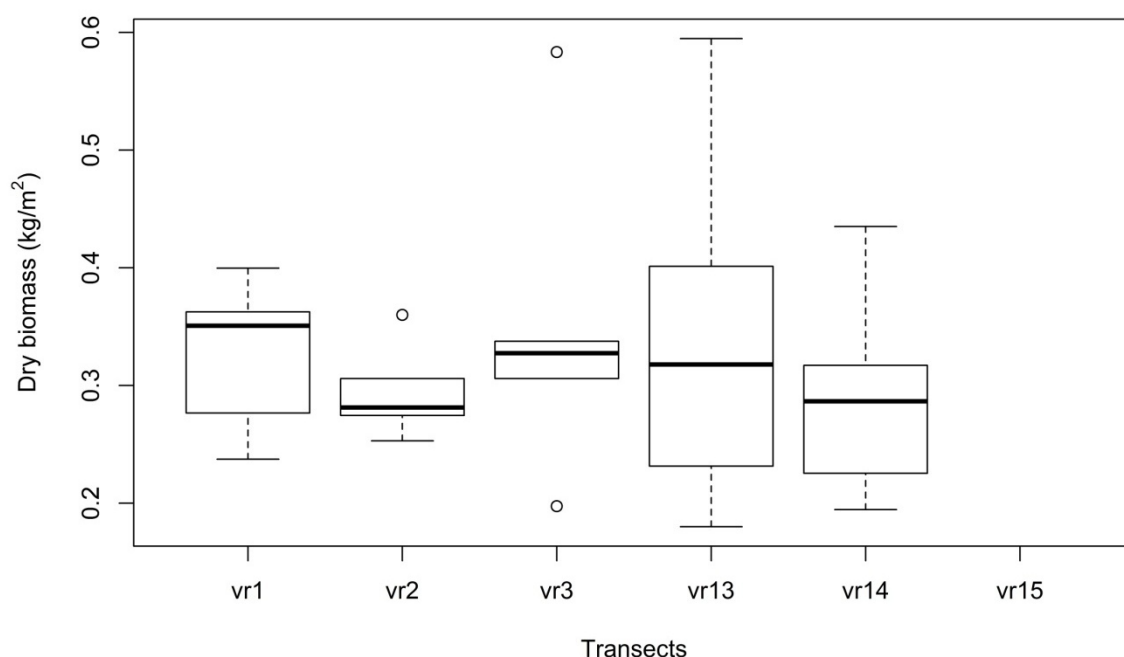


Figure 4.30 Distribution of dry biomass values ($\text{kg}\cdot\text{m}^{-2}$) for pixels matching five transects at Vaal River

similar to the unsmoothed results for validation data but weakened very slightly for calibration data with each iteration of the smoothing algorithm (Table 4.10). The prediction map for the 2-component model using unsmoothed spectra is shown in Figure 4.32a. The prediction map does not show any major speckling and the biomass values predicted around the transects do fall into the range of the values measured in the field. However, there are areas with negative biomass ($-1.1 - 0 \text{ kg}\cdot\text{m}^{-2}$) and areas with very high biomass for a short grassland ($1 - 3.2 \text{ kg}\cdot\text{m}^{-2}$) which supports the fact that the model predictions are not very strong.

The best model for predicting dry biomass with validation data was derived from the far NIR block which consisted of thirteen bands between 1185 and 1306 nm. This model required five components and spectra were smoothed with one application of the smoothing algorithm. The coefficient of determination for this model with calibration data was 0.54 ($P = 0.001$, $n = 16$ pixels) and with validation data was 0.45 ($P = 0.048$, $n = 9$ pixels; Table 4.10). However an examination of the predicted versus measured values showed an inverse relationship between predicted and measured values. Furthermore, the prediction map was

Table 4.10 Exploratory models predicting the dry biomass (kg/m²) on each pixel from the Vaal River Hyperion hyperspectral data cube. The terms “far” and “near” are used as relative terms describing portions of the Hyperion spectrum relative to the visible portion of the spectrum. Raw refers to the spectra as they are, post-FLAASH atmospheric correction, and the terms sm1 – sm4 pertain to the number of iterations of the smoothing algorithm applied to the spectra.

Bands and smoothing model	wavelengths (nm)	Calibration (n = 16 pixels)			Validation (n = 9 pixels)		
		RMSEP (kg/m ²)	Spectral r ² (%)	Biomass r ² (%)	comps	Biomass r ² (%)	P
105 raw	478 - 610, 630 - 885, 1003 - 1084, 1185 - 1306, 1518 - 1780, 2093 - 2113, 2133 - 2264, 2295, 2325 - 2335	0.095	97.61	54.70	2	25.18	0.169
105 sm1		0.092	98.49	50.25	2	25.55	0.165
105 sm2		0.091	98.72	49.29	2	25.53	0.165
105 sm3		0.091	98.84	48.67	2	25.38	0.167
105 sm4		0.091	98.93	48.21	2	25.16	0.169
far SWIR raw	2093 - 2113, 2133 - 2264, 2295, 2325 - 2335	0.128	98.86	55.64	2	6.89	0.495
far SWIR sm1		0.110	99.25	2.90	1	6.30	0.515
far SWIR sm2		0.110	99.41	2.89	1	6.33	0.514
far SWIR sm3		0.129	99.59	26.27	2	6.76	0.499
far SWIR sm4		0.129	99.63	22.85	2	8.83	0.437
near SWIR raw	1518 - 1578, 1629 - 1780	0.107	95.25	49.14	2	20.19	0.225
near SWIR sm1		0.103	97.81	40.38	2	38.95	0.072
near SWIR sm2		0.103	98.56	35.18	2	41.58	0.061
near SWIR sm3		0.103	98.93	32.36	2	43.11	0.055
near SWIR sm4		0.103	99.14	30.59	2	44.41	0.050
Far NIR raw	1185 - 1306	0.128	95.84	33.66	2	29.37	0.132
Far NIR sm1		0.213	99.54	53.93	5	45.11	0.048
Far NIR sm2		0.136	99.52	23.78	3	25.12	0.169
Far NIR sm3		0.116	99.48	19.05	2	17.89	0.257
Far NIR sm4		0.115	99.59	18.90	2	19.12	0.239
Near NIR raw	1003 - 1084	0.194	99.87	49.46	8	10.64	0.392
Near NIR sm1		0.105	98.32	7.93	1	3.45	0.632
Near NIR sm2		0.105	98.92	7.89	1	3.47	0.631
Near NIR sm3		0.185	100.00	47.09	9	17.56	0.262
Near NIR sm4		0.108	99.98	41.14	4	18.22	0.252
VNIR raw	478 - 610, 630 - 885	0.136	99.18	90.01	4	5.61	0.540
VNIR sm1		0.103	98.02	7.95	1	3.28	0.641
VNIR sm2		0.131	99.84	65.29	5	3.38	0.636
VNIR sm3		0.126	99.92	74.07	5	6.42	0.511
VNIR sm4		0.123	99.91	55.41	5	8.29	0.452

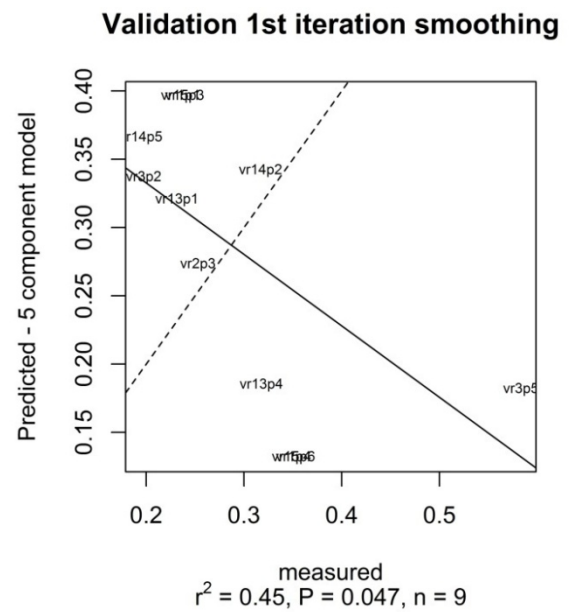
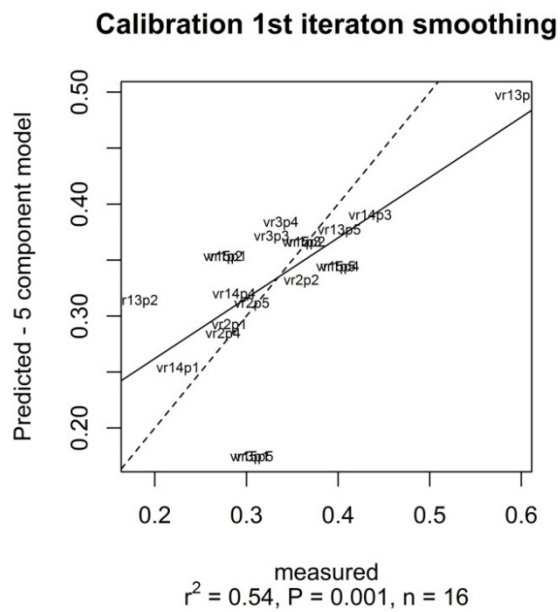
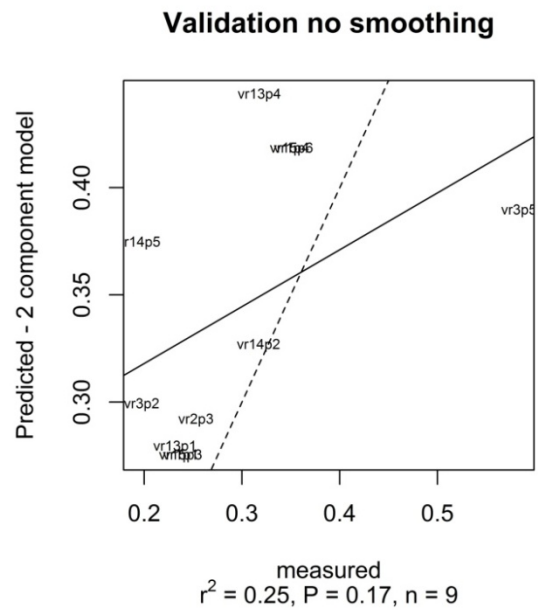
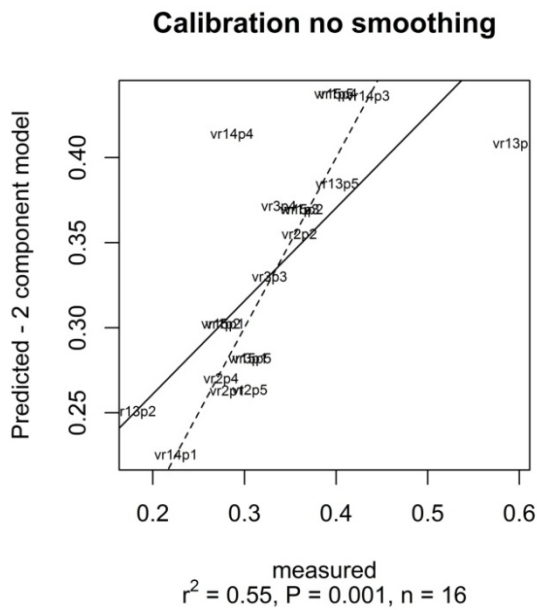
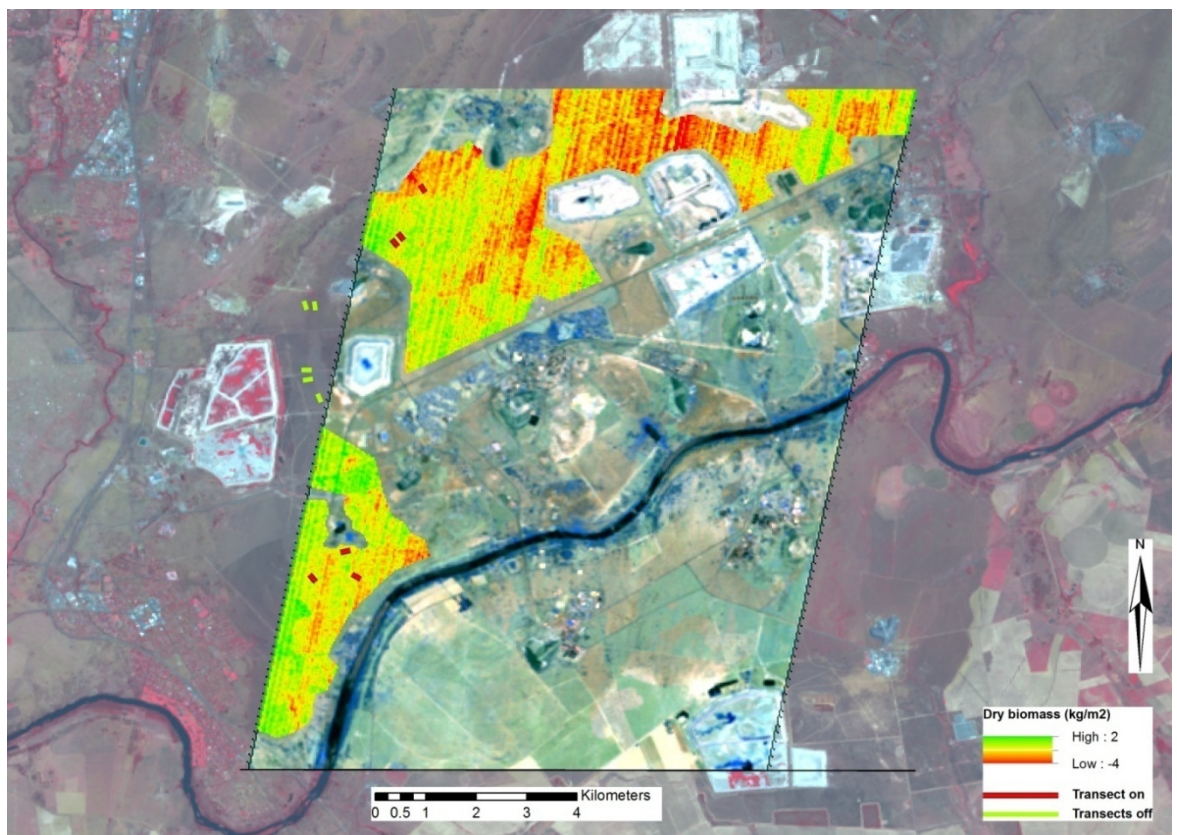
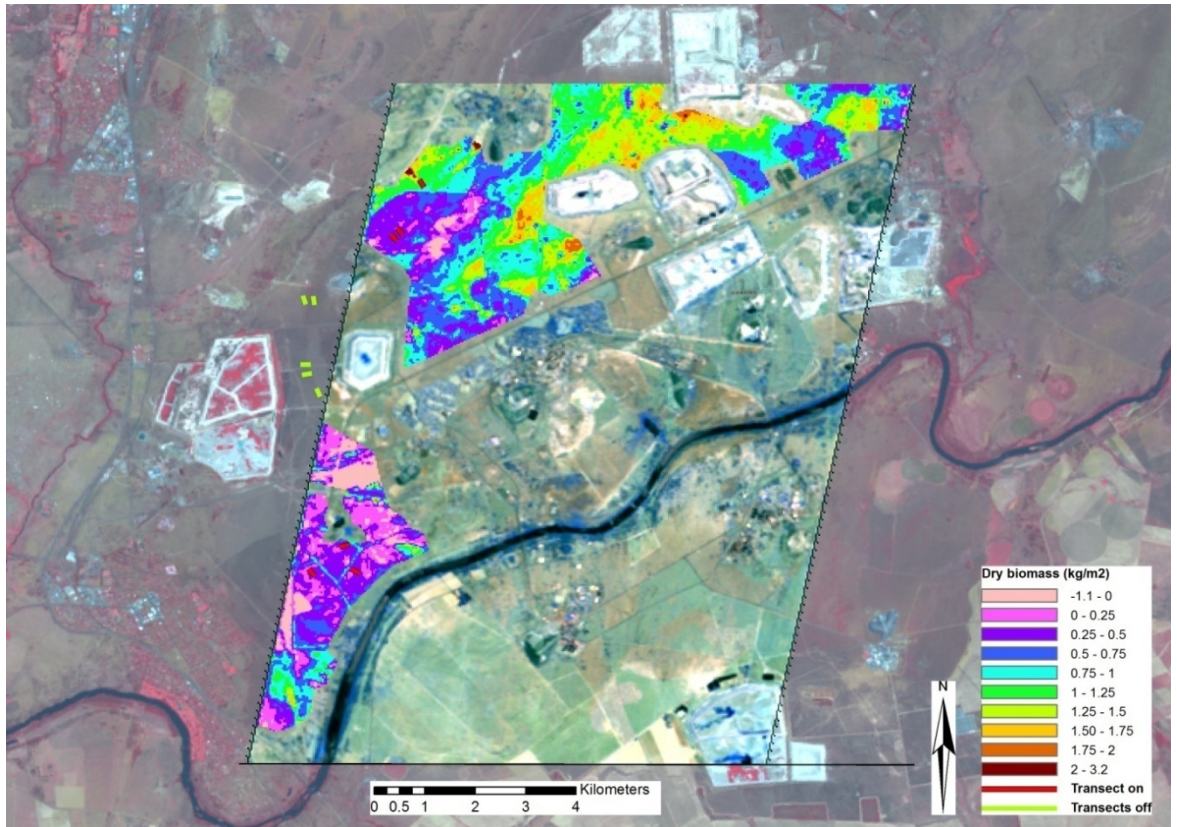


Figure 4.31 Measured versus PLSR predicted dry biomass (kg/m^2) using (a) 105 bands across the whole Hyperion spectrum and a 2-component model with no spectral smoothing, and (b) a 5-component PLSR model derived from 40 bands in the far NIR where the spectra were subjected to one iteration of the smoothing algorithm prior to PLSR modelling.



4698 **Figure 4.32** Dry biomass (kg/m^2) prediction maps for (a) a 105 band, full spectrum, 2-
 4699 component model with no spectral smoothing, and (b) a 5-component PLSR
 4700 model with 1 iteration of the smoothing algorithm on 13 far NIR bands.

4701 very badly striped such that features in the landscape were poorly discernible in the map
4702 (Figure 4.32b). The VNIR, normally a spectral block with clean spectra and fairly strong
4703 models had very weak validation statistics with either unsmoothed or smoothed spectra
4704 and generally required a number of components (4 or 5) to achieve the best validation
4705 model (Table 4.10). The near SWIR, which also has performed strongly with other variables,
4706 was weak in a PLSR model with unsmoothed data but the validation statistics for the best
4707 model (two components in all cases) got stronger with each iteration of the smoothing
4708 algorithm while the calibration statistics were weaker with each iteration (Table 4.10).

4709 4.4 Discussion

4710 4.4.1 The Hyperion Hyperspectral Image

4711 To put these results into context, various confounding factors that underlie this study need
4712 to be understood. The first of these is the time differential between the field work and the
4713 acquisition of the Hyperion image for Vaal River. Data acquisition requests to the USGS
4714 requested that the Hyperion image be acquired between June and September 2011 as this
4715 would coincide with the field work and the phenological phase of the grasslands. However,
4716 the Hyperion image was acquired on 2nd November 2013 which may have moved these
4717 grasslands into a different phenological state, i.e. an earlier summer growth phase, when
4718 the study was done over a winter dormant phase. Semi-arid grasslands are very sensitive to
4719 rainfall and require a certain amount of rain before soil moisture is high enough to induce
4720 breaking of dormancy and sprouting of grasses (Xin *et al.*, 2015, Archibald and Scholes,
4721 2007, Prins, 1988, Prins and Loth, 1988, Dye and Walker, 1987). It is unknown what rainfall
4722 fell in the area over the months of August, September and October of 2013 but spectra from
4723 the natural grasslands do not show much chlorophyll activity (Figure 4.5) when compared to
4724 the sports stadium field which has probably been irrigated (Figure 4.4). Nonetheless, this
4725 possible shift in grass phenology (Knox *et al.*, 2013, Butterfield and Malmström, 2009,
4726 Archibald and Scholes, 2007) does add to the complexity of interpreting the results.

4727 A greater effect of the mismatch between the Hyperion spectra and the field measurements
4728 is brought about by the two-year delay. Not only are grasslands sensitive to soil moisture,
4729 and hence rainfall, in the timing of their phenological phases, but biomass production is also
4730 positively related to mean annual precipitation (O'Connor *et al.*, 2001, Epstein *et al.*, 1996,
4731 Sala *et al.*, 1988) although this may be modulated by a number of factors such as grass
4732 species response (O'Connor *et al.*, 2001, Dye and Walker, 1987), temporal rainfall
4733 distribution (Dye and Walker, 1987) and soil nutrient status (Burke *et al.*, 1997). Higher or
4734 lower biomass would shift the ratio of soil surface versus canopy cover reflectance over the
4735 area of a 30 x 30 m pixel which could change the shape of the spectra, especially in the SWIR
4736 (Numata, 2012, Eva and Lambin, 1998a), from that which would have been recorded if the
4737 Hyperion spectra had been acquired during the field sampling period. If a lower biomass
4738 than that present in 2011 had resulted from a drier 2012/3, this would result in a greater

4739 soil surface reflectance predominating. However, if there had been more grass biomass
4740 produced during the growing season, this would result in greater grass canopy cover and
4741 soil features such as the clay absorption feature would be less exposed and more masked
4742 from the sensor. Once senescence and dormancy sets in, grasslands maintain a standing
4743 biomass (Zedler, 2007), although there is degradation of the standing biomass through
4744 breakage from trampling, grazing, rarely wind (Zedler, 2007), and photodegradation (Austin
4745 and Ballaré, 2010, Austin, 2006, Köchy and Wilson, 1997, Vossbrinck *et al.*, 1979). Biological
4746 soil crusts may also influence the spectra as any moisture would increase their chlorophyll
4747 content (O'Neill, 1994, in Weber *et al.*, 2008) and ground-based spectral sampling showed
4748 that these could produce a slight chlorophyll feature around 670 - 680 nm (Weber *et al.*,
4749 2008). However, it is unclear what effect these would have on the overall pixel spectra but
4750 their inclusion would depend on the degree of canopy cover.

4751 A third factor that confounds the PLSR modelling due to the delay in providing the Hyperion
4752 image is the effect of landscape processes and disturbance. There is no evidence that there
4753 have been major changes in landscape processes between the field sampling in 2011 and
4754 image acquisition in 2013 but there is definite evidence of disturbance, particularly from
4755 fire. Of the five transects that fell within the Hyperion image, three transects, namely vr1,
4756 vr2, and vr3, in the southwestern grassland block show no signs of fire over the 2013 winter
4757 senescent period which is also the predominant fire season. However, transects vr13 and
4758 vr14 were burnt between July 7th and August 16th 2013 giving a minimum of 11 weeks to
4759 recover, while the nearby transect, vr15 was not burnt during 2013. Snyman (2004) found
4760 that grass density, in Highveld grasslands with slightly higher annual rainfall than that at
4761 Vaal River, was reduced for two dominant species four months after an August fire. One
4762 species with low density before the fire, increased its density after the fire, and there was
4763 no change in density for other grass species. These changes in density were still statistically
4764 significant two years after the fire (Snyman, 2004). Furthermore, he found that burnt areas
4765 produced half as much above-ground biomass as unburnt areas did after the first growing
4766 season (Snyman, 2004).

4767 Thus it can be assumed that the burnt areas would have significantly different and, most
4768 likely, lower levels of biomass and canopy cover in 2013 compared to the field
4769 measurements collected in the winter of 2011. Furthermore, what biomass that was present

on these two burnt transects when the Hyperion image was acquired in 2013 was most likely to be in a green photosynthetic state as all the standing senescent material would have been burnt off. However, from the Hyperion spectra extracted for the transects (Figure 4.5), it is not clear that vr13 and vr14 had much photosynthetically active biomass compared to vr1, vr2 and vr3 which were not burnt, and hence still have their standing non-photosynthetic biomass mixed with any new photosynthetic material produced in the spring flush. However, vr1, 2 and 3 have a definite peak in the SWIR region around 2200 nm, which is characteristic of vegetation and caused by lignin/cellulose absorption at 2100 and 2300 nm which is absent in soil spectra (Numata, 2012, Datt *et al.*, 2003). This peak is absent in vr13 and vr14 suggesting that eleven weeks after burning, vr13 and vr14 have very little vegetation present. Transects vr13 and vr14 also have a higher reflectance across the whole spectrum than do vr1, vr2 and vr3 and this may be due to soils often having a brighter spectrum (Marsett *et al.*, 2006, Huete and Jackson, 1987, Heilman and Boyd, 1986). This is supported by transect vr15 which had the brightest set of spectra for all pixels (Figure 4.5) and was slightly northeast of transects vr13 and vr14. Transect vr15 was not burnt in 2013 but in 2011 it was heavily grazed almost to soil level. Cattle were observed every afternoon on this portion of the grassland with the cattle pen situated to the north east of transect vr15. It can therefore be concluded that overgrazing was very likely to be the situation in both 2011 and 2013 and the slightly lower brightness of vr13 and vr14 compared to vr15 may be due to residual ash and charcoal (Eva and Lambin, 1998a, Roberts *et al.*, 1997) which would have been present on vr13 and vr14 but not on vr15 during 2013.

Another factor which undermined the study was the shift in the centroid of the Hyperion image from that requested. The centroid of the Hyperion image was shifted to the east by 4996.7 m and to the north by 3650.2 m from the coordinates supplied in the DAR. The shift to the north is not an issue but the shift to the east meant 5 transects (or approximately 30 pixels) were outside the Hyperion image and could not be part of the PLSR modelling. This had a number of negative effects. First, it reduced the number of pixels from a possible 50 (or 56 if we include LOI and PAI from transect vr15 which was burnt down by cattle herders halfway through sampling) to 25 (or 31) for most indices. The 25 (31) pixels available had to be split on a 2:1 ratio between calibration (16 or 20) and validation (9 or 11) data sets and such low replicates could have substantially weakened the PLSR models.

4801 Second, in shifting the image to the east, the northern and southern grasslands became
4802 disconnected. The experimental design was premised on sampling the slopes following a
4803 small drainage basin from close to the Black Reef, which forms a ridge and watershed
4804 boundary northwest and north of transects vr10 to vr15, down between two tailings storage
4805 facilities (TSF) where transects vr7, vr8 and vr9 were located, and onto the floodplains
4806 where vr1, vr2 and vr3 are situated and where the drainage basin seeps into the Vaal River.
4807 Transects vr4, vr5 and vr6, located just below the tailings dams, as well as vr12 located
4808 above the tailings dams, were burnt before they could be sampled. Thus 15 transects were
4809 planned and marked out at Vaal River in five sets of three to provide a high density
4810 assessment from likely “low” polluted sites high up the drainage basin, to “high” polluted
4811 sites between and below the TSFs. However, the position of the Hyperion band cuts the
4812 drainage basin into disconnected northern and southwestern grasslands and removes the
4813 main source of pollution which is the section between the two TSFs. This also removed the
4814 continuity expected in spectra from pixels following the drainage basin that would have
4815 been available to the PLSR model. Furthermore, the position of available grassland for PLSR
4816 model prediction now extends outside the sampled drainage basin in the northern block
4817 into grasslands to the east where no sampling was done and conditions may be different to
4818 those where sampling was done. Therefore, predicted values derived from the PLSR models
4819 in the northern grassland should be regarded as having a lower probability of being accurate
4820 as one moves east. Lastly, this extension into unsampled and possibly different grasslands
4821 also undermines the assumption that values predicted from the Hyperion data cube for the
4822 grasslands should have a similar distribution as values measured in the field which was used
4823 as one test of the accuracy of predicted values.

4824 Finally, there is the issue of the quality of the hyperspectral image itself. The Hyperion
4825 sensor is a pushbroom sensor in which a dichromic filter reflects VNIR light to sensors
4826 sensitive to the 400 to 1000 nm range while transmitting the region from 900 – 2500 nm to
4827 a SWIR detector (Pearlman *et al.*, 2001). Early radiometric calibration tests reported that the
4828 VNIR was within 5% of expectation but the SWIR was not consistent above 1100 nm (Barry
4829 *et al.*, 2002) yet the same report indicated that the Hyperion instrument had excellent
4830 radiometric quality. The Hyperion data for Vaal River had two types of artifact in the data
4831 cube. The first is vertical striping within a band (x y dimensions), and the second is spectral

noise (z dimension). Striping may be caused by a poorly calibrated detector, which in the most extreme case may be referred to as a bad pixel and have no valid data (Datt *et al.*, 2003) or it may be due to a detector calibration varying spatially or temporally either from within array interactions or the read-out process (Rogass *et al.*, 2014, Datt *et al.*, 2003). Either way, the end result is a vertical stripe or stripes running in the direction of the satellite track. These are most prevalent in the SWIR region but also occur in the VNIR region (Datt *et al.*, 2003). Destriping (Rogass *et al.*, 2014, Scheffler and Karrasch, 2013, Datt *et al.*, 2003), together with other correction issues such as desmiling (Goodenough *et al.*, 2003) and keystone correction (Yokoya *et al.*, 2010), should be done before atmospheric correction, georeferencing and orthorectification (Scheffler and Karrasch, 2013). As the Hyperion image came from the USGS already georectified, it was decided in this initial study to apply the FLAASH correction and then remove the bands with excessive striping. However, even when there was only one major stripe in a band, the removal of that band did not guarantee the removal of the stripe from the PLSR prediction map. Often adjacent bands one or two band position either side of the striped band had to be removed as well, even though none of these adjacent bands had any visible striping, as the PLSR models seemed to be sensitive enough to multiply subtle impressions of the stripe in the adjacent bands into a stripe in the prediction map. Stripes with this adjacent band association were always off transect pixels and therefore not involved in the calibration and validation models but did feature in the prediction maps.

The second kind of noise is spectral noise and according to Datt (2003) has two causes. The first is processing and sensor noise and the second is related to the atmospheric correction process through differences in the modelled and real atmosphere at the time of image acquisition (Datt *et al.*, 2003). Various methods are available to correct the atmospheric effects, one of which is to fit the Hyperion spectra to known reference spectra acquired on the ground, but due to the time difference between field sampling and Hyperion image acquisition this was not considered feasible for the Vaal River Hyperion image. Therefore, a smoothing algorithm was applied (Bertels, 2012) which calculates a weighted mean whereby the central or desired wavelength has twice the weight of its neighbor wavelengths. Increasing the number of iterations of the smoothing algorithm increased the smoothing effect (Figure 4.4) and data cubes with no smoothing or with 1 to 4 iterations of

the smoothing algorithm were tested in this study. The end and beginning of each section of spectra uses a triple weighting for the end value as there is no neighbor on one side (Bertels, 2012). However, these end spectra were found to be quite different to the original spectral value (Figure 4.4) and so together with the excessively striped bands, these end bands on the spectral blocks were removed from the data cube prior to PLSR modelling. It should be noted though that smoothing in this way also smooths out the more subtle absorption features that may be present in the spectra.

4.4.2 **The LFA Indices: Stability, Infiltration and Nutrient Cycling**

The strongest PLSR models for all three indices with Hyperion data came from the VNIR, with the near SWIR also have moderately strong linear relationships with the three LFA indices. The strongest PLSR model predicting any of the three LFA indices from Hyperion hyperspectral data was a 4-component model predicting the nutrient cycling index (validation $r^2 = 0.70$, $P = 0.005$, $n = 9$; Table 4.6) with four iterations of the smoothing algorithm. This model was only slightly stronger than the model with the same bands and no spectral smoothing and both produced fairly clean prediction maps lacking the speckling and striping apparent in some other models (Figure 4.20). The LFA infiltration index also produced a fairly strong models with 35 VNIR bands in both unsmoothed and smoothed forms (validation $r^2 = 0.66$, $P = 0.008$, $n = 9$, 4th iteration smoothing; Table 4.5) with a quality prediction map although the model with smoothed spectra (Figure 4.17b) showed slightly less speckling than the unsmoothed spectra (Figure 4.17a). The LFA stability index produced fairly strong models statistically (unsmoothed 5-component model validation: $r^2 = 0.72$, $P = 0.004$, $n = 9$; Table 4.2) in the VNIR but prediction maps from these models were of poor quality with serious speckling obscuring features present in the landscape (Figure 4.10). Smoothing the spectra did remove some of this speckling (Figure 4.11) but also weakened the prediction statistics (4-component validation model: $r^2 = 0.39$, $P = 0.074$; Table 4.2).

The PLSR modelling results with Hyperion data, to some extent, mirror those modelling the LFA indices from ground data (results presented in Chapter 3) except that the models predicting the LFA infiltration index was the strongest model at both West Wits (5-component model validation: $r^2 = 0.75$, $P < 0.0001$, $n = 19$; Table 3.11) and Vaal River (12-component model validation: $r^2 = 0.72$, $P < 0.0001$, $n = 59$; Table 3.12). Models predicting

the nutrient cycling index from ground-based spectra were slightly mixed with the West Wits producing stronger validation results (6-component model validation $r^2 = 0.73$, $P < 0.0001$, $n = 19$, Table 3.13) compared to Vaal River (4-components validation: $r^2 = 0.54$, $P < 0.0001$, $n = 59$; Table 3.14). The LFA stability index produced the weakest models of the three process-based LFA indices for each site (West Wits 5-component validation model $r^2 = 0.64$, $P < 0.0001$, $n = 19$; Table 3.9; and Vaal River 4-component validation model $r^2 = 0.39$, $P < 0.0001$, $n = 59$; Table 3.10). Generally the results of these PLSR models for the LFA infiltration and nutrient cycling indices with both Hyperion-derived spectra and ground-based spectra are considerably stronger than those obtained with ground-based spectra in the pilot study (Furniss *et al.*, 2009, Furniss, 2008). This supports the argument outlined at the beginning of this study that the three LFA indices of stability, infiltration and nutrient cycling needed to be considered in the context of the LFA methodology, and not adapted to the spectral measurements, as was done in the pilot study (Furniss *et al.*, 2009, Furniss, 2008).

The stability index produced strong model statistics with Hyperion imagery (unsmoothed validation $r^2 = 0.72$, $P = 0.004$, $n = 9$; Table 4.2) but the prediction maps were of poor quality (Figure 4.10). West Wits also produced a fairly strong PLSR model predicting the LFA soil stability index with ground spectra (validation 4-component model $r^2 = 0.63$, $P < 0.0001$, $n = 19$; Table 3.9), whereas the ground-based stability model at Vaal River was weak (validation 4-component model $r^2 = 0.39$, $P < 0.0001$, $n = 59$; Table 3.10). It seems that this weakness in the Vaal River model derived from ground spectra is due to the wide spread in predicted values for a specific measured value, where these predicted values tend to organize themselves on the basis of patch type (Figure 3.21). In other words, the model is predicting different values from differences in spectra due to different patch types. However, the low range of measured values at Vaal River means there is a large overlap in measured values between patch types at Vaal River and this is undermining the PLSR model to accurately predict LFA stability values. In contrast, values measured at West Wits sort more clearly over the range into patch types. Thus bare patches have low LFA stability values and grass and *S. plumosum* patch types have high values allowing PLSR models to more accurately predict values (Figure 3.21).

In this study, the ground-based and satellite-based spectral data were modelled separately such that in the Vaal River models, pixel spectra corresponding to transects were extracted from the Hyperion data cube. The LFA transects were partitioned to match their corresponding pixels and the LFA indices recalculated for these pixel-based transects before paired pixel spectra and LFA data were split into calibration and validation data sets. The calibration data was then used to generate the model and the validation data used to test the selected models before applying the best model to the whole Hyperion data cube to generate predictive maps. In Australia, Ong *et al.* (2008) used a slightly different approach where they calibrated their HyMap airborne data to ground-based spectral measurements of light and dark invariant targets and then derived their PLSR models from ground-based measurements before applying the selected models to airborne data to produce prediction maps. At Goldsworthy decommissioned iron ore mine, PLSR models with airborne data predicting the LFA stability, infiltration and nutrient cycling indices achieved coefficients of determination of 0.53, 0.59 and 0.66 respectively, and at the Huntly Bauxite Mine they were 0.64, 0.46 and 0.88 respectively (Ong *et al.*, 2008). Despite a pixel size difference at ground level of 29.9 x 29.9 m for Hyperion imagery versus 3 x 3 m for the HyMap imagery, the Hyperion results predicting the same indices in this study compare favourably with those described by Ong *et al.* (2008).

In a related study by the same group (Lau *et al.*, 2008) in Huntly Bauxite Mine, they reported coefficients of determination of 0.76, 0.67 and 0.71 for the LFA stability, infiltration and nutrient cycling indices which are comparable to the Vaal River values. However, they quote RMSE values for these model validations of 16.87, 60.97 and 41.77 respectively (Lau *et al.*, 2008), although it is unclear if these are calibration or validation root mean square errors. In the Vaal River study, calibration RMSEP values of 6.38, 2.09 and 3.76 were calculated for stability, infiltration and nutrient cycling respectively, which are considerably smaller than those quoted by Lau (2008). This discrepancy is probably related to the range of LFA index values at the two sites which were much larger for the Lau (2008) study (35 – 85, 20 – 70 and 10 – 60% for stability, infiltration and nutrient cycling respectively) compared to the Vaal River site (57 – 69, 26 – 36 and 15 – 29% for stability, infiltration and nutrient cycling). This much larger range is due to fundamental differences between the two sites where the Huntly Bauxite Mine has a much higher MAP (700 – 1300 mm) and consists of eucalypt

4954 woodland from recently rehabilitated sites to >30 year-old rehabilitation and includes
4955 minimally disturbed natural woodland (Lau *et al.*, 2008), whereas, Vaal River with a MAP of
4956 560 mm is semi-arid grassland and the areas sampled were natural relict grasslands within
4957 the mining environment. The Lau, *et al.* (2008) study also found good correlation between
4958 LFA indices and age of rehabilitation with old rehabilitated areas having high values and
4959 young rehabilitated areas having lower values. The Huntly study site experienced fire in
4960 certain parts of the mine landscape in the year before sampling and were able to
4961 incorporate this into the PLSR modelling and found that these sites had lower LFA index
4962 values than did comparable sites (Lau *et al.*, 2008). At Vaal River, fire occurred two years
4963 after the field sampling and three months before the Hyperion image acquisition and
4964 therefore becomes a confounding variable.

4965 Two other studies have used remote sensing data to predict LFA indices (Gaitán *et al.*, 2013,
4966 García-Gómez and Maestre, 2011), but both used simple linear regression rather than PLSR,
4967 and multispectral satellite imagery rather than hyperspectral imagery. Gaitán *et al.* (2013)
4968 sampled 194 sites across 15 degrees of latitude in Argentina. They randomly selected 154
4969 plots to derive linear regressions predicting the LFA indices and the remaining 40 plots to
4970 validate the models, and repeated the random selection 300 times. They didn't directly use
4971 reflectance in their regressions but calculated a number of vegetation indices (VI) such as
4972 the normalized difference vegetation index (NDVI) which they calculated from average
4973 reflectance over the growing season using MODIS Land subsets (MOD13Q1 product) with a
4974 pixel size of approximately 250 x 250 m (Gaitán *et al.*, 2013). The NDVI gave the best results
4975 when predicting the LFA stability, infiltration and nutrient cycling indices with mean
4976 coefficients of determination of 0.15, 0.38 and 0.50 respectively for validation data (Gaitán
4977 *et al.*, 2013). García-Gómez and Maestre (2011) working in the semi-arid steppes of central
4978 Spain also used the NDVI calculated from a single ASTER image with a pixel resolution of
4979 approximately 15 x 15 m. They also used linear regression and randomly selected 30 plots
4980 for calibration with the remaining 10 used in model validation, and repeated this 300 times
4981 (García-Gómez and Maestre, 2011). Their mean coefficients of determination for validation
4982 were 0.66, 0.75 and 0.53 for LFA stability, infiltration and nutrient cycling, with all tests
4983 significant at the 5% level (García-Gómez and Maestre, 2011). The generally stronger results
4984 of García-Gómez and Maestre (2011), which are comparable with those obtained from

hyperspectral data and PLSR modelling (Furniss *et al.*, 2009, Lau *et al.*, 2008, Ong *et al.*, 2008), compared to Gaitán *et al.* (2013) are probably partially related to a much smaller pixel size (15 m versus 250 m respectively) whereby sample plots are likely to have less variability in the smaller pixels (Gaitán *et al.*, 2013). Furthermore, these two results, together with the results from Vaal River, Ong *et al.* (2008) and Lau *et al.* (2008), support Gaitán *et al.* (2013) in asserting that higher spectral resolution may improve predictions of LFA indices from spectral imagery.

Often regression coefficients and loadings can be examined and interpreted to identify features in the environment that have contributed to the PLSR models (Ong *et al.*, 2008, Mevik and Wehrens, 2007). The features in the loadings from the West Wits and Vaal River PLSR models with ground-based spectra predicting the LFA infiltration index show a mix of both soil and plant features (Figure 3.25). Furthermore, the correlation between the loadings and mean spectra for each patch type showed that in each model, the first component was dominated by soil characteristics while the second loading was dominated by vegetation related spectral features (Figure 3.26). Mayor and Bautista (2012) examined the relationship between actual infiltration and run-off and the LFA infiltration index using rainfall simulation and sediment concentration at the scale of patches and inter-patches, and at hillslope and catchment scales, they used records of natural rainfall and runoff and sediment production collected over four years. They found that at the patch/inter-patch scale, LFA infiltration values correctly correlated with patches and inter-patches but did not adequately characterize the hydrological variation within patches and inter-patches (Mayor and Bautista, 2012). At the hillslope and catchment scale, they found that there was an inverse relationship between run-off and the infiltration index for bare patches, otherwise often referred to as inter-patches, but not with the global infiltration index which is calculated as a weighted sum of mean infiltration for each patch type where the weights are based on the proportion of transect occupied by each patch type (Mayor and Bautista, 2012). Their results are to some extent supported by the results from West Wits and Vaal River where the first and third loadings for the ground-based infiltration models showed a strong inverse correlation at the patch/inter-patch scale with bare patches/biological soil crust/bare grass patch types (Figure 3.26). These are all patch types that are principally related to exposed soil surface. The first loading in the Vaal River infiltration model

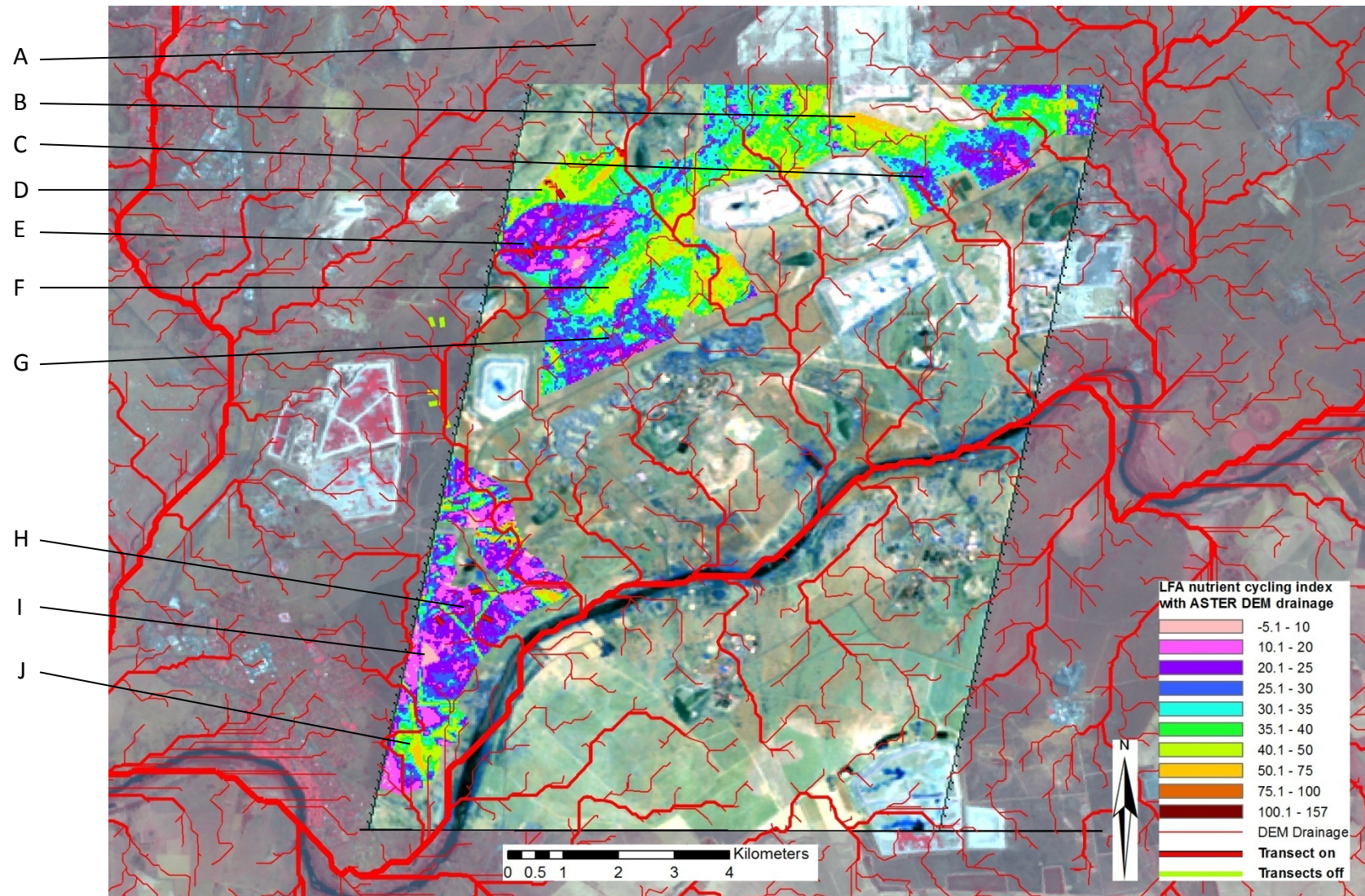
accounted for 64.8% of the variability in the spectra, while the second loading accounted for 31.5% of the remaining spectral variability in the model (Figure 3.25). The second loading is slightly positively correlated with vegetation and slightly negatively correlated with soil-dominated patch types (Figure 3.26).

For the LFA indices of stability, infiltration and nutrient cycling predicted from Hyperion spectra, prediction maps with the best quality, as in clearness of landscape features and reduced speckling or stripes (Rogass *et al.*, 2014, Scheffler and Karrasch, 2014, Scheffler and Karrasch, 2013, Datt *et al.*, 2003), came from the VNIR portion of the Hyperion spectrum between 478 and 885 nm. However, despite the PLSR prediction maps being generally of a high visual quality, regression coefficients and loadings were difficult to interpret for absorption features that may be involved in the models because the spectral noise appears in the first component. Smoothing progressively dampens the noise with each iteration but does not remove it. At the same time though, smoothing also dampens any real signals in the spectra (Bertels, 2012) and therefore also the regression coefficients and loadings. To complicate interpretation of the regression coefficients and loadings, the grasslands were in a senesced form with no pigments and minimal liquid water content which are the spectrally active features in vegetation used to discriminate green vegetation from soil spectra (Roberts *et al.*, 1997, Elvidge, 1990). The process of senescence withdraws moisture and pigments and leads to a decrease in NIR-to-red ratios in senesced vegetation and differences between soil and senesced grassland canopy spectra may be very small and difficult to separate (Nagler *et al.*, 2000, Roberts *et al.*, 1997, Roberts *et al.*, 1993). This can be seen in Figure 3.14 where biological soil crust patches, bare grass patches and grass patches have very similar mean spectra in the VNIR and PLSR patch models found it difficult to separate these three patch types.

An examination of the first loading, which explained between 97 and 98 percent of the spectral variability in all three models, shows no strong absorption features for the PLSR stability model (Figure 4.13, loadings for infiltration and nutrient cycling models not shown) although there was progressively more influence on the PLSR models from blue-green wavelengths (500 nm) to NIR wavelengths (850 nm). However, the nature of this influence varied between models with the first loading for the LFA stability and nutrient cycling PLSR models getting more positive towards the 850 nm position, and LFA infiltration model

getting more negative. As the first component of the PLSR models generally showed a weak linear relationship with values for the three indices (stability $r^2 = 0.15$, infiltration $r^2 = 0.01$, nutrient cycling $r^2 = 0.15$), the explanatory power of the first loading in relation to the indices is also weak and especially so for the infiltration index. The second loading for all three indices showed broad absorption features around 750 and 800 nm which are quite well developed. However, it is not clear what may be causing these as both senesced grass and soils are usually relatively featureless in these areas (Nagler *et al.*, 2000). Soils have a broad iron oxide absorption feature around 830 nm (Huete and Escadafal, 1991) and the mean spectra for patch types from the ground-based spectral sampling having an inflection point around 730 - 740 nm in grass-based patches, possibly related to a very weakly developed red-edge (Figure 3.14) which may be contributing to the 750 and 800 nm features in the loadings. The profile of the third loading in the infiltration and nutrient cycling model (not shown) have absorption features around 650 and 730 nm that more closely match the iron oxide soil feature (Huete and Escadafal, 1991) and the dry grass inflection point suggesting that this loadings may be a mixture of soil and senescent grass features. This is supported by the fact that in all three LFA PLSR models, this third loading accounts for the greatest proportion of the explained variation in the response variable in these models (stability $r^2 = 0.22$, infiltration $r^2 = 0.30$ and nutrient $r^2 = 0.35$ for the third loading). So overall, it remains unclear what features in the landscape are contributing to these models derived from Hyperion spectra. Furthermore, because the pixel size is 30 by 30 m and the patch sizes are considerably smaller, the difference in scale between the pixel size and patch sizes meant it was not possible to determine correlations between loadings and patch types composing pixel-scaled transects.

Interpreting the prediction maps for the three LFA indices of stability, infiltration and nutrient cycling are made more complex by the confounding factors discussed in section 4.4.1. The terrain at Vaal River is gently inclined plain with long gradual slopes and low flat ridges with mine infrastructure such as waste rock piles and tailings storage facilities generally towering of the landscape. Therefore to get some idea of the topography, Figure 4.33 presents a Landsat false infrared image with the Hyperion SWIR image as background to the nutrient cycling prediction map overlaid with a theoretical drainage derived from an



5077

5078

5079

Figure 4.33 Drainage (red) derived from an ASTER GDEM overlaid on the prediction map for the nutrient cycling index derived from a 4-component PLSR model using forty VNIR bands with four iterations of the smoothing algorithm.

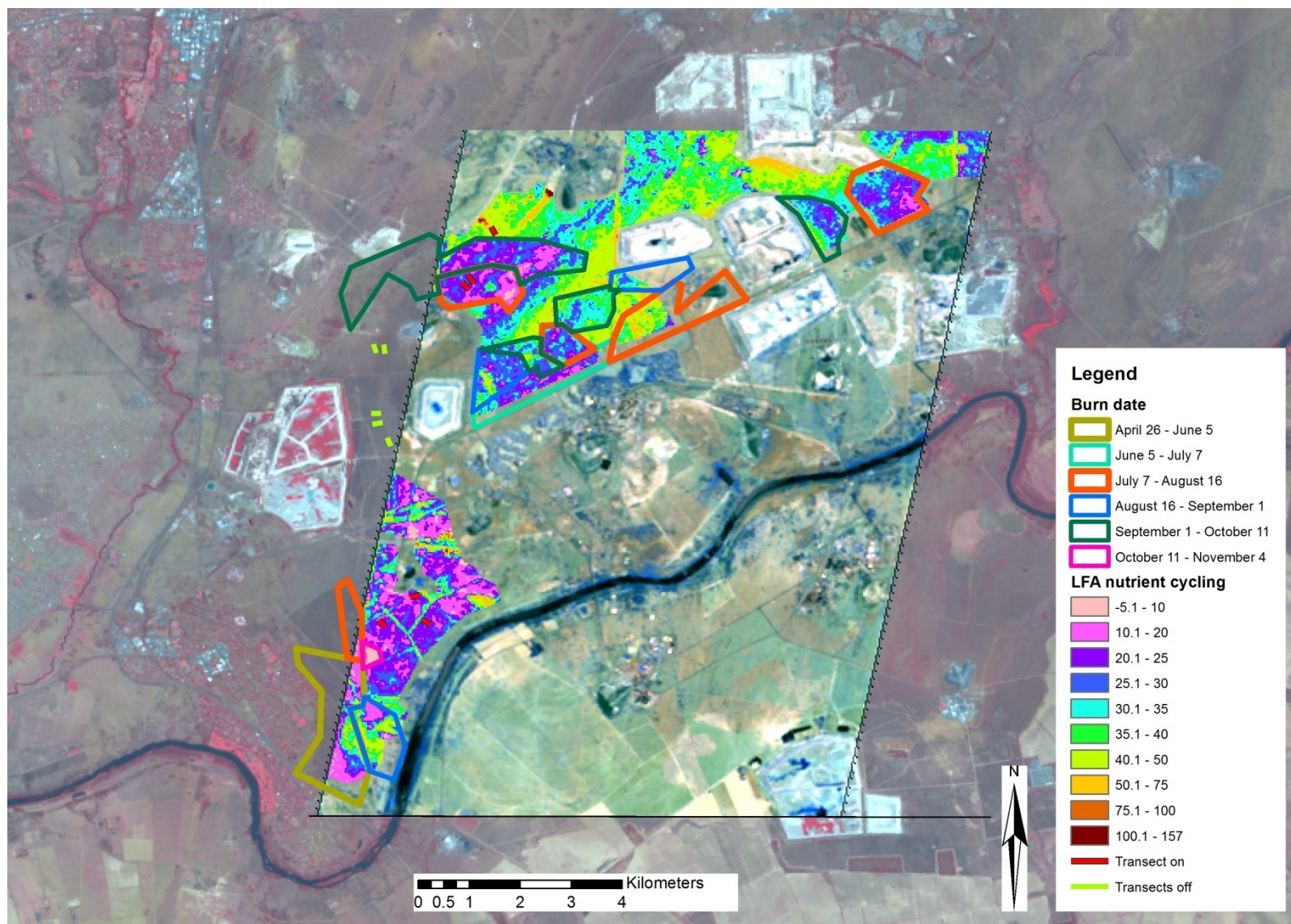
ASTER GDEM (Meti and NASA, 2009). It can be seen that this derived drainage is not 100% accurate, especially along the Vaal River which is not surprising considering the DEM is derived from stereoscopic ASTER images and tall trees lining the banks, together with old river beds nearby may be confounding the DEM derivation process. Furthermore, it should also be noted that this is an environment subjected to extensive engineering practices some of which are involved with controlling runoff such as toe paddocks, canals and return water dams and these were not taken into account when deriving the drainage. Nonetheless it does provide some indication of where high ground or low lying ephemeral drainage lines are. Also apparent is the ridge formed by the black reef which forms the main watershed north of the Vaal River and shows as a dark wavy line labeled A in Figure 4.33.

The ranges for LFA field-measured indices corresponding with pixels are 57.2 – 69.1%, 26.8 – 36.1% and 15.0 – 29.4% for LFA stability, infiltration and nutrient cycling respectively which corresponds well with predictions for the areas around transects (Labels E and H in Figure 4.33) where the range was between 55.1 – 80.0% (Figure 4.11), 25.1 – 35.0% (Figure 4.17b), and 10.1 – 30.0% (Figure 4.20b and 4.31) respectively. The area labeled F in Figure 4.33 has rather high prediction values for nutrient cycling (35 – 55%). This area is slightly higher ground forming a low plateau in the area although altitudinal change is quite small with the black reef ridge at approximately 1340 MASL, transects vr13 and vr14 at ± 1326 MASL, transect vr15 at 1325 m and the high ground at E at ± 1330 MASL. The factors regulating nutrient cycling are extremely complex but a general trend for nutrients is that the crest and top of slopes tend to have lower values and the bottom of slopes higher values (Rezaei *et al.*, 2006, Verchot *et al.*, 2002, Wang *et al.*, 2001) as resources generally move down slope due to gravity and run-off, and accumulate at the slope foot. The LFA infiltration index for this area likewise showed higher values ranging from 40 – 60% (Figure 4.17b) compared to transect vr13 and vr14. Despite the differences in LFA infiltration index, the very shallow slopes and essential flat nature of Vaal River mean infiltration is much more likely to be influenced by local factors (Bergkamp, 1998) rather than very gradual changes in topography. The predicted LFA stability index though, is going against the trend in the other two indices with lower values in the upper slope position (25 – 55%; Figure 4.11) compared to the higher values for the other two indices. This is also against the trend generally found for LFA indices where the three indices generally increase or decrease together indicating

5111 higher or low ecological functionality for a site. This area of anomalous values for the three
5112 LFA indices of stability, infiltration and nutrient cycling extends from the area indicated at E
5113 in a northeast direction, through a drainage area and north and west of a tailings storage
5114 complex (Figure 4.33) implying that slope position is not the cause. However, this zone does
5115 lie parallel with the Black Reef and so possibly this pattern is indicating an edaphic and/or
5116 underlying geological pattern in the landscape which is being mirrored in the predicted LFA
5117 values.

5118 An analysis of the fire history of the prediction area for 2013 may provide some more insight
5119 into the mapped LFA values (Figure 4.34). It can be seen from Figure 4.34 that nine separate
5120 fires burnt through the areas labeled C and G between 5th June and 11th October 2013.
5121 Many of these fire scars show similar values for nutrient cycling as those predicted for the
5122 area associated with transect vr13 and vr14 (Figure 4.34) which was burnt three months
5123 prior to the Hyperion image acquisition. While for areas outside the burn areas, the LFA
5124 nutrient cycling values are higher than the range of measured values. A similar pattern is
5125 observed in the prediction maps for the LFA stability (Figure 4.11) and infiltration indices
5126 (Figure 4.17b). This suggests that the presence or absence of fire in the northern prediction
5127 block is having a major effect on prediction values for all three indices in this block and
5128 raises the question of how accurate the predicted values may be in this block. It should be
5129 noted though that the prediction values in the fire blocks don't show a consistent pattern in
5130 terms of date of burn, or prediction values for the three indices across the individual burns.
5131 This is consistent with conclusions drawn by other researchers that fire scars can have a
5132 large spectral diversity (Smith and McDermid, 2014) through variation in the type of fire, its
5133 severity, the recovery of burnt vegetation and the time since burn of satellite overpass
5134 (Smith and McDermid, 2014, Bastarrika *et al.*, 2011, Picotte and Robertson, 2011, Chuvieco
5135 *et al.*, 2002). Another feature of the northern prediction block is the very high values
5136 recorded for the LFA infiltration and nutrient cycling indices in the area labeled B and D
5137 (Figure 4.33) which are indicative of tailings spillages.

5138 The area of the transects (vr1, vr2 and vr3) in the southern prediction block, labeled H in
5139 Figure 4.33, was not burnt and only has burn scars on its southern tip and western border
5140 (Figure 4.34). Apart from the LFA stability prediction map (Figure 4.17) where the PLSR
5141 model struggled to produce a quality prediction map, the maps for the LFA infiltration and



5142

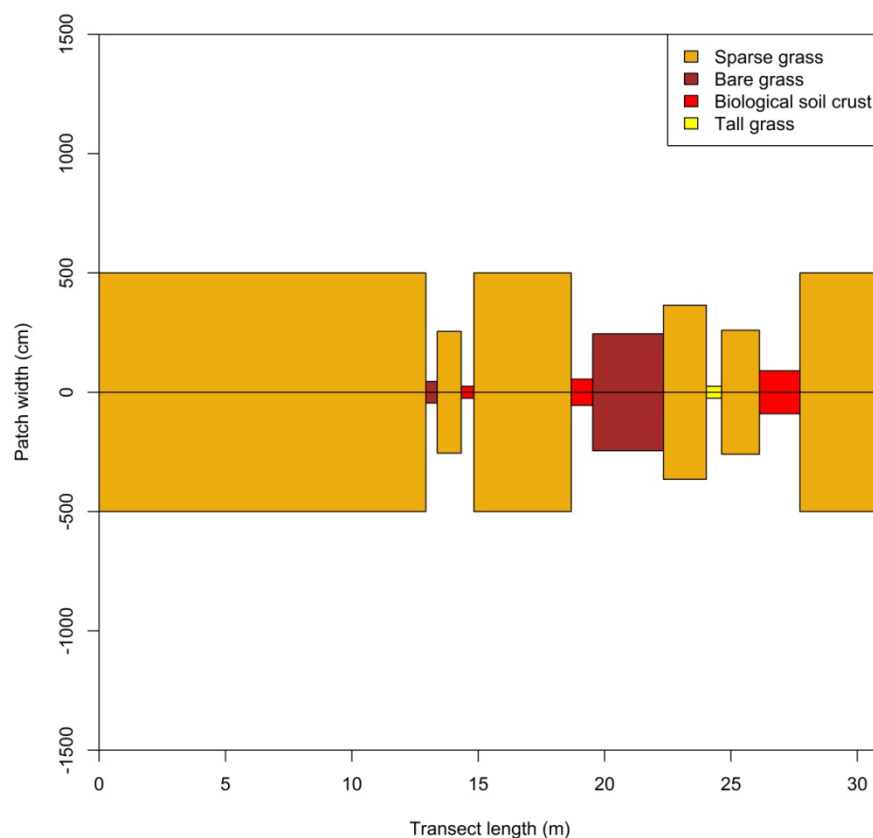
5143 **Figure 4.34** Occurrence and extent of fire during 2013 in the PLSR prediction area laid over the nutrient cycling prediction map.

5144 nutrient cycling indices (Figure 4.17 and 4.20b respectively) show clear separation between
5145 dirt roads and grassland. Furthermore, the predicted values in the grassland areas are within
5146 the range of those measured in the field. Among the burnt areas, one small burn stands out
5147 as a homogenous pink triangle (label I, Figure 4.33) which was burnt very late, between
5148 October 11th and November 4th and shows limited variability in the predictions for these
5149 pixels, presumably because it was burnt so close to the acquisition of the Hyperion image
5150 that the fire area was spectrally homogenous.

5151 The area labeled J (Figure 4.33) seems to illustrate how fire has had an impact on the PLSR
5152 modelling and subsequent predictions. During the field campaign in 2011, all measurements
5153 were in unburnt grassland so in deriving the models, the response variable has been a
5154 measurement of unburnt grassland. However, due to a reduction in pixels available, spectra
5155 were used from both burnt and unburnt grassland to derive the Hyperion models. Thus the
5156 values predicted for the area labeled J in the southern grassland block, which was burnt,
5157 reflect unburnt areas predicted for the northern grassland block (Figure 3.32). Furthermore,
5158 unburnt pixels in the southern block have similar values to burnt pixels in the northern
5159 block. To confound issues, some burnt areas in the southern block have similar values as
5160 unburnt areas in the southern block (Figure 4.34). Possibly, there are fundamental
5161 differences in soil moisture, or some other factor, between these two prediction areas and
5162 the PLSR models are confounding this with burnt and unburnt which are essentially the two
5163 dominant spectral classes in the pixels used in the PLSR models. If the Hyperion image had
5164 included all the transects measured and not left out the ones linking the northern and
5165 southern prediction areas, this might have been more apparent and less confounding.

5166 4.4.3 The LFA Landscape Organization and Patch Area Indices

5167 Of the five LFA indices predicted from PLSR models using Hyperion imagery in this study, the
 5168 landscape organization index (LOI) and the patch area index (PAI) are the only empirical
 5169 indices as they are calculated from measurements made with a tape measure (Figure 4.35)
 5170 rather than from observer decisions. The LOI is a one-dimensional measurement taking only
 5171 the length of patches along the transect, not including inter-patches (red polygons in Figure
 5172 4.35, biological soil crust patches), and the transect length (the central line in Figure 4.35)
 5173 into account. Whereas the PAI is a two dimensional measurement based on the total patch
 5174 area, not including inter-patches, as defined by the length of patch and the width of each
 5175 patch at the widest point, and the total transect area where the width in both cases is a
 5176 maximum of 5 m either side of the transect line (Figure 4.35). It does not consider the shape
 5177 and so all patches are essentially rectangular in shape. This difference in dimensions
 5178 underlying each index is presumably the main reason the PLSR models predicting these two



5179

5180 **Figure 4.35** The layout of patches along a transect in a pixel space (vr14 pixel 4) to show
 5181 the difference between the LOI and PAI. The real transect ran slightly
 5182 obliquely across the pixel but for diagrammatic purposes, the pixel and
 5183 transect have been aligned.

5184 indices had such divergent results. PLSR models predicting LOI were very weak (Table 4.7)
5185 and those predicting PAI indices were much stronger (Table 4.9).

5186 Neither full spectrum PLSR models, nor those using specific spectral blocks in the PLSR
5187 modelling, were able to produce very strong models to predict the LOI index. The near SWIR
5188 region from 1518 – 1780 nm produced the strongest model with a 3-component PLSR model
5189 with unsmoothed spectra producing a coefficient of determination of 0.47 ($P = 0.019$; Table
5190 4.7) with validation data. However this spectral block produced very poor quality prediction
5191 maps. The VNIR spectral block produced the best quality prediction maps from a speckling
5192 and striping point of view (Figure 4.25a and b) but the model statistics were weak with the
5193 strongest being a 1-component model with unsmoothed spectra (validation $r^2 = 0.36$, $P =$
5194 0.05 ; Table 4.7 and Figure 4.24). Garcia-Gomez and Maestre (2011) used the NDVI
5195 vegetation index to predict LOI with linear regression and had strong predictions with a
5196 range of validation r^2 values between 0.64 and 0.92 (all significant at $P < 0.05$) for 300
5197 random allocations of their data into calibration and validation data sets. It should be noted
5198 though that their pixel size was 15 m with ASTER imagery (García-Gómez and Maestre,
5199 2011) which is half the roughly 30 m pixel size of the Hyperion imagery, and the NDVI is
5200 generally a vegetation index focused on green photosynthetic vegetation (Huete and
5201 Jackson, 1987, Tucker, 1979).

5202 The weakness in predictions for the LOI in this study and the strength of those achieved by
5203 Garcia-Gomez and Maestre (2011) raises questions about combining the grass tufts, which is
5204 the real patch size in these grasslands, into homogenous patches. This meant that the LOI is
5205 being calculated from two different patch definitions: a large scale consisting of a number of
5206 smaller patches (grass tufts) for the grass matrix (sparse grass; Figure 4.33) and a smaller
5207 scale based on individual units (bare/biological soil crust patch or single tall grass tuft; Figure
5208 4.33). This may have resulted in distorted LOI values and thus contributed to the difference
5209 in results obtained in this study compared to those of Garcia-Gomez and Maestre (2011).
5210 There is also a question mark around how well the line intercept method is at characterizing
5211 all components of a landscape. During the fieldwork, it became apparent that the
5212 distribution and presence of *S. plumosum* patches, which are not common on the six
5213 transects covered by the Hyperion imagery but were common on the five transects missed
5214 by the Hyperion imagery, were not being well characterized by the line intercept method

that forms the basis of the LFA technique (D. Tongway, pers. comm.). It is likely that this may be the case with other small patch types such as tall grass patches which, from observations while collecting the field data, had a variable distribution through the landscape ranging from fairly high local densities, to individual scattered tufts, to areas of no presence. This will be explored in further detail in the discussion of the PAI results.

The PAI models were exceptionally strong whether using full spectrum models, specific blocks of the spectrum, or applying smoothing or no smoothing to the spectra (Table 4.9 and Figure 4.27a and b). For instance the full spectrum PLSR model with two components and no spectral smoothing achieved a coefficient of determination of 0.78 ($P = 0.0003$) for validation data. It does not seem as if any other author has attempted to predict the PAI from remote sensing imagery. Gaitán *et al* (2013) did find a weak linear relationship ($r^2 = 0.31$) between ground cover (basal cover) and the NDVI derived from the MODIS sensor but it should be noted that their pixel size was quite large at 250 x 250 m. Essentially the PAI is an indicator of density of patch organization in a landscape (Figure 4.33), although in calculating patch area, it does not take into account patches outside the patch width intercepting the transect but within the 5 m limit either side of the tape measure. Generally, remote sensing studies have studied landscape spatial pattern at larger scales such as land cover and land use (Betbeder *et al.*, 2014, Petropoulos *et al.*, 2012, Hirche *et al.*, 2011, Coetzer *et al.*, 2010, Brown, 1998) than the sub-pixel scale of this study at Vaal River where the fundamental scale of organization is grass tufts. Furthermore, the weakness of any LOI models and the strength of the PAI models is surprising as field measures of these two indices scaled to pixel size are fairly strongly correlated ($r = 0.66$, $P < 0.001$, Figure 4.7), and should be, as the PAI is only using one more bit of information (patch width) than does the LOI (patch length). The purpose of measuring the width in the PAI is to get an indicator of the resource trapping ability of patches (Dlamini *et al.*, 2011, Bautista *et al.*, 2007, Bergkamp, 1998, Tongway and Ludwig, 1996). Considering the strongly confounding affects of fire already shown for the process-based LFA indices of stability, infiltration and nutrient cycling, an examination of what happens to grassland recovering from fire may shed some light on the why the LOI and PAI had such divergent model results.

The immediate effect of a fire is to remove all or most of the above-ground vegetation, depending on the type and characteristics of the fire, leaving behind charred basal grass tuft

stumps (Figure 4.36c) and possibly some standing seed culms depending on the position of the flames due to the type of fire, fuel properties and the grass species (Snyman, 2003). Fire generally does not appear to alter the species composition of grasslands but can shift the relative abundance of species (Snyman and Bredenkamp, 2005, Snyman, 2003). Snyman (2003), in a semi-arid grassland west of the Vaal River sites, showed that basal cover was a third that of unburnt treatments two months after a fire and that plant density (plants/m²) was reduced by between 14 and 25% two months after the burn. He further showed that basal diameter had not recovered to unburnt levels two years after the burn but that plant density had recovered by the end of the first year (Snyman, 2003). However it should be noted that Snyman (2003) only took into account the living parts in the basal cover estimates and not the dead parts which at Vaal River were still present after a burn (Figure 4.36c). Thus a remote sensing pixel would still include the remaining dead part in the mix of spectra that would make up the spectrum recorded by a sensor for a particular pixel. Removal of above-ground growth stimulates the development of tillers leading to recovery of the aerial parts of the grass plant (Tainton and Mentis, 1984). In the Serengeti, fire stimulated the production of biomass such that burnt plots had higher primary productivity in comparison to unburnt plots in the weeks following burning (Hassan, 2011). In the season after the burn, Snyman (2003) found that burnt areas produced 37% less above ground biomass than did unburnt areas through reduced photosynthetic biomass, increased runoff from burnt areas, and ability to respond to available soil moisture (Nippert *et al.*, 2006, Snyman, 2003, Georgiadis *et al.*, 1989). The amount of litter produced two months after the burn treatment was 5% to 10% that of the litter produced in the unburnt treatments depending on the type of fire (Snyman, 2003).

Despite the removal of all above-ground biomass by fire, it can be seen from Figure 4.36c that once the ash has blown or washed away, or broken down into the soil, a presence of the individual grass tufts still remains. A bird's eye view would show that bare patches and biological soil crust patches would be a little larger than before the fire through the removal of the surrounding grass canopy. The matrix or dominant cover in the grassland, which had been classified as sparse grass, would have switched from grass-dominated to soil-dominated. However, the density of grass tufts within this matrix, which includes living and dead material, would still be the same. From a spectral perspective, there would be a



(a)



(b)



(c)



(d)

Figure 4.36 (a) Fire approaching transect vr15 in 2011. (b) Transect vr15 after the fire. (c) Post-fire recovery during September 2011 in a fire break near transect vr7. (d) Recovery of grassland during October 2011 from an undated fire east of vr13, vr14 and vr15.

rearrangement of the spectral components making up a pixel spectrum. Before the fire, the spectral signal would be a mix of soil, senescent grass and possibly some green photosynthetic grass (Huete and Jackson, 1987, Tucker, 1978, Tucker, 1977). Immediately after the fire, as seen in Figure 4.36a and b, the spectral signal would be dominated by ash and charcoal with some soil spectra (Eva and Lambin, 1998b), and the soil spectra would increase with time as the ash got blown off site or broken down (Eva and Lambin, 1998b). The spectral signal for the vegetation would switch from senescent vegetation with limited photosynthetic pigments, to a mix of charcoal and ash, and then to a mix of charcoal and young green photosynthetic material (Figure 3.36c) with the green photosynthetic signal gradually dominating the charcoal signal (Figure 4.36d) as time progressed. Furthermore, the more time after the burn, the less soil and charred remains would be exposed, and the greater the development of a canopy as seen in Figure 4.36b, c and d (Eva and Lambin, 1998b). So from a sensor's perspective and viewing the grass tufts as points or circles, what would change are the spectral properties of the points (Eva and Lambin, 1998b) and the size of points, but not their density.

Essentially the measurements made in the field to calculate the LOI and PAI (Tongway and Hindley, 2004) were based on identifying relatively homogenous areas of grassland as a patch, and the boundaries of these grassland patches would be defined by a change in patch type or a change in density. So although the actual structure of the grassland has been dramatically changed by fire (Snyman and Bredenkamp, 2005, Snyman, 2003), the density of grass tufts, including living and dead material in the sward, has not. However, as the LOI is a measure of points in contact with the transect line, a reduction in the size of the points means many fewer points that would hit the transect line. The boundary of the area defining a homogenous density of points would change slightly, becoming a little smaller as the point size shrunk. Possibly, because the PAI is an area measure that covers more of the area of a pixel and the density of grass tufts has not changed as a result of fire, the PAI is more resilient to the kind of changes invoked in a landscape by a fire. In contrast, the LOI, being a line measure across the area of the pixel, would likely be something very different after a fire. Therefore, despite the spectral signal being very different before and after a fire, it is concluded that the PLSR modelling would still be able to map the PAI values in a relative manner as measured in the field, with spectra involving a mix of burnt and unburnt pixels,

whereas the PLSR modelling to predict LOI values in such conditions would be far more confounded. The one outstanding question is how would the density of grass tufts have changed over the two years? Generally perennial grass species are long lived so survivorship or recruitment of grasses during the two years (Lauenroth and Adler, 2008, Garnier and Dajoz, 2001) between field sampling and Hyperion image acquisition is believed to have had minimal impact.

4.4.4 **Above-Ground Biomass**

The PLSR models predicting biomass in the Vaal River region were all generally weak with the far NIR block (1185 - 1306 nm) giving the strongest validation statistics ($r^2 = 0.45$, $P = 0.048$) for a 5-component model (Table 4.10). However the prediction map for this model was excessively striped which was often the case for models with many components (Figure 4.32b). The VNIR spectral block, which usually produced a quality prediction map, had an inverse relationship between PLSR predicted and field measured values (Figure 4.31). The full spectrum model had a positive relationship between field measured and PLSR predicted values which was fairly strong for calibration data ($r^2 = 0.55$, $P = 0.001$) but substantially weaker for validation data ($r^2 = 0.25$, $P = 0.17$; Table 4.10 and Figure 4.31a) and did produce a prediction map free of speckling and stripes (Figure 4.32a). However, this prediction map presents trends that were previously noted with the three LFA indices of stability, infiltration and nutrient cycling. In essence, the range of above-ground biomass values in the burnt areas in the northern prediction block is very similar to the unburnt areas in the southern prediction block. Furthermore, the range in some of the burnt areas in the southern block is similar to unburnt areas in the northern prediction block. This suggests that the burn scars and recovery from fire are acting as a major confounding factor rendering the prediction map for above-ground biomass subject to errors in prediction. Under the circumstances, these weak models are not surprising as above-ground biomass is the variable of those predicted from PLSR modelling in this study that is most likely to be affected by the two-year difference between field-measured values and the acquisition of the spectra. Above-ground biomass is entirely dependent on rainfall over the season of growth (O'connor *et al.*, 2001, Epstein *et al.*, 1996, Sala *et al.*, 1988, Dye and Walker, 1987, Deshmukh, 1984) and standing biomass at the beginning of the growing season (Scurlock *et al.*, 2002). Furthermore, fire would have had a major impact on the above-ground biomass

(Snyman, 2006) as it would have removed all standing biomass from the grassland and so any above-ground biomass in the burn areas when the Hyperion images were collected would have been exclusively from fire recovery (Snyman, 2006), whereas any above-ground biomass in the unburnt areas would have been the remaining standing biomass from the previous growth season (Scurlock *et al.*, 2002). Thus the above-ground biomass values used as the response variable in the PLSR models were very likely something quite different to those responsible for the spectra which were the predictor variables in the models.

In contrast to the Hyperion derived PLSR models predicting above-ground biomass, the ground-based hyperspectral models from both West Wits and Vaal River performed substantially better at predicting above-ground biomass. West wits calibration models ($r^2 = 0.70$, $P < 0.0001$, $n = 57$; Table 3.15 and Figure 3.30a) were stronger than validation models ($r^2 = 0.55$, $P = 0.0003$, $n = 19$; Table 3.15 and Figure 3.30a). Vaal River had the reverse pattern with the calibration model ($r^2 = 0.67$, $P < 0.0001$, $n = 180$; Table 3.16 and Figure 3.30b) slightly weaker than validation model ($r^2 = 0.79$, $P < 0.0001$, $n = 59$; Table 3.16 and Figure 3.30b). The slightly weaker results for West Wits in comparison to Vaal River may be a function of West Wits being sampled at the beginning of the winter senescent period compared to Vaal River being sampled from the middle to end of the winter senescent period. West Wits grasses still had substantial green photosynthetic material at the base of the grass tufts despite the canopy appearing brown and dry. At Vaal River, no green photosynthetic material was present in the grasses when the sampling was conducted. These results illustrate very clearly how the delay in acquiring the Hyperion imagery and the subsequent burning of the grassland has undermined the ability to predict above-ground biomass from Hyperion imagery.

Cho *et al.* (2007) compared PLSR under three spectral treatments to vegetation indices derived from HyMap hyperspectral imagery as predictors of biomass. They found that PLSR provided better predictions of green grass/herb biomass than the vegetation indices, and that continuum removed (Clark and Roush, 1984) spectra using selected spectra with PLSR gave stronger prediction statistics ($r^2 = 0.83$ for calibration) than the no treatment and first derivative spectra (Cho *et al.*, 2007). These results by Cho *et al.* (2007) are slightly stronger than those from the ground-based spectra but no spectral selection or continuum removal was performed on the ground-based spectra from either West Wits or Vaal River.

5376 4.5 **Conclusion**

5377 The best models using Hyperion spectra and PLSR were obtained with the LFA PAI index
5378 which is an index of the area along a transect occupied by patches. It is believed that this
5379 model was strong, despite a mix of unburnt and burnt pixels, because the PAI is a two-
5380 dimensional index and the density of grass tufts may have been able to compensate for the
5381 reduction in canopy cover and a switch from mixed senesced and green photosynthetic
5382 material to purely photosynthetic material in the burnt pixels. This conclusion was partially
5383 supported by the LAI which produced very weak models. The LFA LAI is a one-dimensional
5384 index of landscape organization and thus the burn would have substantially altered the
5385 kinds of values that would have been measured on the ground when the Hyperion image
5386 was acquired (2013) compared to those obtained two years earlier when the field campaign
5387 was conducted (2011).

5388 The three LFA indices of stability, infiltration and nutrient cycling had quite strong
5389 predictions from PLSR models when the VNIR block using wavelengths between 478 nm and
5390 885 nm from Hyperion spectra were used in the model, but full spectrum PLSR models were
5391 considerably weaker. Prediction maps using these models with VNIR spectra for the
5392 infiltration and nutrient cycling indices were of high quality with minimal striping and
5393 speckling due to the quality of the Hyperion imagery. However, for the stability index, even
5394 the best prediction map was of poor quality. Furthermore, these prediction maps seemed to
5395 be severely undermined by the burnt pixels corresponding with transects in the northern
5396 block. This resulted in burnt areas in the northern grasslands having similar prediction
5397 values to unburnt areas in the southern block, and some unburnt pixels in the northern
5398 block having similar values to burnt pixels in the southern grassland block. The PLSR models
5399 derived from ground-based spectra collected at the same time as the LFA indices produced
5400 strong predictions for the LFA infiltration and nutrient cycling indices at both West Wits and
5401 Vaal River. However, for the stability index, only ground-based spectra and measurements
5402 from West Wits produced a strong PLSR model while those from Vaal River were weak. This
5403 weakness seems to be related to low variability in the measured LFA stability values,
5404 possibly due to the flatness of the terrain at Vaal River, such that different patch types had
5405 overlapping ranges for the LFA stability index. However the spectra for each patch type
5406 were distinct and this resulted in a large range of predicted stability values for a specific

5407 measured stability value, with the predicted values tending to order themselves according
5408 to patch type.

5409 The PLSR models predicting biomass from Hyperion spectra were weak, whereas those
5410 derived from ground-based spectra were fairly strong for both West Wits and Vaal River.
5411 This discrepancy between Hyperion derived PLSR models and ground-based models is
5412 believed to be due to the effects of the two-year separation between the field
5413 measurements and the acquisition of the Hyperion hyperspectral imagery. This meant that
5414 biomass in some pixels used in the PLSR models had been subjected to fire three months
5415 prior to the Hyperion image acquisition and therefore reflected post-fire recovery
5416 vegetation in a green-photosynthetic state where field measurements had been of winter-
5417 senesced dry-season vegetation. This delay in Hyperion acquisition also meant that unburnt
5418 pixels used in the model had spectra reflecting biomass produced under a different
5419 hydrological regime and a shift into a slightly later phenological state than was the case
5420 during the field sampling. Thus the delay in acquiring the Hyperion hyperspectral imagery,
5421 together with the shift in a different phenological state, and the burning of some pixels, all
5422 undermined the PLSR modelling to predict above-ground biomass from Hyperion
5423 hyperspectral imagery.

5424 In summary, this research has shown that satellite-based hyperspectral imagery used
5425 together with PLSR modelling has the potential to predict and map LFA indices across large
5426 heterogeneous landscapes such as mining environments. Obtaining strong models,
5427 however, requires good temporal matching between the hyperspectral imagery and the
5428 field work, and may also depend on other factors that require further investigation.

5429

5. Conclusions and Ways Forward

5430 5.1 Conclusions

5431 The purpose of this study was to develop a method to use remotely-sensed hyperspectral
5432 data to predict LFA indices in grasslands associated with deep-level gold mining activities
5433 and thereby produce ecological risk maps across the grassland areas, and to test the
5434 method using Hyperion hyperspectral data acquired from NASA/USGS EO-1 satellite
5435 platform (Ungar *et al.*, 2003, Pearlman *et al.*, 2001). This aim was broken down into three
5436 objectives: (1) the collection and evaluation of the LFA field data (Tongway and Hindley,
5437 2004) against other ecological data collected at the same time, (2) using ground-based
5438 hyperspectral spectra and PLSR models (Mevik and Wehrens, 2007, Geladi and Kowalski,
5439 1986b) to predict the LFA indices and other ecological variables, and (3) the use of satellite-
5440 acquired Hyperion hyperspectral imagery (Ungar *et al.*, 2003, Ungar, 2001) to develop PLSR
5441 models predicting five LFA indices and above-ground biomass. These models allow the
5442 production of risk maps for each of the six variables across natural relict grasslands within
5443 the mining environment. The method presented essentially partitions the landscape to
5444 match a pixel and its spectrum rather than the more common method of spectral mixture
5445 analysis whereby the spectrum is split using pure endmembers (Roberts *et al.*, 1998, Huete,
5446 1986) to read the variability in a landscape over a single pixel.

5447 The study began with an ecological field study of the grasslands in two mining regions, Vaal
5448 River and West Wits. The method used was landscape function analysis (LFA), which is a
5449 transect-and-patch method for monitoring ecological and environmental processes such as
5450 landscape organization, erosion, infiltration/runoff, and nutrient cycling (Tongway and
5451 Hindley, 2004). Soil cores and above-ground vegetation were also sampled with a minimum
5452 of 299 points at Vaal River and 75 points at West Wits. There was one modification made to
5453 the LFA technique where the grass patches, which most often were a carpet of individual
5454 grass tufts, were collapsed into homogenous areas. This was essentially in response to scale
5455 and economy issues related to the size of Hyperion pixels and their unknown position at the
5456 time of the field sampling. The effect of this is that the definition of patch type operated at
5457 two scales: one for the matrix grasslands which was based on homogenous cover, and a
5458 second based on a change in structure for patch types such as the edge of the canopy for

the shrub *S. plumosum*, or the edge of exposed soil that defined bare patches. LFA transects were 180 m long to accommodate a minimum of five Hyperion pixels of 30 x 30 m each. Eleven transects were sampled at Vaal River mining region and 5 transects in West Wits mining region for a total of 2 880 m of transect. Patch types were shown to possibly represent various successional trajectories (Macdonald, 1978) although this was not conclusive. There was evidence that cattle and other hoofed animals may be factors in some processes active in these grasslands through the trampling of biological soil crusts and the reduction of stable soil surfaces to fine powder. Wind deflation, rain splash and sheet wash are the prevalent active erosion processes as grass tufts were often slightly pedestalled with some fine lag material in the interstitial spaces. Water flow and infiltration also play a role as some bare grass patches were often either associated with runoff or temporary pooling of water, or with some other physical disturbance of the soil surface. Subsidiary ecological data recorded along with the LFA sampling provided support for the values calculated for the LFA stability and infiltration indices. However, subsidiary data for the LFA nutrient cycling index seemed to suggest that above-ground nutrient cycling was decoupled from below-ground processes (Abbadie *et al.*, 1992), possibly in response to a regular fire regime. Therefore it is suggested that the LFA nutrient cycling index adequately characterised above-ground nutrient cycling but not below-ground nutrient cycling.

At the same time as the ecological measurements were collected, ground-based spectra in the wavelength range of 400 – 2500 nm were collected for every point sampled. These spectra were used to predict LFA indices and edaphic and vegetation properties using PLSR modelling. The PLSR models predicting patch types from ground-based spectra were strong for patch types that were spectrally well-defined, such as *S. plumosum* and bare/biological soil crust patches, but struggled to separate patch types that formed a spectral continuum, such as bare soil/biological soil crust - bare grass – sparse grass patches. However, by clumping patch types that were essentially functionally and compositionally similar, such as sparse grass and grass patch types, the PLSR models predicting patch types were substantially improved, although the spectral similarity problem was not entirely overcome. Models predicting the LFA stability, infiltration and nutrient cycling indices from ground-based spectra were strong at both mine sites. The PLSR models predicting above-ground biomass and soil moisture were also strong for both mine sites. West Wits PLSR models

5490 predicting SOM were weak for validation data whereas for Vaal River, both calibration and
5491 validation results were strong when predicting SOM. Likewise soil EC models were weak at
5492 West Wits but fairly strong at Vaal River. This result was due to EC values at West Wits being
5493 very low, whereas at Vaal River a wide range of EC values for soil solutions was recorded
5494 with very high values associated with surface evaporites believed to be gypsum. All other
5495 ecological variables produced very weak PLSR models with ground-based hyperspectral
5496 spectra.

5497 Hyperion hyperspectral images were downloaded from the USGS Glovis site. The Vaal River
5498 image consisted of 220 images plus metadata. These images were stacked into a data cube,
5499 converted to at-sensor radiance, re-projected to a local projection, and atmospherically
5500 corrected with FLAASH to surface reflectance. Pixels matching LFA transect lengths longer
5501 than 20 m were extracted from the data cube for PLSR model calibration and validation.
5502 Pixel spectra consisted of a maximum of 105 wavelengths spread across the 478 – 2335 nm
5503 range at a minimum of 10 nm intervals after removing Hyperion bands with excessive
5504 stripping or atmospheric water noise. The LFA index that produced the strongest PLSR
5505 models with Hyperion spectra was the PAI or patch area index, an index that relates patch
5506 area to transect area. In contrast, the LOI or landscape organisation index performed very
5507 poorly with PLSR modelling and Hyperion spectra. The most likely reason these two very
5508 similar LFA indices produced such different PLSR model results is related to the PAI being a
5509 two dimensional index and therefore adequately reflecting landscape structure at the pixel
5510 scale (30 x 30 m) of the Hyperion imagery. The LOI, on the other hand, is one-dimensional,
5511 being a straight line slice of landscape structure. Considering the landscape history just prior
5512 to the Hyperion image acquisition, it is postulated that fire changed the size/diameter of
5513 individual structural points (grass tuft and shrub boundaries) but did not change the
5514 arrangement of these in terms of dispersion patterns. Thus the Hyperion-derived PAI still
5515 reflected strong correspondence with the LFA patch dimensions recorded 2 years before the
5516 Hyperion acquisition because these are both 2-dimensional measurements. The LOI, being a
5517 one-dimensional line, was markedly affected by the changes in patch diameter and
5518 therefore was quite different between field sampling and Hyperion sampling, and thus
5519 produced weak models with Hyperion imagery.

The temporal mismatch between the field data and the Hyperion spectral data by two phenological cycles, as well as a shift into the early stages of the next phase of the phenological cycle, and the effect of fire prior to the spectral measurements, which had been purposefully excluded from the ground measurements two years earlier, proved to be major confounding factors in modelling the other LFA indices and above-ground biomass. This essentially caused a collapse in the meaningfulness of the prediction maps for the three LFA indices of stability, infiltration and nutrient cycling, as well as that for above-ground biomass. Furthermore, the poor quality of Hyperion imagery severely undermined the PLSR derived maps. The data exhibited excessive striping, particularly in the SWIR but not limited to the SWIR, as well as poorly calibrated spectra along the entire length of the Hyperion spectrum whereby spectral values tended to oscillate between successive bands. These weaknesses in the Hyperion imagery often restricted the PLSR modelling to the VNIR spectral blocks to reduce these effects on the output prediction maps. It is thought that this predominance of the VNIR for strong models and quality prediction maps is partly the result of shorter wavelengths possibly having more energy and therefore greater signal to noise ratios in the spectra. Smoothing was applied but this didn't necessarily improve the modelling results although it did reduce speckling in prediction maps.

5.2 **Recommendations going forward**

In the LFA, the question of collapsing grassland patch structure into homogenous units and whether this still reflects the natural landscape organization remains unanswered. This means that the scale at which patches were measured in the field is at least one step removed from the actual scale that stresses on the landscape would play out on the vegetation component. The scale that stress would act on is that of the individual living unit, which is the grass-tuft scale in these grasslands and is regulated by the establishment, resilience and mortality of individual grass tufts. The decision to collapse the grass tufts into homogenous units was made purely from an economic perspective in terms of time constraints for field sampling. However, despite the good results with the PAI, the PAI values might have been quite different if the true patch size, i.e. at grass tuft scale, had been used. As this is a truer reflection of landscape organisation than the homogenous grassland units used in this study, and therefore of the underlying factors determining the organization of landscape, it remains an important question that should be explored further.

5551 Another area of future exploration could be to determine what, if any, of the ecological
5552 variables measured in this study are related to the different patch types found in these
5553 grasslands. Such a question can be explored through the use of nominal multinomial logistic
5554 regression (Agresti, 2007) which allows for modelling the patch types as nominal response
5555 variables against all the ecological variables as predictor variables. This would show which
5556 ecological variables have a relationship to specific patch types and which are unrelated to
5557 any patch definition. Answering this question might provide some insight into how these
5558 different patch types form and aid in management decisions for the management of these
5559 threatened and relict grasslands.

5560 From a hyperspectral perspective, the most important recommendation is the need to
5561 temporally match the spectral data acquisition with the field measurements. To this end, an
5562 ongoing step in this study is the use of the Landsat imagery acquired at the same time as the
5563 field data to test PLSR modelling with a low spectral-resolution, broad-band sensor with
5564 identical pixel size as the Hyperion data. Another aspect that deserves more research is the
5565 applicability of the PAI to characterise the landscape from a sensors perspective. The
5566 problem of mixed pixels in remote sensing is ubiquitous and the dominant approach has
5567 been spectral mixture analysis (Kuemmerle *et al.*, 2006, Bateson *et al.*, 2000, Drake *et al.*,
5568 1999, Roberts *et al.*, 1998, Huete, 1986). However the PAI presents a method whereby one
5569 can resolve the landscape heterogeneity into a single value rather than trying to unmix the
5570 spectral profile into its components with pure endmember spectra. To this end, the LOI may
5571 also be useful despite performing poorly in this study. This is expected to be due to the
5572 effect of fire on the landscape, and the delay between the field data and Hyperion image
5573 acquisition, rather than any inherent weakness in the LOI *per se*. A further area of future study
5574 should include how much of a transect needs to fall on a pixel to regard that transect as
5575 representative of the pixel area. In this study, the rule of two thirds of a pixel length, which
5576 is 20 m of transect, was an entirely arbitrary decision. Airborne sensors, because they fly
5577 closer to the earth than satellites, can provide higher spatial resolution that more closely
5578 matches the scale of landscape organisation for the same or higher spectral resolution, but
5579 are expensive. Considering 2880 m of grassland was measured with LFA transects for this
5580 study, of which 1080 m provided 31 pixel spectra (910 m maximum), the labour cost of
5581 redundant field work is high. Thus the trade off between pixel size and spectral resolution

5582 with satellite sensors, together with sub-transect redundancy because of unknown pixel
5583 positions, means an idea of what proportion of a pixel on the ground can be regarded as
5584 representative of that pixel would be valuable.

5585 This study is valuable because hyperspectral remote sensing is a young discipline and this
5586 project presents an alternative method to model and map landscape processes and
5587 structure at sub-pixel scales using spectra acquired by satellite. Most sensors are
5588 multispectral with the Faculty of Geo-Information Science and Earth Observation of the
5589 University of Twente (ITC) currently listing 328 sensors on its list of satellite remote-sensing
5590 systems currently active or planned. As far as is known, there are only two hyperspectral
5591 sensors active: the EO-1 Hyperion sensor was the first, and more recently the Chinese HJ-1A
5592 HSI is actively acquiring images (Zhang *et al.*, 2013). Others expected soon include the
5593 European HSI on EnMAP (Steffler *et al.*, 2009, Steffler *et al.*, 2007), NASA's Hyperspectral
5594 InfraRed Imager or HypIRI (Lee *et al.*, 2015) and the Italian PRISMA (PREcursore
5595 IperSpettrale della Missione Applicativa) hyperspectral imager (Labate *et al.*, 2009). Spatial
5596 resolution for these proposed sensors is 30 m for the HSI and PRISMA sensors which is the
5597 same spatial resolution used to test the methods in this study, and both sensors have
5598 expected swath widths of 30 km, which is a better spread of the landscape than Hyperion's
5599 7.5 km swath width.

5600 Notwithstanding the problems caused by the delay in acquiring the Hyperion imagery, LFA,
5601 PLSR and hyperspectral imagery have shown they can be used to produce LFA prediction
5602 maps and characterise ecological risk across large landscapes (Lau *et al.*, 2008, Ong *et al.*,
5603 2008). The performance of the PAI also vindicates the decision to switch from adapting the
5604 LFA technique to fit the spectral measurements as was done in the pilot study (Furniss *et al.*,
5605 2009, Furniss, 2008), to using the entire LFA method to structure the measurements and
5606 adapting, on a pixel basis, the transect lengths to match the spectral measurements. Thus it
5607 can be seen that this study presents an important contribution to the fields of hyperspectral
5608 imaging and landscape ecology.

5609 6. **Reference**

- 5610 Abbadie, L., Mariotti, A. & Menaut, J.-C. 1992. Independence of savanna grasses from soil
5611 organic matter for their nitrogen supply. *Ecology*, 73, 608-613.
- 5612 Abrahams, A. D., Parsons, A. J. & Wainwright, J. 1995. Effects of vegetation change on
5613 interrill runoff and erosion, Walnut Gulch, southern Arizona. *Geomorphology*, 13, 37-
5614 48.
- 5615 Acocks, J. P. H. 1988. *Veld types of South Africa, edn 3*, Pretoria., Botanical Research
5616 Institute, .
- 5617 Aerts, R., Bakker, C. & Caluwe, H. D. 1992. Root turnover as determinant of the cycling of C,
5618 N, and P in a dry heathland ecosystem. *Biogeochemistry*, 15, 175-190.
- 5619 Agresti, A. 2007. *An introduction to categorical data analysis (2nd ed.)*, Hoboken, New
5620 Jersey, John Wiley & Sons, Inc.
- 5621 Aguiar, M. R. & Sala, O. E. 1999. Patch structure, dynamics and implications for the
5622 functioning of arid ecosystems. *Trends in Ecology & Evolution*, 14, 273-277.
- 5623 Aide, M. 2009. Three missouri alfisols impacted by aeolian deposition of lead-zinc-cadmium-
5624 silver-indium bearing mine tailings. *Soil & Sediment Contamination*, 18, 43-54.
- 5625 Akcil, A. & Koldas, S. 2006. Acid mine drainage (amd): Causes, treatment and case studies.
5626 *Journal of Cleaner Production*, 14, 1139-1145.
- 5627 Allison, F. E. 1973. Chapter 11: Nonbiological immobilization of nitrogen. In: ALLISON, F. E.
5628 (ed.) *Developments in soil science*. Elsevier.
- 5629 Anderson, J. E. & Robbins, E. I. 1998. Spectral reflectance and detection of iron-oxide
5630 precipitates associated with acidic mine drainage. *Photogrammetric Engineering &*
5631 *Remote Sensing*, 64, 1201-1208.
- 5632 Anderson, R. C. 2006. Evolution and origin of the central grassland of North America:
5633 Climate, fire, and mammalian grazers. *The Journal of the Torrey Botanical Society*,
5634 133, 626-647.
- 5635 Anderson, T.-H. 2003. Microbial eco-physiological indicators to asses soil quality.
5636 *Agriculture, Ecosystems & Environment*, 98, 285-293.
- 5637 Anderson, T.-H. & Domsch, K. H. 1989. Ratios of microbial biomass carbon to total organic
5638 carbon in arable soils. *Soil Biology and Biochemistry*, 21, 471-479.
- 5639 Archibald, S. & Scholes, R. J. 2007. Leaf green-up in a semi-arid African savanna -separating
5640 tree and grass responses to environmental cues. *Journal of Vegetation Science*, 18,
5641 583-594.
- 5642 Asrar, G., Kanemasu, E. T., Miller, G. P. & Weiser, R. L. 1986. Light interception and leaf area
5643 estimates from measurements of grass canopy reflectance. *Geoscience and Remote*
5644 *Sensing, IEEE Transactions on*, GE-24, 76-82.
- 5645 Austin, A. T. & Ballaré, C. L. 2010. Dual role of lignin in plant litter decomposition in
5646 terrestrial ecosystems. *Proceedings of the National Academy of Sciences of the USA*,
5647 107, 4618–4622.
- 5648 Austin, A. T. V. L. 2006. Plant litter decomposition in a semi-arid ecosystem controlled by
5649 photodegradation. *Nature*, 442, 555-558.
- 5650 Bagnold, R. A. 1941. *The physics of blown sand and desert dunes*, London, Methuen & Co.
5651 Ltd.
- 5652 Bannari, A., Guedon, A. M., El-Harti, A., Cherkaoui, F. Z. & El-Ghmari, A. 2008.
5653 Characterization of slightly and moderately saline and sodic soils in irrigated

5654 agricultural land using simulated data of Advanced Land Imaging (EO-1) sensor.
 5655 *Communications in Soil Science & Plant Analysis*, 39, 2795-2811.

5656 Barbour, M. G., Burk, J. H. & Pitts, W. D. 1987. *Terrestrial plant ecology*, Menlo Park,
 5657 California, The Benjamin/Cummings Publishing Company, Inc.

5658 Barrett, J. E. & Burke, I. C. 2000. Potential nitrogen immobilization in grassland soils across a
 5659 soil organic matter gradient. *Soil Biology and Biochemistry*, 32, 1707-1716.

5660 Barry, P., Segal, C., Pearlman, J., Jarecke, P. & Folkman, M. 2002. Hyperion data collection:
 5661 Performance assessment and science application. *In: Aerospace Conference*
 5662 *Proceedings*, 2002. IEEE, 2002. 3-1493-3-1499 vol.3.

5663 Bartley, R., Roth, C. H., Ludwig, J., Mcjannet, D., Liedloff, A., Corfield, J., Hawdon, A. &
 5664 Abbott, B. 2006. Runoff and erosion from Australia's tropical semi-arid rangelands:
 5665 Influence of ground cover for differing space and time scales. *Hydrological Processes*,
 5666 20, 3317-3333.

5667 Bastarrika, A., Chuvieco, E. & Martín, M. P. 2011. Mapping burned areas from Landsat
 5668 TM/ETM+ data with a two-phase algorithm: Balancing omission and commission
 5669 errors. *Remote Sensing of Environment*, 115, 1003-1012.

5670 Bastida, F., Zsolnay, A., Hernández, T. & García, C. 2008. Past, present and future of soil
 5671 quality indices: A biological perspective. *Geoderma*, 147, 159-171.

5672 Bateson, C. A., Asner, G. P. & Wessman, C. A. 2000. Endmember bundles: A new approach to
 5673 incorporating endmember variability into spectral mixture analysis. *Geoscience and*
 5674 *Remote Sensing, IEEE Transactions on*, 38, 1083-1094.

5675 Bautista, S., Mayor, A. G., Bourakhouadar, J. & Bellot, J. 2007. Plant spatial pattern predicts
 5676 hillslope runoff and erosion in a semiarid mediterranean landscape. *Ecosystems*, 10,
 5677 987-998.

5678 Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.-M. & Mcbratney, A. 2010.
 5679 Critical review of chemometric indicators commonly used for assessing the quality of
 5680 the prediction of soil attributes by nir spectroscopy. *TrAC Trends in Analytical*
 5681 *Chemistry*, 29, 1073-1081.

5682 Belnap, J. 1995. Surface disturbances: Their role in accelerating desertification.
 5683 *Environmental Monitoring and Assessment*, 37, 39-57.

5684 Belnap, J. & Gillette, D. A. 1998. Vulnerability of desert biological soil crusts to wind erosion:
 5685 The influences of crust development, soil texture, and disturbance. *Journal of Arid*
 5686 *Environments*, 39, 133-142.

5687 Belnap, J., Welter, J. R., Grimm, N. B., Barger, N. & Ludwig, J. A. 2005. Linkages between
 5688 microbial and hydrologic processes in arid and semiarid watersheds. *Ecology*, 86,
 5689 298-307.

5690 Ben-Dor, E. & Banin, A. 1994. Visible and near-infrared (0.4-1.1 μm) analysis of arid and
 5691 semiarid soils. *Remote Sensing of Environment*, 48, 261-274.

5692 Ben-Dor, E. & Banin, A. 1995. Near-infrared analysis as a rapid method to simultaneously
 5693 evaluate several soil properties. *Soil Sci. Soc. Am. J.*, 59, 364-372.

5694 Ben-Dor, E., Chabrillat, S., Demattê, J. a. M., Taylor, G. R., Hill, J., Whiting, M. L. & Sommer,
 5695 S. 2009. Using imaging spectroscopy to study soil properties. *Remote Sensing of*
 5696 *Environment*, 113, S38-S55.

5697 Ben-Dor, E., Inbar, Y. & Chen, Y. 1997. The reflectance spectra of organic matter in the
 5698 visible near-infrared and short wave infrared region (400-2500 nm) during a
 5699 controlled decomposition process. *Remote Sensing of Environment*, 61, 1-15.

- 5700 Ben-Dor, E., Patkin, K., Banin, A. & Karnieli, A. 2002. Mapping of several soil properties using
5701 dais-7915 hyperspectral scanner data - a case study over clayey soils in Israel.
5702 *International Journal of Remote Sensing*, 23, 1043-1062.
- 5703 Ben-Hur, M. & Lado, M. 2008. Effect of soil wetting conditions on seal formation, runoff, and
5704 soil loss in arid and semiarid soils — a review. *Australian Journal of Soil Research*, 46,
5705 191-202.
- 5706 Bengtsson, G., Bengtson, P. & Månsson, K. F. 2003. Gross nitrogen mineralization-,
5707 immobilization-, and nitrification rates as a function of soil C/N ratio and microbial
5708 activity. *Soil Biology and Biochemistry*, 35, 143-154.
- 5709 Benton Jones Jr., J. 2001. *Laboratory guide for conducting soil tests and plant analysis*, CRC
5710 Press LLC.
- 5711 Bergkamp, G. 1998. A hierarchical view of the interactions of runoff and infiltration with
5712 vegetation and microtopography in semiarid shrublands. *CATENA*, 33, 201-220.
- 5713 Bertels, L. 2012. Colibri: The envi/idl code library.
- 5714 Betbeder, J., Gond, V., Frappart, F., Baghdadi, N. N., Briant, G. & Bartholome, E. 2014.
5715 Mapping of Central Africa forested wetlands using remote sensing. *Selected Topics in*
5716 *Applied Earth Observations and Remote Sensing, IEEE Journal of*, 7, 531-542.
- 5717 Bhark, E. W. & Small, E. E. 2003. Association between plant canopies and the spatial
5718 patterns of infiltration in shrubland and grassland of the Chihuahuan desert, New
5719 Mexico. *Ecosystems*, 6, 0185-0196.
- 5720 Bhetanabhotla, M. N., Crowell, B. A., Coucouvinos, A., Hill, R. D. & Rinker, R. G. 1985.
5721 Simulation of trace species production by lightning and corona discharge in moist air.
5722 *Atmospheric Environment (1967)*, 19, 1391-1397.
- 5723 Bielders, C. L. & Grymonprez, B. 2010. Raindrop impact: A neglected but potentially major
5724 contributor to particle mobilization in soils. *Soil Science Society of America Journal*,
5725 74, 1446-1456.
- 5726 Bigham, J. M., Schwertmann, U. & Pfab, G. 1996. Influence of pH on mineral speciation in a
5727 bioreactor simulating acid mine drainage. *Applied Geochemistry*, 11, 845-849.
- 5728 Bogner, C., Gaul, D., Kolb, A., Schmiedinger, I. & Huwe, B. 2010. Investigating flow
5729 mechanisms in a forest soil by mixed-effects modelling. *European Journal of Soil*
5730 *Science*, 61, 1079-1090.
- 5731 Bolton Jr, H., Smith, J. L. & Link, S. O. 1993. Soil microbial biomass and activity of a disturbed
5732 and undisturbed shrub-steppe ecosystem. *Soil Biology and Biochemistry*, 25, 545-
5733 552.
- 5734 Borisova, D., Nikolov, H., Danov, M. & Tsanev, V. 2007. Comparison between
5735 reflectance/emittance spectra of iron-containing minerals. *Recent Advances in Space*
5736 *Technologies, 2007. RAST '07. 3rd International Conference on*.
- 5737 Bouwman, A. F., Grinsven, J. J. M. V. & Eickhout, B. 2010. Consequences of the cultivation of
5738 energy crops for the global nitrogen cycle. *Ecological Applications*, 20, 101-109.
- 5739 Boyer, M., Miller, J., Belanger, M., Hare, E. & Wu, J. 1988. Senescence and spectral
5740 reflectance in leaves of northern pin oak (*Quercus palustris muenchh.*). *Remote*
5741 *Sensing of Environment*, 25, 71-87.
- 5742 Breshears, D. & Barnes, F. 1999. Interrelationships between plant functional types and soil
5743 moisture heterogeneity for semiarid landscapes within the grassland/forest
5744 continuum: A unified conceptual model. *Landscape Ecology*, 14, 465-478.

- 5745 Brown, J. F. 1998. Mapping global grassland ecosystems: A comparison of four data sets. *In*:
5746 Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS '98. 1998
5747 IEEE International, 1998. 2107-2111 vol.4.
- 5748 Burger, M. & Jackson, L. E. 2003. Microbial immobilization of ammonium and nitrate in
5749 relation to ammonification and nitrification rates in organic and conventional
5750 cropping systems. *Soil Biology and Biochemistry*, 35, 29-36.
- 5751 Burke, I. C., Lauenroth, W. K. & Parton, W. J. 1997. Regional and temporal variation in net
5752 primary production and nitrogen mineralization in grasslands. *Ecology*, 78, 1330-
5753 1340.
- 5754 Burke, I. C., Lauenroth, W. K., Vinton, M. A., Hook, P. B., Kelly, R. H., Epstein, H. E., Aguiar, M.
5755 R., Robles, M. D., Aguilera, M. O., Murphy, K. L. & Gill, R. A. 1998. Plant-soil
5756 interactions in temperate grasslands. *Biogeochemistry*, 42, 121-143.
- 5757 Burns, D. A. 2004. The effects of atmospheric nitrogen deposition in the rocky mountains of
5758 colorado and southern wyoming, USA—a critical review. *Environmental Pollution*,
5759 127, 257-269.
- 5760 Burylo, M., Rey, F., Mathys, N. & Dutoit, T. 2012. Plant root traits affecting the resistance of
5761 soils to concentrated flow erosion. *Earth Surface Processes and Landforms*, 37, 1463-
5762 1470.
- 5763 Butterfield, H. S. & Malmström, C. M. 2009. The effects of phenology on indirect measures
5764 of aboveground biomass in annual grasses. *International Journal of Remote Sensing*,
5765 30, 3133-3146.
- 5766 Cammeraat, E., Van Beek, R. & Kooijman, A. 2005. Vegetation succession and its
5767 consequences for slope stability in SE Spain. *Plant & Soil*, 278, 135-147.
- 5768 Campbell, P. K. E., Middleton, E. M., Thome, K. J., Kokaly, R. F., Huemmrich, K. F.,
5769 Lagomasino, D., Novick, K. A. & Brunsell, N. A. 2013. EO-1 Hyperion reflectance time
5770 series at calibration and validation sites: Stability and sensitivity to seasonal
5771 dynamics. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE*
5772 *Journal of*, 6, 276-290.
- 5773 Carling, P. A., Glaister, M. S. & Flinham, T. P. 1997. The erodibility of upland soils and the
5774 design of preafforestation drainage networks in the United Kingdom. *Hydrological*
5775 *Processes*, 11, 1963-1980.
- 5776 Carmona, D. M., Faz Cano, Á. & Arocena, J. M. 2008. Minerals in salt efflorescence and
5777 leaching of metals in mined soils in SE Spain – preliminary results. *In*: KERTÉSZ, Á.,
5778 VÁRALLYAY, G., NÉMETH, T., FÜLEKY, G. & SISÁK, I., eds. 15th International Congress
5779 of ISCO, 2008 Budapest.
- 5780 Carter, G. A. & Knapp, A. K. 2001. Leaf optical properties in higher plants: Linking spectral
5781 characteristics to stress and chlorophyll concentration. *American Journal of Botany*,
5782 88, 677-684.
- 5783 Carvalho, I. G., Cidu, R., Fanfani, L., Pitsch, H., Beaucaire, C. & Zuddas, P. 2005.
5784 Environmental impact of uranium mining and ore processing in the Lagoa Real
5785 district, Bahia, Brazil. *Environmental Science & Technology*, 39, 8646-8652.
- 5786 Casenave, A. & Valentin, C. 1992. A runoff capability classification system based on surface
5787 features criteria in semi-arid areas of West Africa. *Journal of Hydrology*, 130, 231-
5788 249.
- 5789 Chabrilat, S., Goetz, A. F. H., Krosley, L. & Olsen, H. W. 2002. Use of hyperspectral images in
5790 the identification and mapping of expansive clay soils and the role of spatial
5791 resolution. *Remote Sensing of Environment*, 82, 431-445.

- 5792 Chaney, K. & Swift, R. S. 1984. The influence of organic matter on aggregate stability in some
5793 British soils. *Journal of Soil Science*, 35, 223-230.
- 5794 Chang, C. W., Laird, D. A., Mausbach, M. J. & Hurburgh, C. R. 2001. Near-infrared reflectance
5795 spectroscopy—principal components regression analysis of soil properties. *Soil Sci.*
5796 *Soc. Am. J.*, 65, 480-490.
- 5797 Charley, J. L. & West, N. E. 1975. Plant-induced soil chemical patterns in some shrub-
5798 dominated semi-desert ecosystems of Utah. *Journal of Ecology*, 63, 945-963.
- 5799 Chartier, M. P., Rostagno, C. M. & Videla, L. S. 2013. Selective erosion of clay, organic carbon
5800 and total nitrogen in grazed semiarid rangelands of northeastern Patagonia,
5801 Argentina. *Journal of Arid Environments*, 88, 43-49.
- 5802 Chejara, V. K., Kriticos, D. J., Kristiansen, P., Sindel, B. M., Whalley, R. D. B. & Nadolny, C.
5803 2010. The current and future potential geographical distribution of *Hyparrhenia*
5804 *hirta*. *Weed Research*, 50, 174-184.
- 5805 Cheng, W. & Coleman, D. C. 1990. Effect of living roots on soil organic matter
5806 decomposition. *Soil Biology and Biochemistry*, 22, 781-787.
- 5807 Cho, M. A., Skidmore, A., Corsi, F., Van Wieren, S. E. & Sobhan, I. 2007. Estimation of green
5808 grass/herb biomass from airborne hyperspectral imagery using spectral indices and
5809 partial least squares regression. *International Journal of Applied Earth Observation*
5810 *and Geoinformation*, 9, 414-424.
- 5811 Cho, M. A. & Skidmore, A. K. 2006. A new technique for extracting the red edge position
5812 from hyperspectral data: The linear extrapolation method. *Remote Sensing of*
5813 *Environment*, 101, 181-193.
- 5814 Chuvieco, E., Martín, M. P. & Palacios, A. 2002. Assessment of different spectral indices in
5815 the red-near-infrared spectral domain for burned land discrimination. *International*
5816 *Journal of Remote Sensing*, 23, 5103-5110.
- 5817 Clark, R. N. & Roush, T. L. 1984. Reflectance spectroscopy: Quantitative analysis techniques
5818 for remote sensing applications. *Journal of Geophysical Research: Solid Earth*, 89,
5819 6329-6340.
- 5820 Clark, R. N., 1999, Chapter 1: Spectroscopy of Rocks and Minerals, and Principles of
5821 Spectroscopy, in *Manual of Remote Sensing*, Volume 3, Remote Sensing for the Earth
5822 Sciences, (A.N. Rencz, ed.) John Wiley and Sons, New York, p 3- 58
- 5823 Clark, R. N., Swayze, G. A., Livo, K. E., Kokaly, R. F., King, T. V. V., Dalton, J. B., J. S. Vance,
5824 Rockwell, B. W., Hoefen, T. & McDougal, R. R. 2002. Surface reflectance calibration of
5825 terrestrial imaging spectroscopy data: A tutorial using aviris. In: Proceedings of the
5826 10th Airborne Earth Science Workshop, 2002. JPL Publication.
- 5827 Clark, R. N., Swayze, G. A., Wise, R., Livo, E., Hoefen, T., Kokaly, R. & Sutley, S. J. 2007. USGS
5828 digital spectral library splib06a. In: U.S. GEOLOGICAL SURVEY, D. D. S. (ed.).
- 5829 Cleveland, C. & Liptzin, D. 2007. C:N:P stoichiometry in soil: Is there a “redfield ratio” for the
5830 microbial biomass? *Biogeochemistry*, 85, 235-252.
- 5831 Cleveland, C. C., Townsend, A. R., Schimel, D. S., Fisher, H., Howarth, R. W., Hedin, L. O.,
5832 Perakis, S. S., Latty, E. F., Von Fischer, J. C., Elseroad, A. & Wasson, M. F. 1999. Global
5833 patterns of terrestrial biological nitrogen (N₂) fixation in natural ecosystems. *Global*
5834 *Biogeochemical Cycles*, 13, 623-645.
- 5835 Clevers, J. G. P. W. 1989. Application of a weighted infrared-red vegetation index for
5836 estimating leaf area index by correcting for soil moisture. *Remote Sensing of*
5837 *Environment*, 29, 25-37.

- 5838 Coetzer, K. L., Erasmus, B. F. N., Witkowski, E. T. F. & Bachoo, A. K. 2010. Land-cover change
5839 in the Kruger to Canyons biosphere reserve (1993-2006): A first step towards
5840 creating a conservation plan for the subregion. *South African Journal of Science*, 106,
5841 1-10.
- 5842 Cohen, C. 1937. Some aspects of the reproduction of *Stoebe vulgaris*, and certain local grass
5843 and weed seeds. *South African Journal of Science*, 32, 351-354.
- 5844 Connolly, R. D. 1998. Modelling effects of soil structure on the water balance of soil-crop
5845 systems: A review. *Soil and Tillage Research*, 48, 1-19.
- 5846 Cornelis, W. M. & Gabriels, D. 2004. A simple model for the prediction of the deflation
5847 threshold shear velocity of dry loose particles. *Sedimentology*, 51, 39-51.
- 5848 Craine, J. M., Morrow, C. & Stock, W. D. 2008. Nutrient concentration ratios and co-
5849 limitation in South African grasslands. *New Phytologist*, 179, 829-836.
- 5850 Cresswell, H., Smiles, D. & Williams, J. 1992. Soil structure, soil hydraulic properties and the
5851 soil water balance. *Australian Journal of Soil Research*, 30, 265-283.
- 5852 Curran, P. J., Dungan, J. L., Macler, B. A. & Plummer, S. E. 1991. The effect of a red leaf
5853 pigment on the relationship between red edge and chlorophyll concentration.
5854 *Remote Sensing of Environment*, 35, 69-76.
- 5855 D'elia, C. F., Steudler, P. A. & Corwin, N. 1977. Determination of total nitrogen in aqueous
5856 samples using persulfate digestion. *Limnology and Oceanography*, 22, 760-764.
- 5857 D'odorico, P., Caylor, K., Okin, G. S. & Scanlon, T. M. 2007. On soil moisture-vegetation
5858 feedbacks and their possible effects on the dynamics of dryland ecosystems. *J.*
5859 *Geophys. Res.*, 112, G04010.
- 5860 D'acqui, L. P., Pucci, A. & Janik, L. J. 2010. Soil properties prediction of western
5861 mediterranean islands with similar climatic environments by means of mid-infrared
5862 diffuse reflectance spectroscopy. *European Journal of Soil Science*, 61, 865-876.
- 5863 Da Silva, R. & Rosolem, C. 2003. Early development and nutrition of cover crop species as
5864 affected by soil compaction. *Journal of Plant Nutrition*, 26, 1635.
- 5865 Daly, E. & Porporato, A. 2005. A review of soil moisture dynamics: From rainfall infiltration
5866 to ecosystem response. *Environmental Engineering Science*, 22, 9-24.
- 5867 Daryanto, S., Eldridge, D. J. & Wang, L. 2013. Ploughing and grazing alter the spatial
5868 patterning of surface soils in a shrub-encroached woodland. *Geoderma*, 200-201,
5869 67-76.
- 5870 Datt, B., Mcvicar, T. R., Van Niel, T. G., Jupp, D. L. B. & Pearlman, J. S. 2003. Preprocessing
5871 EO-1 Hyperion hyperspectral data to support the application of agricultural indexes.
5872 *Geoscience and Remote Sensing, IEEE Transactions on*, 41, 1246-1259.
- 5873 Daughtry, C. S. T., Hunt, E. R., Jr., Doraiswamy, P. C., McMurtrey, J. E., Iii & Russ, A. L. 2003.
5874 Remote sensing of crop residue cover and soil tillage intensity. In: *Geoscience and*
5875 *Remote Sensing Symposium*, 2003. IGARSS '03. Proceedings. 2003 IEEE
5876 International, 2003. 2192-2194 vol.4.
- 5877 Daughtry, C. S. T., Hunt, E. R. & McMurtrey, J. E. 2004. Assessing crop residue cover using
5878 shortwave infrared reflectance. *Remote Sensing of Environment*, 90, 126-134.
- 5879 Daughtry, C. S. T., Serbin, G., Reeves, J. B., Doraiswamy, P. C. & Hunt, E. R. 2009. Wheat
5880 straw composition and spectral reflectance changes during decomposition. In:
5881 *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2009.
5882 WHISPERS '09. First Workshop on, 2009. 1-4.
- 5883 Dawson, T. P. & Curran, P. J. 1998. A new technique for interpolating the reflectance red
5884 edge position. *International Journal of Remote Sensing*, 19, 2133-2139.

- 5885 De Baets, S., Poesen, J., Knapen, A. & Galindo, P. 2007. Impact of root architecture on the
5886 erosion-reducing potential of roots during concentrated flow. *Earth Surface*
5887 *Processes and Landforms*, 32, 1323-1345.
- 5888 De Boer, W. & Kowalchuk, G. A. 2001. Nitrification in acid soils: Micro-organisms and
5889 mechanisms. *Soil Biology and Biochemistry*, 33, 853-866.
- 5890 De Jong, S. M. 1991. The use of spectroscopy to monitor soils and soil crusts. *In: Geoscience*
5891 *and Remote Sensing Symposium*, 1991. IGARSS '91. Remote Sensing: Global
5892 Monitoring for Earth Management., International, 1991. 1187-1189.
- 5893 De Jong, S. M., Addink, E. A., Van Beek, L. P. H. & Duijsings, D. 2011. Physical
5894 characterization, spectral response and remotely sensed mapping of mediterranean
5895 soil surface crusts. *CATENA*, 86, 24-35.
- 5896 Demattê, J. a. M., Sousa, A. A., Alves, M. C., Nanni, M. R., Fiorio, P. R. & Campos, R. C. 2006.
5897 Determining soil water status and other soil characteristics by spectral proximal
5898 sensing. *Geoderma*, 135, 179-195.
- 5899 Deshmukh, I. K. 1984. A common relationship between precipitation and grassland peak
5900 biomass for east and Southern Africa. *African Journal of Ecology*, 22, 181-186.
- 5901 Dexter, A. R. & Czyż, E. A. 2007. Applications of s-theory in the study of soil physical
5902 degradation and its consequences. *Land Degradation & Development*, 18, 369-381.
- 5903 Dijkstra, F. A., Pendall, E., Morgan, J. A., Blumenthal, D. M., Carrillo, Y., Lecain, D. R., Follett,
5904 R. F. & Williams, D. G. 2012. Climate change alters stoichiometry of phosphorus and
5905 nitrogen in a semiarid grassland. *New Phytologist*, 196, 807-815.
- 5906 Dixon, R. M. 1995. Water infiltration control at the soil surface: Theory and practice. *Journal*
5907 *of Soil and Water Conservation*, 50, 450.
- 5908 Dlamini, P., Orchard, C., Jewitt, G., Lorentz, S., Titshall, L. & Chaplot, V. 2011. Controlling
5909 factors of sheet erosion under degraded grasslands in the sloping lands of KwaZulu-
5910 Natal, South Africa. *Agricultural Water Management*, 98, 1711-1718.
- 5911 Dojani, S., Büdel, B., Deutschewitz, K. & Weber, B. 2011. Rapid succession of biological soil
5912 crusts after experimental disturbance in the succulent karoo, South Africa. *Applied*
5913 *Soil Ecology*, 48, 263-269.
- 5914 Drake, N. A., Mackin, S. & Settle, J. J. 1999. Mapping vegetation, soils, and geology in
5915 semiarid shrublands using spectral matching and mixture modeling of SWIR AVIRIS
5916 imagery. *Remote Sensing of Environment*, 68, 12-25.
- 5917 Dultz, S. & Kühn, P. 2005. Occurrence, formation, and micromorphology of gypsum in soils
5918 from the central-German Chernozem region. *Geoderma*, 129, 230-250.
- 5919 Dutkiewicz, A., Lewis, M. & Ostendorf, B. 2009a. Evaluation and comparison of
5920 hyperspectral imagery for mapping surface symptoms of dryland salinity.
5921 *International Journal of Remote Sensing*, 30, 693-719.
- 5922 Dutkiewicz, A., Lewis, M. & Ostendorf, B. 2006. Mapping surface symptoms of dryland
5923 salinity with hyperspectral imagery. *ISPRS Commission VII Mid-term Symposium*
5924 *"Remote Sensing: From Pixels to Processes"*. Enschede, the Netherlands.
- 5925 Dye, P. J. & Walker, B. H. 1987. Patterns of shoot growth in a semi-arid grassland in
5926 Zimbabwe. *Journal of Applied Ecology*, 24, 633-644.
- 5927 Egboka, B. C. E. & Orajaka, I. P. 1987. Soil and gully erosion models for effective control
5928 programmes. *Geoforum*, 18, 333-341.
- 5929 Eklundh, L. 1998. Estimating relations between AVHRR NDVI and rainfall in East Africa at 10-
5930 day and monthly time scales. *International Journal of Remote Sensing*, 19, 563-570.

- 5931 Eldridge, D. J., Zaady, E. & Shachak, M. 2000. Infiltration through three contrasting biological
5932 soil crusts in patterned landscapes in the Negev, Israel. *CATENA*, 40, 323-336.
- 5933 Eldridge, D. J., Zaady, E. & Shachak, M. 2002. Microphytic crusts, shrub patches and water
5934 harvesting in the Negev desert: The *shikim* system. *Landscape Ecology*, 17, 587-597.
- 5935 Ellis, J. F., Ngigi, S., Rex, H. G., Allan, K., Human, C., Sutton, M., Geel, C., Grond, E.,
5936 Weiersbye, I. M., Kotze, D. J., Merwe, M. V. D. & Labuschagne, P. 2009. Anglogold
5937 Ashanti Ltd, Vaal River operations, environmental management programme.
- 5938 Elvidge, C. D. 1990. Visible and near infrared reflectance characteristics of dry plant
5939 materials. *International Journal of Remote Sensing*, 11, 1775-1795.
- 5940 Engman, E. T. 1984. Remote sensing based continuous hydrologic modeling. *Advances in*
5941 *Space Research*, 4, 201-209.
- 5942 Epstein, H. E., Lauenroth, W. K., Burke, I. C. & Coffin, D. P. 1996. Ecological responses of
5943 dominant grasses along two climatic gradients in the great plains of the United
5944 States. *Journal of Vegetation Science*, 7, 777-788.
- 5945 Eva, H. & Lambin, E. F. 1998a. Burnt area mapping in Central Africa using ATSR data.
5946 *International Journal of Remote Sensing*, 19, 3473-3497.
- 5947 Eva, H. & Lambin, E. F. 1998b. Remote sensing of biomass burning in tropical regions:
5948 Sampling issues and multisensor approach. *Remote Sensing of Environment*, 64, 292-
5949 315.
- 5950 Farifteh, J., Tolpekin, V., Van Der Meer, F. & Sukchan, S. 2010. Salinity modelling by inverted
5951 gaussian parameters of soil reflectance spectra. *International Journal of Remote*
5952 *Sensing*, 31, 3195 - 3210.
- 5953 Farifteh, J., Van Der Meer, F., Atzberger, C. & Carranza, E. J. M. 2007. Quantitative analysis
5954 of salt-affected soil reflectance spectra: A comparison of two adaptive methods
5955 (PLSR and ANN). *Remote Sensing of Environment*, 110, 59-78.
- 5956 Farrar, T. J., Nicholson, S. E. & Lare, A. R. 1994. The influence of soil type on the relationships
5957 between NDVI, rainfall, and soil moisture in semiarid Botswana. ii. NDVI response to
5958 soil oisture. *Remote Sensing of Environment*, 50, 121-133.
- 5959 Fernández, O. A., Gil, M. E. & Distel, R. A. 2009. The challenge of rangeland degradation in a
5960 temperate semiarid region of argentina: The caldenal. *Land Degradation &*
5961 *Development*, 20, 431-440.
- 5962 Fidêncio, P. H., Poppi, R. J., De Andrade, J. C. & Cantarella, H. 2002. Determination of organic
5963 matter in soil using near-infrared spectroscopy and partial least squares regression.
5964 *Communications in Soil Science and Plant Analysis*, 33, 1607-1615.
- 5965 Filella, I. & Penuelas, J. 1994. The red edge position and shape as indicators of plant
5966 chlorophyll content, biomass and hydric status. *International Journal of Remote*
5967 *Sensing*, 15, 1459 - 1470.
- 5968 Fischer, J. & Lindenmayer, D. B. 2007. Landscape modification and habitat fragmentation: A
5969 synthesis. *Global Ecology & Biogeography*, 16, 265-280.
- 5970 Forman, R. T. T. & Godron, M. 1981. Patches and structural components for a landscape
5971 ecology. *Bioscience*, 31, 733-740.
- 5972 Foster, J. R., Kingdon, C. C. & Townsend, P. A. 2002. Predicting tropical forest carbon from
5973 EO-1 hyperspectral imagery in Noel Kempff Mercado National Park, Bolivia. *In:*
5974 *Geoscience and Remote Sensing Symposium*, 2002. IGARSS '02. 2002 IEEE
5975 International, 2002. 3108-3110 vol.6.

- 5976 Fourty, T., Baret, F., Jacquemoud, S., Schmuck, G. & Verdebout, J. 1996. Leaf optical
5977 properties with explicit description of its biochemical composition: Direct and
5978 inverse problems. *Remote Sensing of Environment*, 56, 104-117.
- 5979 Frank, I. E. & Friedman, J. H. 1993. A statistical view of some chemometrics regression tools.
5980 *Technometrics*, 35, 109-135.
- 5981 Fuhlendorf, S. D., Briske, D. D. & Smeins, F. E. 2001. Herbaceous vegetation change in
5982 variable rangeland environments: The relative contribution of grazing and climatic
5983 variability. *Applied Vegetation Science*, 4, 177-188.
- 5984 Furniss, D., Weiersbye, I., Tongway, D., Stark, R., Margalit, N., Nel, H., Grond, E. &
5985 Witkowski, E. T. 2009. Deriving indices of landscape function from spectral
5986 reflectance of grassland and savanna on gold mines in South Africa. *Geoscience and*
5987 *Remote Sensing Symposium, IEEE International, IGARSS*.
- 5988 Furniss, D. G. 2008. *Can indices of landscape function analysis (LFA) be derived from ground-*
5989 *based spectroscopy? A case study from gold mines on the highveld of South Africa.*
5990 Master of Science, University of the Witwatersrand.
- 5991 Fynn, R. W. S., Morris, C. D. & Kirkman, K. P. 2005. Plant strategies and trait trade-offs
5992 influence trends in competitive ability along gradients of soil fertility and
5993 disturbance. *Journal of Ecology*, 93, 384-394.
- 5994 Gaitán, J. J., Bran, D., Oliva, G., Ciari, G., Nakamatsu, V., Salomone, J., Ferrante, D., Buono,
5995 G., Massara, V., Humano, G., Celdrán, D., Opazo, W. & Maestre, F. T. 2013.
5996 Evaluating the performance of multiple remote sensing indices to predict the spatial
5997 variability of ecosystem structure and functioning in Patagonian steppes. *Ecological*
5998 *Indicators*, 34, 181-191.
- 5999 Gallardo, A. & Schlesinger, W. H. 1992. Carbon and nitrogen limitations of soil microbial
6000 biomass in desert ecosystems. *Biogeochemistry*, 18, 1-17.
- 6001 Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P.,
6002 Asner, G. P., Cleveland, C. C., Green, P. A., Holland, E. A., Karl, D. M., Michaels, A. F.,
6003 Porter, J. H., Townsend, A. R. & Vöosmarty, C. J. 2004. Nitrogen cycles: Past, present,
6004 and future. *Biogeochemistry*, 70, 153-226.
- 6005 García-Gómez, M. & Maestre, F. T. 2011. Remote sensing data predict indicators of soil
6006 functioning in semi-arid steppes, central Spain. *Ecological Indicators*, 11, 1476-1481.
- 6007 Garnier, L. K. M. & Dajoz, I. 2001. The influence of fire on the demography of a dominant
6008 grass species of West African savannas, *Hyparrhenia diplandra*. *Journal of Ecology*,
6009 89, 200-208.
- 6010 Geladi, P. & Kowalski, B. R. 1986a. An example of 2-block predictive partial least-squares
6011 regression with simulated data. *Analytica Chimica Acta*, 185, 19-32.
- 6012 Geladi, P. & Kowalski, B. R. 1986b. Partial least-squares regression: A tutorial. *Analytica*
6013 *Chimica Acta*, 185, 1-17.
- 6014 Georgiadis, N. J., Ruess, R. W., Mcnaughton, S. J. & Western, D. 1989. Ecological conditions
6015 that determine when grazing stimulates grass production. *Oecologia*, 81, 316-322.
- 6016 Gibbens, R. P., Tromble, J. M., Hennessy, J. T. & Cardenas, M. 1983. Soil movement in
6017 mesquite dunelands and former grasslands of southern New Mexico from 1933 to
6018 1980. *Journal of Range Management*, 36, 145-148.
- 6019 Giraldo, M. A., Bosch, D., Madden, M., Usery, L. & Kvien, C. 2008. Landscape complexity and
6020 soil moisture variation in South Georgia, USA, for remote sensing applications.
6021 *Journal of Hydrology*, 357, 405-420.

- 6022 Gitelson, A. & Merzlyak, M. N. 1994. Quantitative estimation of chlorophyll-a using
6023 reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of*
6024 *Photochemistry and Photobiology B: Biology*, 22, 247-252.
- 6025 Gitelson, A. A., Kaufman, Y. J. & Merzlyak, M. N. 1996. Use of a green channel in remote
6026 sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58,
6027 289-298.
- 6028 Gitelson, A. A. & Merzlyak, M. N. 1996. Signature analysis of leaf reflectance spectra:
6029 Algorithm development for remote sensing of chlorophyll. *Journal of Plant*
6030 *Physiology*, 148, 494-500.
- 6031 Gitelson, A. A., Zur, Y., Chivkunova, O. B. & Merzlyak, M. N. 2002. Assessing carotenoid
6032 content in plant leaves with reflectance spectroscopy. *Photochemistry and*
6033 *Photobiology*, 75, 272-281.
- 6034 Goldshleger, N., Ben-Dor, E., Lugassi, R. & Eshel, G. 2010. Soil degradation monitoring by
6035 remote sensing: Examples with three degradation processes. *Soil Sci. Soc. Am. J.*, 74,
6036 1433-1445.
- 6037 Gomez, C., Viscarra Rossel, R. A. & Mcbratney, A. B. 2008. Soil organic carbon prediction by
6038 hyperspectral remote sensing and field vis-nir spectroscopy: An australian case
6039 study. *Geoderma*, 146, 403-411.
- 6040 Goodenough, D. G., Dyk, A., Niemann, K. O., Pearlman, J. S., Hao, C., Tian, H., Murdoch, M. &
6041 West, C. 2003. Processing Hyperion and ALI for forest classification. *Geoscience and*
6042 *Remote Sensing, IEEE Transactions on*, 41, 1321-1331.
- 6043 Greene, R. 1992. Soil physical properties of three geomorphic zones in a semi-arid mulga
6044 woodland. *Australian Journal of Soil Research*, 30, 55 - 69.
- 6045 Greene, R. S. B. & Hairsine, P. B. 2004. Elementary processes of soil–water interaction and
6046 thresholds in soil surface dynamics: A review. *Earth Surface Processes and*
6047 *Landforms*, 29, 1077-1091.
- 6048 Greig-Smith, P. 1979. Pattern in vegetation. *Journal of Ecology*, 67, 755.
- 6049 Guerrero-Campo, J. & Montserrat-Martí, G. 2000. Effects of soil erosion on the floristic
6050 composition of plant communities on marl in northeast Spain. *Journal of Vegetation*
6051 *Science*, 11, 329-336.
- 6052 Guerrero-Campo, J. & Montserrat-Martí, G. 2004. Comparison of floristic changes on
6053 vegetation affected by different levels of soil erosion in Miocene clays and Eocene
6054 marls from northeast Spain. *Plant Ecology*, 173, 83-93.
- 6055 Guntiñas, M. E., Leirós, M. C., Trasar-Cepeda, C. & Gil-Sotres, F. 2012. Effects of moisture
6056 and temperature on net soil nitrogen mineralization: A laboratory study. *European*
6057 *Journal of Soil Biology*, 48, 73-80.
- 6058 Gustafson, E. J. 1998. Quantifying landscape spatial pattern: What is the state of the art?
6059 *Ecosystems*, 1, 143-156.
- 6060 Haines, B. L. 1977. Nitrogen uptake. Apparent pattern during old field succession in
6061 southeastern U.S. *Oecologia*, 26, 295-303.
- 6062 Han, G., Hao, X., Zhao, M., Wang, M., Ellert, B. H., Willms, W. & Wang, M. 2008. Effect of
6063 grazing intensity on carbon and nitrogen in soil and vegetation in a meadow steppe
6064 in Inner Mongolia. *Agriculture, Ecosystems & Environment*, 125, 21-32.
- 6065 Hartanto, H., Prabhu, R., Widayat, A. S. E. & Asdak, C. 2003. Factors affecting runoff and soil
6066 erosion: Plot-level soil loss monitoring for assessing sustainability of forest
6067 management. *Forest Ecology and Management*, 180, 361-374.

- 6068 Hassan, S. N. 2011. Influence of early dry season fires on primary production in western
6069 Serengeti grasslands, Tanzania. *Open Journal of Ecology*, 1, 24-34.
- 6070 Hattingh, E. R. 1953. Observations on the ecology of *seriphium plumosum* levyns. *Empire*
6071 *Journal of Experimental Agriculture*, 21, 84.
- 6072 Heal, O. W. & Harrison, A. F. 1990. *Keynote paper: Turnover of nutrients: A technological*
6073 *challenge*, London, Elsevier Applied Science.
- 6074 Heilman, J. L. & Boyd, W. E. 1986. Soil background effects on the spectral response of a
6075 three-component rangeland scene. *Remote Sensing of Environment*, 19, 127-137.
- 6076 Heinen, M. 2006. Simplified denitrification models: Overview and properties. *Geoderma*,
6077 133, 444-463.
- 6078 Herrero, J., Artieda, O. & Hudnall, W. H. 2009. Gypsum, a tricky material. *Soil Science Society*
6079 *of America Journal*, 73, 1757-1763.
- 6080 Hijmans, R. J. 2014. Raster: Geographic data analysis and modeling. R package version 2.2-
6081 31.
- 6082 Hino, M., Fujita, K. & Shutto, H. 1987. A laboratory experiment on the role of grass for
6083 infiltration and runoff processes. *Journal of Hydrology*, 90, 303-325.
- 6084 Hirche, A., Salamani, M., Abdellaoui, A., Benhouhou, S. & Valderrama, J. 2011. Landscape
6085 changes of desertification in arid areas: The case of south-west Algeria.
6086 *Environmental Monitoring and Assessment*, 179, 403-420.
- 6087 Hoeft, I., Keuter, A., Quiñones, C. M., Schmidt-Walter, P., Veldkamp, E. & Corre, M. D. 2014.
6088 Nitrogen retention efficiency and nitrogen losses of a managed and phytodiverse
6089 temperate grassland. *Basic and Applied Ecology*.
- 6090 Holland, E. A. & Detling, J. K. 1990. Plant response to herbivory and belowground nitrogen
6091 cycling. *Ecology*, 71, 1040-1049.
- 6092 Hooper, A., Vannelli, T., Bergmann, D. & Arciero, D. 1997. Enzymology of the oxidation of
6093 ammonia to nitrite by bacteria. *Antonie van Leeuwenhoek*, 71, 59-67.
- 6094 Horler, D. N. H., Barber, J. & Barringer, A. R. 1980. Effects of heavy metals on the absorbance
6095 and reflectance spectra of plants *International Journal of Remote Sensing*, 1, 121-
6096 136.
- 6097 Houlton, B. Z. W. Y.-P. P. M. F. C. B. 2008. A unifying framework for dinitrogen fixation in the
6098 terrestrial biosphere. *Nature*, 454, 327-330.
- 6099 Huete, A. R. 1986. Separation of soil-plant spectral mixtures by factor analysis. *Remote*
6100 *Sensing of Environment*, 19, 237-251.
- 6101 Huete, A. R. & Escadafal, R. 1991. Assessment of biophysical soil properties through spectral
6102 decomposition techniques. *Remote Sensing of Environment*, 35, 149-159.
- 6103 Huete, A. R. & Jackson, R. D. 1987. Suitability of spectral indices for evaluating vegetation
6104 characteristics on arid rangelands. *Remote Sensing of Environment*, 23, 213-232, IN1-
6105 IN8.
- 6106 Huete, A. R., Miura, T. & Xiang, G. 2003. Land cover conversion and degradation analyses
6107 through coupled soil-plant biophysical parameters derived from hyperspectral EO-1
6108 Hyperion. *Geoscience and Remote Sensing, IEEE Transactions on*, 41, 1268-1276.
- 6109 Iversen, J. D. & White, B. R. 1982. Saltation threshold on earth, mars and venus.
6110 *Sedimentology*, 29, 111-119.
- 6111 Jackson, R. B. & Caldwell, M. M. 1993. Geostatistical patterns of soil heterogeneity around
6112 individual perennial plants. *Journal of Ecology*, 81, 683-692.

- 6113 Jaeger, C. H., Monson, R. K., Fisk, M. C. & Schmidt, S. K. 1999. Seasonal partitioning of
6114 nitrogen by plants and soil microorganisms in an alpine ecosystem. *Ecology*, 80,
6115 1883-1891.
- 6116 Jalali, M., Mahdavi, S. & Ranjbar, F. 2014. Nitrogen, phosphorus and sulfur mineralization as
6117 affected by soil depth in rangeland ecosystems. *Environmental Earth Sciences*, 1-14.
- 6118 Jarrell, W. M., Armstrong, D. E., Grigal, D. F., Kelly, E. F., Monger, H. C. & Wedin, D. A. 1999.
6119 Soil water and temperature status. In: ROBERTSON, G. P., COLEMAN, D. C., BLEDSE, C. S. & SOLLINS, P. (eds.) *Standard soil methods for long-term ecological research*.
6120 Oxford: Oxford University Press Inc.
- 6121
- 6122 Jenkinson, D. S., Andrew, S. P. S., Lynch, J. M., Goss, M. J. & Tinker, P. B. 1990. The turnover
6123 of organic carbon and nitrogen in soil [and discussion]. *Philosophical Transactions:
6124 Biological Sciences*, 329, 361-368.
- 6125 Jenkinson, D. S. & Rayner, J. H. 1977. The turnover of soil organic matter in some of the
6126 rothamsted classical experiments. *Soil Science*, 123, 298-305.
- 6127 Ji, L. & Peters, A. J. 2003. Assessing vegetation response to drought in the northern great
6128 plains using vegetation and drought indices. *Remote Sensing of Environment*, 87, 85-
6129 98.
- 6130 Jobbágy, E. G. & Jackson, R. 2001. The distribution of soil nutrients with depth: Global
6131 patterns and the imprint of plants. *Biogeochemistry*, 53, 51-77.
- 6132 Jobbágy, E. G. & Jackson, R. B. 2000. The vertical distribution of soil organic carbon and its
6133 relation to climate and vegetation. *Ecological Applications*, 10, 423-436.
- 6134 Joeckel, R. M., Ang Clement, B. J. & Vanfleet Bates, L. R. 2005. Sulfate-mineral crusts from
6135 pyrite weathering and acid rock drainage in the dakota formation and graneros
6136 shale, jefferson county, nebraska. *Chemical Geology*, 215, 433-452.
- 6137 Jordaan, D. 2009. Bankruptbush (slangbos) – a silent threat to grasslands? *Grassroots*, 9, 40-
6138 42.
- 6139 Jordan, C. F. 1969. Derivation of leaf-area index from quality of light on the forest floor.
6140 *Ecology*, 50, 663-666.
- 6141 Kakembo, V. 2009. Vegetation patchiness and implications for landscape function: The case
6142 of pteronia incana invader species in ngqushwa rural municipality, eastern cape,
6143 South Africa. *CATENA*, 77, 180-186.
- 6144 Kalwij, J. M., De Boer, W. F., Mucina, L., Prins, H. H. T., Skarpe, C. & Winterbach, C. 2010.
6145 Tree cover and biomass increase in a southern African savanna despite growing
6146 elephant population. *Ecological Applications*, 20, 222-233.
- 6147 Kätterer, T., Reichstein, M., Andrén, O. & Lomander, A. 1998. Temperature dependence of
6148 organic matter decomposition: A critical review using literature data analyzed with
6149 different models. *Biology and Fertility of Soils*, 27, 258-262.
- 6150 Keller, L. P., McCarthy, G. J. & Richardson, J. L. 1986. Mineralogy and stability of soil
6151 evaporites in north dakota. *Soil Science Society of America Journal*, 50, 1069-1071.
- 6152 Kelley, K. R. & Stevenson, F. J. 1996. Organic forms of n in soil. In: PICCOLO, A. (ed.) *Humic
6153 substances in terrestrial ecosystems*. Amsterdam, The Netherlands: ELSEVIER
6154 SCIENCE B.V.
- 6155 Kelly, R. D. & Walker, B. H. 1976. The effects of different forms of land use on the ecology of
6156 a semi-arid region in south-eastern Rhodesia. *Journal of Ecology*, 64, 553-576.
- 6157 Klotz, M. G. & Stein, L. Y. 2008. Nitrifier genomics and evolution of the nitrogen cycle. *FEMS
6158 Microbiology Letters*, 278, 146-156.

- 6159 Knox, N. M., Skidmore, A. K., Van Der Werff, H. M. A., Groen, T. A., De Boer, W. F., Prins, H.
6160 H. T., Kohi, E. & Peel, M. 2013. Differentiation of plant age in grasses using remote
6161 sensing. *International Journal of Applied Earth Observation and Geoinformation*, 24,
6162 54-62.
- 6163 Köchy, M. & Wilson, S. D. 1997. Litter decomposition and nitrogen dynamics in aspen forest
6164 and mixed-grass prairie. *Ecology*, 78, 732-739.
- 6165 Koerner, S. & Collins, S. 2013. Small-scale patch structure in North American and South
6166 African grasslands responds differently to fire and grazing. *Landscape Ecology*, 28,
6167 1293-1306.
- 6168 Kooistra, L., Salas, E. a. L., Clevers, J. G. P. W., Wehrens, R., Leuven, R. S. E. W., Nienhuis, P.
6169 H. & Buydens, L. M. C. 2004. Exploring field vegetation reflectance as an indicator of
6170 soil contamination in river floodplains. *Environmental Pollution*, 127, 281-290.
- 6171 Kooistra, L., Wehrens, R., Leuven, R. S. E. W. & Buydens, L. M. C. 2001. Possibilities of visible-
6172 near-infrared spectroscopy for the assessment of soil contamination in river
6173 floodplains. *Analytica Chimica Acta*, 446, 97-105.
- 6174 Kuemmerle, T., Röder, A. & Hill, J. 2006. Separating grassland and shrub vegetation by
6175 multirate pixel-adaptive spectral mixture analysis. *International Journal of Remote
6176 Sensing*, 27, 3251-3271.
- 6177 Kumar, L., Rietkerk, M., Van Langevelde, F., Van De Koppel, J., Van Andel, J., Hearne, J., De
6178 Ridder, N., Stroosnijder, L., Skidmore, A. K. & Prins, H. H. T. 2002. Relationship
6179 between vegetation growth rates at the onset of the wet season and soil type in the
6180 sahel of burkina faso: Implications for resource utilisation at large scales. *Ecological
6181 Modelling*, 149, 143-152.
- 6182 Kumar, L., Schmidt, K., Dury, S. & Skidmore, A. 2010. Imaging spectrometry and vegetation
6183 science In: VAN DER MEER, F. & DE JONG, S. (eds.) *Imaging spectroscopy: Basic
6184 principles and prospective applications*. Enschede, Netherlands: ITC, International
6185 Institute for Geo-Information Science and Earth Observation.
- 6186 Kusumo, B., Hedley, M., Hedley, C. & Tuohy, M. 2011. Measuring carbon dynamics in field
6187 soils using soil spectral reflectance: Prediction of maize root density, soil organic
6188 carbon and nitrogen content. *Plant and Soil*, 338, 233-245.
- 6189 Labate, D., Ceccherini, M., Cisbani, A., De Cosmo, V., Galeazzi, C., Giunti, L., Melozzi, M.,
6190 Pieraccini, S. & Stagi, M. 2009. The prisma payload optomechanical design, a high
6191 performance instrument for a new hyperspectral mission. *Acta Astronautica*, 65,
6192 1429-1436.
- 6193 Lacy, H. W. B., File, T. & Biggs, B. 2008. Ecosystem function analysis: Measuring and
6194 monitoring for mine closure and completion in Australia and abroad - 1994 to 2008.
6195 In: FOURIE, A., TIBBETT, M., WEIERSBYE, I. & DYE, P. (eds.) *Mine Closure 2008*.
6196 Johannesburg, South Africa: Australian Centre for Geomechanics, Perth.
- 6197 Ladd, J. N., Parsons, J. W. & Amato, M. 1977. Studies of nitrogen immobilization and
6198 mineralization in calcareous soils—i: Distribution of immobilized nitrogen amongst
6199 soil fractions of different particle size and density. *Soil Biology and Biochemistry*, 9,
6200 309-318.
- 6201 Lagacherie, P., Baret, F., Feret, J.-B., Madeira Netto, J. & Robbez-Masson, J. M. 2008.
6202 Estimation of soil clay and calcium carbonate using laboratory, field and airborne
6203 hyperspectral measurements. *Remote Sensing of Environment*, 112, 825-835.

- 6204 Lakshmi, V., Bolten, J., Guha, A. & Zehrhuhs, D. 2001. Remote sensing of soil moisture at
6205 various spatial scales. *In: Geoscience and Remote Sensing Symposium, 2001. IGARSS*
6206 *'01. IEEE 2001 International, 2001. 9-11 vol.1.*
- 6207 Lal, R. 2001. Soil degradation by erosion. *Land Degradation & Development*, 12, 519-539.
- 6208 Lau, I. C., Hewson, R. D., Ong, C. C. H. & Tongway, D. J. 2008. Remote mine site rehabilitation
6209 monitoring using airborne hyperspectral imaging and landscape function analysis
6210 (LFA). *In: CHEN, J., JIANG, J. & PELED, A., eds. The International Archives of the*
6211 *Photogrammetry, Remote Sensing and Spatial Information Sciences, 2008. 325 - 330.*
- 6212 Lauenroth, W. K. & Adler, P. B. 2008. Demography of perennial grassland plants: Survival,
6213 life expectancy and life span. *Journal of Ecology*, 96, 1023-1032.
- 6214 Lauenroth, W. K. & Coffin, D. P. 1992. Belowground processes and the recovery of semiarid
6215 grasslands from disturbance. *In: WALL, M. K. (ed.) Ecosystem rehabilitation: Vol 2.*
6216 *Ecosystem analysis and synthesis. The Hague, The Netherlands: SPB Academic*
6217 *Publishing.*
- 6218 Lebauer, D. S. & Treseder, K. K. 2008. Nitrogen limitation of net primary productivity in
6219 terrestrial ecosystems is globally distributed. *Ecology*, 89, 371-379.
- 6220 Lecatsas, G. 1962. Solar radiation as a factor in the establishment of *Stoebe vulgaris*. *South*
6221 *African Journal of Science*, 58, 304-306.
- 6222 Lee, C. M., Cable, M. L., Hook, S. J., Green, R. O., Ustin, S. L., Mandl, D. J. & Middleton, E. M.
6223 2015. An introduction to the NASA hyperspectral infrared imager (HYSPIRI) mission
6224 and preparatory activities. *Remote Sensing of Environment*, 167, 6-19.
- 6225 Lee, X., Wu, H.-J., Sigler, J., Oishi, C. & Siccama, T. 2004. Rapid and transient response of soil
6226 respiration to rain. *Global Change Biology*, 10, 1017-1026.
- 6227 Legates, D. R., Mahmood, R., Levia, D. F., Deliberty, T. L., Quiring, S. M., Houser, C. & Nelson,
6228 F. E. 2011. Soil moisture: A central and unifying theme in physical geography.
6229 *Progress in Physical Geography*, 35, 65-86.
- 6230 Legout, C., Legu  dois, S., Le Bissonnais, Y. & Malam Issa, O. 2005. Splash distance and size
6231 distributions for various soils. *Geoderma*, 124, 279-292.
- 6232 Levin, S. A. & Paine, R. T. 1974. Disturbance, patch formation, and community structure.
6233 *Proceedings of the National Academy of Sciences of the United States of America*, 71,
6234 2744-2747.
- 6235 Lewis, M. 1998. Numeric classification as an aid to spectral mapping of vegetation
6236 communities. *Plant Ecology*, 136, 133-133.
- 6237 Li, X. Y., Hu, X., Zhang, Z. H., Peng, H. Y., Zhang, S. Y., Li, G. Y., Li, L. & Ma, Y. J. 2013. Shrub
6238 hydopedology: Preferential water availability to deep soil layer. *Vadose Zone*
6239 *Journal*, 12.
- 6240 Liaw, Y. P., Sisterson, D. L. & Miller, N. L. 1990. Comparison of field, laboratory, and
6241 theoretical estimates of global nitrogen fixation by lightning. *Journal of Geophysical*
6242 *Research: Atmospheres*, 95, 22489-22494.
- 6243 Liu, H., Lei, T. W., Zhao, J., Yuan, C. P., Fan, Y. T. & Qu, L. Q. 2011. Effects of rainfall intensity
6244 and antecedent soil water content on soil infiltrability under rainfall conditions using
6245 the run off-on-out method. *Journal of Hydrology*, 396, 24-32.
- 6246 Lloyd, J. & Taylor, J. A. 1994. On the temperature dependence of soil respiration. *Functional*
6247 *Ecology*, 8, 315-323.
- 6248 L  pez-Pamo, E., Barettino, D., Ant  n-Pacheco, C., Ortiz, G., Arr  nz, J. C., Gumiel, J. C.,
6249 Mart  nez-Pledel, B., Aparicio, M. & Montouto, O. 1999. The extent of the Aznalc  llar

- pyritic sludge spill and its effects on soils. *Science of The Total Environment*, 242, 57-88.
- Lottermoser, B. G., Ashley, P. M. & Lawie, D. C. 1999. Environmental geochemistry of the Gulf Creek copper mine area, north-eastern New South Wales, Australia. *Environmental Geology*, 39, 61-74.
- Lovett, G. M., Jones, C. G., Turner, M. G. & Weathers, K. C. 2005. Ecosystem function in heterogeneous landscapes. In: LOVETT, G. M., JONES, C. G., TURNER, M. G. & WEATHERS, K. C. (eds.) *Ecosystem function in heterogeneous landscapes*. USA: Springer Science+Business Media Inc.
- Lucas, Y., Schmitt, A. D., Chabaux, F., Clément, A., Fritz, B., Elsass, P. & Durand, S. 2010. Geochemical tracing and hydrogeochemical modelling of water-rock interactions during salinization of alluvial groundwater (upper Rhine Valley, France). *Applied Geochemistry*, 25, 1644-1663.
- Ludwig, J. A., Wiens, J. A. & Tongway, D. J. 2000. A scaling rule for landscape patches and how it applies to conserving soil resources in savannas. *Ecosystems*, 3, 84-97.
- Macdonald, I. a. W. 1978. Pattern and process in a semi-arid grassveld in Rhodesia. *Proceedings of the Annual Congresses of the Grassland Society of Southern Africa*, 13, 103-109.
- Maestre, F. T. & Cortina, J. 2004. Insights into ecosystem composition and function in a sequence of degraded semiarid steppes. *Restoration Ecology*, 12, 494-502.
- Mark, H. & Workman, J. J. 2007a. Calculating the solution for regression techniques: Part 3 -- partial least squares regression made simple. *Chemometrics in spectroscopy*. Amsterdam: Academic Press.
- Mark, H. & Workman, J. J. 2007b. A simple question: The meaning of chemometrics pondered. *Chemometrics in spectroscopy*. Amsterdam: Academic Press.
- Marnewecke, G. & Kotze, D. 1999. Appendix w6: Guidelines for delineation of wetland boundary and wetland zones. In: MACKAY, H. (ed.) *Water resources protection policy implementation*. Department of Water Affairs and Forestry.
- Mars, J. C. & Crowley, J. K. 2003. Mapping mine wastes and analyzing areas affected by selenium-rich water runoff in southeast Idaho using AVIRIS imagery and digital elevation data. *Remote Sensing of Environment*, 84, 422-436.
- Marschner, H. 1995. *Mineral nutrition of higher plants*, London, Academic Press Limited.
- Marsett, R. C., Qi, J., Heilman, P., Sharon, H. B., Watson, M. C., Amer, S., Weltz, M., Goodrich, D. & Marsett, R. 2006. Remote sensing for grassland management in the arid southwest. *Rangeland Ecology & Management*, 59, 530-540.
- Martens, H. 2001. Reliable and relevant modelling of real world data: A personal account of the development of PLS regression. *Chemometrics and Intelligent Laboratory Systems*, 58, 85-95.
- Martens, H. & Næs, T. 1989. *Multivariate calibration*, Chichester, John Wiley & Sons.
- Mary, B., Beaudoin, N., Justes, E. & Machet, J. M. 1999. Calculation of nitrogen mineralization and leaching in fallow soil using a simple dynamic model. *European Journal of Soil Science*, 50, 549-566.
- Mayor, Á. G. & Bautista, S. 2012. Multi-scale evaluation of soil functional indicators for the assessment of water and soil retention in mediterranean semiarid landscapes. *Ecological Indicators*, 20, 332-336.
- Mcgroddy, M. E., Daufresne, T. & Hedin, L. O. 2004. Scaling of C:N:P stoichiometry in forests worldwide: Implications of terrestrial redfield-type ratios. *Ecology*, 85, 2390-2401.

- 6297 Melillo, J., Aber, J., Linkins, A., Ricca, A., Fry, B. & Nadelhoffer, K. 1989. Carbon and nitrogen
6298 dynamics along the decay continuum: Plant litter to soil organic matter. *Plant and*
6299 *Soil*, 115, 189-198.
- 6300 Melillo, J. M., Aber, J. D. & Muratore, J. F. 1982. Nitrogen and lignin control of hardwood
6301 leaf litter decomposition dynamics. *Ecology*, 63, 621-626.
- 6302 Menon, M., Hermle, S., Günthardt-Goerg, M. S. & Schulin, R. 2007. Effects of heavy metal
6303 soil pollution and acid rain on growth and water use efficiency of a young model
6304 forest ecosystem. *Plant & Soil*, 297, 171-183.
- 6305 Meti & NASA 2009. ASTER GDEM: Global digital elevation model. Ministry of Economy,
6306 Trade, and Industry (METI) of Japan and the United States National Aeronautics and
6307 Space Administration (NASA).
- 6308 Mevik, B.-H. & Cederkvist, H. R. 2004. Mean squared error of prediction (MSEP) estimates
6309 for principal component regression (PCR) and partial least squares regression (PLSR).
6310 *Journal of Chemometrics*, 18, 422-429.
- 6311 Mevik, B.-H. & Wehrens, R. 2007. The pls package: Principal component and partial least
6312 squares regression in r. *Journal of statistical software*, 18, 1-24.
- 6313 Meza-Figueroa, D., Maier, R. M., De La O-Villanueva, M., Gómez-Alvarez, A., Moreno-
6314 Zazueta, A., Rivera, J., Campillo, A., Grandlic, C. J., Anaya, R. & Palafox-Reyes, J. 2009.
6315 The impact of unconfined mine tailings in residential areas from a mining town in a
6316 semi-arid environment: Nacozari, Sonora, Mexico. *Chemosphere*, 77, 140-147.
- 6317 Miller, J. R., Hare, E. W. & Wu, J. 1990. Quantitative characterization of the vegetation red
6318 edge reflectance 1. An inverted-gaussian reflectance model. *International Journal of*
6319 *Remote Sensing*, 11, 1755 - 1773.
- 6320 Minasny, B. & Mcbratney, A. 2013. Why you don't need to use RPD. *Pedometron*, 33, 14-15.
- 6321 Moore, J. C., Mccann, K. & De Ruiter, P. C. 2005. Modeling trophic pathways, nutrient
6322 cycling, and dynamic stability in soils. *Pedobiologia*, 49, 499-510.
- 6323 Mooshammer, M., Wanek, W., Hämmerle, I., Fuchslueger, L., Hofhansl, F., Knoltsch, A.,
6324 Schneckner, J., Takriti, M., Watzka, M., Wild, B., Keiblinger, K. M., Zechmeister-
6325 Boltenstern, S. & Richter, A. 2014. Adjustment of microbial nitrogen use efficiency to
6326 carbon:nitrogen imbalances regulates soil nitrogen cycling. *Nat Commun*, 5.
- 6327 Morales, V. L., Parlange, J. Y. & Steenhuis, T. S. 2010. Are preferential flow paths
6328 perpetuated by microbial activity in the soil matrix? A review. *Journal of Hydrology*,
6329 393, 29-36.
- 6330 Morecroft, M. D., Sellers, E. K. & Lee, J. A. 1994. An experimental investigation into the
6331 effects of atmospheric nitrogen deposition on two semi-natural grasslands. *Journal*
6332 *of Ecology*, 82, 475-483.
- 6333 Moreno, T., Oldroyd, A., Mcdonald, I. & Gibbons, W. 2007. Preferential fractionation of trace
6334 metals–metalloids into pm₁₀ resuspended from contaminated gold mine tailings at
6335 Rodalquilar, Spain. *Water, Air & Soil Pollution*, 179, 93-105.
- 6336 Mucina, L. & Rutherford, M. C. 2006. *The vegetation of South Africa, Lesotho and*
6337 *Swaziland.*, Pretoria, South African National Biodiversity Institute.
- 6338 Mulvaney, R. L. 1996. Nitrogen—inorganic forms. In: BIGHAM, J. M. (ed.) *Methods of soil*
6339 *analysis part 3: Chemical methods*. Madison, WI.: Soil Science Society of America,
6340 Inc.
- 6341 Nagler, P. L., Daughtry, C. S. T. & Goward, S. N. 2000. Plant litter and soil reflectance.
6342 *Remote Sensing of Environment*, 71, 207-215.

- Naicker, K., Cukrowska, E. & McCarthy, T. S. 2003. Acid mine drainage arising from gold mining activity in Johannesburg, South Africa and environs. *Environmental Pollution*, 122, 29-40.
- National Geo-spatial Information (NGI), 2010, Standard for Ortho-Rectification and Mosaicking of Imagery. Department of Rural Development and Land Reform. Republic of South Africa.
<http://www.ngi.gov.za/index.php/technical-information/standards-menu>
- Naumann, J., Anderson, J. & Young, D. 2010. Remote detection of plant physiological responses to TNT soil contamination. *Plant and Soil*, 329, 239-248.
- Naumann, J. C., Anderson, J. E. & Young, D. R. 2008. Linking physiological responses, chlorophyll fluorescence and hyperspectral imagery to detect salinity stress using the physiological reflectance index in the coastal shrub, *Myrica cerifera*. *Remote Sensing of Environment*, 112, 3865-3875.
- Nawar, S., Buddenbaum, H., Hill, J. & Kozak, J. 2014. Modeling and mapping of soil salinity with reflectance spectroscopy and Landsat data using two quantitative methods (PLSR and MARS). *Remote Sensing*, 6, 10813-10834.
- Neff, J. C., Reynolds, R. L., Belnap, J. & Lamothe, P. 2005. Multi-decadal impacts of grazing on soil physical and biogeochemical properties in southeast Utah. *Ecological Applications*, 15, 87-95.
- Nicholson, S. E. & Farrar, T. J. 1994. The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana. I. NDVI response to rainfall. *Remote Sensing of Environment*, 50, 107-120.
- Nippert, J., Knapp, A. & Briggs, J. 2006. Intra-annual rainfall variability and grassland productivity: Can the past predict the future? *Plant Ecology*, 184, 65-74.
- Noy-Meir, I. 1973. Desert ecosystems: Environment and producers. *Annual Review of Ecology & Systematics*, 4, 25-51.
- Noy-Meir, I. 1981. Spatial effects in modelling of arid ecosystems. In: GOODALL, D. W. & PERRY, R. A. (Eds.) *Arid-land ecosystems: Structure, functioning and management*. Sydney: Cambridge University Press.
- Numata, I. 2012. Characterization on pastures using field and imaging spectrometers. In: PRASAD S. THENKABAIL, JOHN G. LYON & ALFREDO HUETE (eds.) *Hyperspectral remote sensing of vegetation*. Boca Raton: CRC Press.
- Numata, I., Roberts, D. A., Chadwick, O. A., Schimel, J., Sampaio, F. R., Leonidas, F. C. & Soares, J. V. 2007. Characterization of pasture biophysical properties and the impact of grazing intensity using remotely sensed data. *Remote Sensing of Environment*, 109, 314-327.
- Numata, I., Soares, J. V., Roberts, D. A., Leonidas, F. C., Chadwick, O. A. & Batista, G. T. 2003. Relationships among soil fertility dynamics and remotely sensed measures across pasture chronosequences in Rondônia, Brazil. *Remote Sensing of Environment*, 87, 446-455.
- O'connor, T. G., Haines, L. M. & Snyman, H. A. 2001. Influence of precipitation and species composition on phytomass of a semi-arid African grassland. *Journal of Ecology*, 89, 850-860.
- O'connor, T. G. & Bredenkamp, G. J. 1997. Grassland. In: COWLING, R. M., RICHARDSON, D. M. & PIERCE, S. M. (eds.) *Vegetation of southern Africa*. Cambridge University Press.
- Oades, J. 1984. Soil organic matter and structural stability: Mechanisms and implications for management. *Plant and Soil*, 76, 319-337.

- 6390 Oades, J. M. 1993. The role of biology in the formation, stabilization and degradation of soil
6391 structure. *Geoderma*, 56, 377-400.
- 6392 Okin, G. S., Gillette, D. A. & Herrick, J. E. 2006. Multi-scale controls on and consequences of
6393 aeolian processes in landscape change in arid and semi-arid environments. *Journal of*
6394 *Arid Environments*, 65, 253-275.
- 6395 Oliveira, V., Makeschin, F., Sano, E. & Lorz, C. 2014. Physical and chemical analyses of bare
6396 soil sites in western central Brazil: A case study. *Environmental Earth Sciences*, 1-9.
- 6397 Ong, C., Lau, I., Hewson, R., Tongway, D. J. & Eisele, A. 2008. Meriwa project M385: The
6398 development of a hyperspectral environmental measurement tool for monitoring
6399 mining related infrastructure and rehabilitation. CSIRO, Australia.
- 6400 Ong, C. C. H. & Cudahy, T. J. 2014. Mapping contaminated soils: Using remotely-sensed
6401 hyperspectral data to predict pH. *European Journal of Soil Science*, 65, 897-906.
- 6402 Pannkuk, T. R., Aitkenhead-Peterson, J. A., Steinke, K., Thomas, J. C., Chalmers, D. R. &
6403 White, R. H. 2011. Carbon, nitrogen, and orthophosphate leaching from soil under
6404 single- and mixed-species landscapes. *HortScience*, 46, 1533-1539.
- 6405 Paruelo, J. M. & Lauenroth, W. K. 1995. Regional patterns of normalized difference
6406 vegetation index in North American shrublands and grasslands. *Ecology*, 76, 1888-
6407 1898.
- 6408 Paul, E. A., Harris, D., Klug, M. J. & Ruess, R. W. 1999. The determination of microbial
6409 biomass. In: ROBERTSON, G. P., COLEMAN, D. C., BLEDSOE, C. S. & SOLLINS, P. (eds.)
6410 *Standard soil methods for long-term ecological research*.
- 6411 Pearlman, J., Carman, S., Segal, C., Jarecke, P., Clancy, P. & Browne, W. 2001. Overview of
6412 the Hyperion imaging spectrometer for the NASA EO-1 mission. In: Geoscience and
6413 Remote Sensing Symposium, 2001. IGARSS '01, 2001. 3036-3038.
- 6414 Peltre, C., Thuriès, L., Barthès, B., Brunet, D., Morvan, T., Nicolardot, B., Parnaudeau, V. &
6415 Houot, S. 2011. Near infrared reflectance spectroscopy: A tool to characterize the
6416 composition of different types of exogenous organic matter and their behaviour in
6417 soil. *Soil Biology and Biochemistry*, 43, 197-205.
- 6418 Peñuelas, J., Filella, I., Biel, C., Serrano, L. & Savé, R. 1993. The reflectance at the 950-970
6419 nm region as an indicator of plant water status. *International Journal of Remote*
6420 *Sensing*, 14, 1887-1905.
- 6421 Peters, A. J., Eve, M. D., Holt, E. H. & Whitford, W. G. 1997. Analysis of desert plant
6422 community growth patterns with high temporal resolution satellite spectra. *Journal*
6423 *of Applied Ecology*, 34, 418-432.
- 6424 Petropoulos, G. P., Arvanitis, K. & Sigrimis, N. 2012. Hyperion hyperspectral imagery analysis
6425 combined with machine learning classifiers for land use/cover mapping. *Expert*
6426 *Systems with Applications*, 39, 3800-3809.
- 6427 Pickup, G., Bastin, G. N. & Chewings, V. H. 1998. Identifying trends in land degradation in
6428 non-equilibrium rangelands. *Journal of Applied Ecology*, 35, 365.
- 6429 Picotte, J. J. & Robertson, K. 2011. Timing constraints on remote sensing of wildland fire
6430 burned area in the southeastern US. *Remote Sensing*, 3, 1680-1690.
- 6431 Pockman, W. & Small, E. 2010. The influence of spatial patterns of soil moisture on the grass
6432 and shrub responses to a summer rainstorm in a Chihuahuan desert ecotone.
6433 *Ecosystems*, 13, 511-525.
- 6434 Porporato, A., D'odorico, P., Laio, F. & Rodriguez-Iturbe, I. 2003. Hydrologic controls on soil
6435 carbon and nitrogen cycles. I. Modeling scheme. *Advances in Water Resources*, 26,
6436 45-58.

- 6437 Poth, M. 1986. Dinitrogen production from nitrite by a *Nitrosomonas* isolate. *Applied and*
6438 *Environmental Microbiology*, 52, 957-959.
- 6439 Poth, M. & Focht, D. D. 1985. ^{15}N kinetic analysis of N_2O production by *Nitrosomonas*
6440 *europaea*: An examination of nitrifier denitrification. *Applied and Environmental*
6441 *Microbiology*, 49, 1134-1141.
- 6442 Preissler, H. & Loercher, G. 1995. Extraction of soil properties from laboratory and imaging
6443 spectrometry data. In: Geoscience and Remote Sensing Symposium, 1995. IGARSS
6444 '95. 'Quantitative Remote Sensing for Science and Applications', 1995. 1968-1970.
- 6445 Prescott, C. E., Maynard, D. G. & Laiho, R. 2000. Humus in northern forests: Friend or foe?
6446 *Forest Ecology and Management*, 133, 23-36.
- 6447 Prins, H. H. T. 1988. Plant phenology patterns in Lake Manyara National Park, Tanzania.
6448 *Journal of Biogeography*, 15, 465-480.
- 6449 Prins, H. H. T. & Loth, P. E. 1988. Rainfall patterns as background to plant phenology in
6450 northern Tanzania. *Journal of Biogeography*, 15, 451-463.
- 6451 Prosser, J. I. & Nicol, G. W. 2008. Relative contributions of archaea and bacteria to aerobic
6452 ammonia oxidation in the environment. *Environmental Microbiology*, 10, 2931-2941.
- 6453 Pyke, D. A., Herrick, J. E., Shaver, P. & Pellant, M. 2002. Rangeland health attributes and
6454 indicators for qualitative assessment. *Journal of Range Management*, 55, 584-597.
- 6455 R Core Team 2013. R: A language and environment for statistical computing. Vienna,
6456 Austria.: R Foundation for Statistical Computing.
- 6457 Ravi, S., D'odorico, P. & Okin, G. S. 2007. Hydrologic and aeolian controls on vegetation
6458 patterns in arid landscapes. *Geophysical Research Letters*, 34, L24S23.
- 6459 Redfield, A. C. 1958. The biological control of chemical factors in the environment. *American*
6460 *Scientist*, 46, 230A-221.
- 6461 Reich, P. B., Oleksyn, J. & Tilman, G. D. 2004. Global patterns of plant leaf N and P in relation
6462 to temperature and latitude. *Proceedings of the National Academy of Sciences of the*
6463 *United States of America*, 101, 11001-11006.
- 6464 Rezaei, S. A., Arzani, H. & Tongway, D. 2006. Assessing rangeland capability in Iran using
6465 landscape function indices based on soil surface attributes. *Journal of Arid*
6466 *Environments*, 65, 460-473.
- 6467 Riaz, M., Mian, I. A. & Cresser, M. S. 2011. How much does $\text{NH}_4\text{-N}$ contribute to mineral-N
6468 losses in N-impacted acid soils under grassland in the UK? A microcosm study.
6469 *Chemistry and Ecology*, 28, 25-36.
- 6470 Richard, Y. & Pocard, I. 1998. A statistical study of NDVI sensitivity to seasonal and
6471 interannual rainfall variations in Southern Africa. *International Journal of Remote*
6472 *Sensing*, 19, 2907-2920.
- 6473 Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G. & Green, R. O. 1998. Mapping
6474 chaparral in the Santa Monica mountains using multiple endmember spectral
6475 mixture models. *Remote Sensing of Environment*, 65, 267-279.
- 6476 Roberts, D. A., Green, R. O. & Adams, J. B. 1997. Temporal and spatial patterns in vegetation
6477 and atmospheric properties from AVIRIS. *Remote Sensing of Environment*, 62, 223-
6478 240.
- 6479 Roberts, D. A., Smith, M. O. & Adams, J. B. 1993. Green vegetation, nonphotosynthetic
6480 vegetation, and soils in AVIRIS data. *Remote Sensing of Environment*, 44, 255-269.
- 6481 Rock, B. N., Hoshizaki, T. & Miller, J. R. 1988. Comparison of in situ and airborne spectral
6482 measurements of the blue shift associated with forest decline. *Remote Sensing of*
6483 *Environment*, 24, 109-127.

- 6484 Rogass, C., Mielke, C., Scheffler, D., Boesche, N. K., Lausch, A., Lubitz, C., Brell, M., Spengler,
6485 D., Eisele, A., Segl, K. & Guanter, L. 2014. Reduction of uncorrelated striping noise-
6486 applications for hyperspectral pushbroom acquisitions. *Remote Sensing*, 6, 11082-
6487 11106.
- 6488 Roux, E. 1969a. Bankrotbos. In: ROUX, W. M. (ed.) *Grass: A story of Frankenwald*. Cape
6489 Town: Oxford University Press.
- 6490 Roux, E. 1969b. Secondary succession: Later views. In: ROUX, W. M. (ed.) *Grass: A story of*
6491 *Frankenwald*. Cape Town: Oxford University Press.
- 6492 Royston, J. P. 1982. Algorithm as 181: The w test for normality. *Applied Statistics*, 31, 176-
6493 180.
- 6494 Rudorff, C. M., Galvão, L. S. & Novo, E. M. L. M. 2009. Reflectance of floodplain waterbodies
6495 using EO-1 Hyperion data from high and receding flood periods of the Amazon River.
6496 *International Journal of Remote Sensing*, 30, 2713 - 2720.
- 6497 Sala, O. E., Parton, W. J., Joyce, L. A. & Lauenroth, W. K. 1988. Primary production of the
6498 central grassland region of the United States. *Ecology*, 69, 40-45.
- 6499 Sanger, J. E. 1971. Quantitative investigations of leaf pigments from their inception in buds
6500 through autumn coloration to decomposition in falling leaves. *Ecology*, 52, 1075-
6501 1089.
- 6502 Schaaf, W., Gast, M., Wilden, R., Scherzer, J., Blechschmidt, R. & Hüttel, R. 1999. Temporal
6503 and spatial development of soil solution chemistry and element budgets in different
6504 mine soils of the Lusatian lignite mining area. *Plant and Soil*, 213, 169-179.
- 6505 Scheffler, D. & Karrasch, P. 2013. Preprocessing of hyperspectral images: A comparative
6506 study of destriping algorithms for EO-1 Hyperion. In: BRUZZONE, L., ed. Proc. SPIE,
6507 Image and Signal Processing for Remote Sensing XIX, 2013. 88920H-88920H-15.
- 6508 Scheffler, D. & Karrasch, P. 2014. Destriping of hyperspectral image data: An evaluation of
6509 different algorithms using EO-1 Hyperion data. *Journal of Applied Remote Sensing*, 8,
6510 083645-083645.
- 6511 Schimel, J. P. & Bennett, J. 2004. Nitrogen mineralization: Challenges of a changing
6512 paradigm. *Ecology*, 85, 591-602.
- 6513 Schlesinger, W. H. & Pilmanis, A. M. 1998. Plant-soil interactions in deserts.
6514 *Biogeochemistry*, 42, 169-187.
- 6515 Schlesinger, W. H., Raikes, J. A., Hartley, A. E. & Cross, A. F. 1996. On the spatial pattern of
6516 soil nutrients in desert ecosystems. *Ecology*, 77, 364-374.
- 6517 Schmidtlein, S. & Sassini, J. 2004. Mapping of continuous floristic gradients in grasslands
6518 using hyperspectral imagery. *Remote Sensing of Environment*, 92, 126-138.
- 6519 Schomberg, H. H., Wietholter, S., Griffin, T. S., Reeves, D. W., Cabrera, M. L., Fisher, D. S.,
6520 Endale, D. M., Novak, J. M., Balkcom, K. S., Raper, R. L., Kitchen, N. R., Locke, M. A.,
6521 Potter, K. N., Schwartz, R. C., Truman, C. C. & Tyler, D. D. 2009. Assessing indices for
6522 predicting potential nitrogen mineralization in soils under different management
6523 systems. *Soil Science Society of America Journal*, 73, 1575-1586.
- 6524 Scurlock, J. M. O., Johnson, K. & Olson, R. J. 2002. Estimating net primary productivity from
6525 grassland biomass dynamics measurements. *Global Change Biology*, 8, 736-753.
- 6526 Serrano, L., Peñuelas, J. & Ustin, S. L. 2002. Remote sensing of nitrogen and lignin in
6527 mediterranean vegetation from AVIRIS data: Decomposing biochemical from
6528 structural signals. *Remote Sensing of Environment*, 81, 355-364.
- 6529 Shachak, M. & Lovett, G. M. 1998. Atmospheric deposition to a desert ecosystem and its
6530 implications for management. *Ecological Applications*, 8, 455-463.

- 6531 Sherman, D. M. & Waite, T. D. 1985. Electronic spectra of Fe³⁺ oxides and oxide hydroxides
6532 in the near IR to near UV. *American Mineralogist*, 70, 1262-1269.
- 6533 Singer, A., Kirsten, W. F. A. & Bühmann, C. 1999. A proposed fog deposition mechanism for
6534 the formation of salt efflorescences in the Mpumalanga highveld, republic of South
6535 Africa. *Water, Air, and Soil Pollution*, 109, 313-325.
- 6536 Skousen, J., Ziemkiewicz, P. & Venable, C. 2006. Tree recruitment and growth on 20-year-
6537 old, unreclaimed surface mined lands in West Virginia. *International Journal of*
6538 *Mining, Reclamation and Environment*, 20, 142 - 154.
- 6539 Smith, B. & McDermid, G. J. 2014. Examination of fire-related succession within the dry
6540 mixed-grass subregion of Alberta with the use of MODIS and Landsat. *Rangeland*
6541 *Ecology & Management (Allen Press Publishing Services Inc.)*, 67, 307-317.
- 6542 Smith, M. O., Ustin, S. L., Adams, J. B. & Gillespie, A. R. 1990a. Vegetation in deserts: I. A
6543 regional measure of abundance from multispectral images. *Remote Sensing of*
6544 *Environment*, 31, 1-26.
- 6545 Smith, M. O., Ustin, S. L., Adams, J. B. & Gillespie, A. R. 1990b. Vegetation in deserts: II.
6546 Environmental influences on regional abundance. *Remote Sensing of Environment*,
6547 31, 27-52.
- 6548 Snyman, H. A. 2003. Fire and the dynamics of semi-arid grassland: Influence on plant
6549 survival, productivity and water-use efficiency. *African Journal of Range & Forage*
6550 *Science*, 20, 29-39.
- 6551 Snyman, H. A. 2004. Short-term response in productivity following an unplanned fire in a
6552 semi-arid rangeland of South Africa. *Journal of Arid Environments*, 56, 465-485.
- 6553 Snyman, H. A. 2006. Estimating grassland production loss due to fire for a semi-arid climate.
6554 In: Proc. 41st Congress of the South African Society for Animal Science, 2006. South
6555 African Society for Animal Science, 38-41.
- 6556 Snyman, H. A. 2009a. Germination potential of *Seriphium plumosum* (bankrupt bush,
6557 slangbos or vaalbos). *Grassroots: Newsletter of the Grassland Society of Southern*
6558 *Africa*, 9, 43-48.
- 6559 Snyman, H. A. 2009b. A philosophical approach to the distribution and spread of *Seriphium*
6560 *plumosum*. *Grassroots: Newsletter of the Grassland Society of Southern Africa*, 9, 29-
6561 37.
- 6562 Snyman, H. A. 2010. Allelopathic potential, seed ecology and germination of the encroacher
6563 shrub *Seriphium plumosum*. *African Journal of Range & Forage Science*, 27, 29-37.
- 6564 Snyman, H. A. 2011. Short-term response of the encroacher shrub *Seriphium plumosum* to
6565 fire. *African Journal of Range & Forage Science*, 28, 65-77.
- 6566 Snyman, H. A. 2012. Habitat preferences of the encroacher shrub, *Seriphium plumosum*.
6567 *South African Journal of Botany*, 81, 34-39.
- 6568 Snyman, H. A. & Bredenkamp, G. J. 2005. Influence of fire on root distribution, seasonal root
6569 production and root/shoot ratios in grass species in a semi-arid grassland of South
6570 Africa. *South African Journal of Botany*, 71, 133-144.
- 6571 Sollins, P., Glassman, C., Paul, E. A., Swanston, C., Lajtha, K., Heil, J. W. & Elliott, E. T. 1999.
6572 Soil carbon and nitrogen: Pools and fractions. In: ROBERTSON, G. P., COLEMAN, D. C.,
6573 BLEDSOE, C. S. & SOLLINS, P. (eds.) *Standard soil methods for long-term ecological*
6574 *research*. New York: Oxford University Press.
- 6575 Sparling, G. P. 1997. Soil microbial biomass, activity and nutrient cycling as indicators of soil
6576 health. In: PANKHURST, C., DOUBE, B. M. & GUPTA, V. V. S. R. (eds.) *Biological*
6577 *indicators of soil health*. Wallingford, UK: CAB INTERNATIONAL.

- 6578 Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M. & Wetterlind, J. 2010. Visible and near
6579 infrared spectroscopy in soil science. *In: DONALD, L. S. (ed.) Advances in agronomy.*
6580 Academic Press.
- 6581 Steven, M. D., Malthus, T. J. & Clark, J. A. 1992. Achievements and unresolved problems in
6582 vegetation monitoring. *In: MATHER, P. M. (ed.) Terra-1: Understanding the terrestrial*
6583 *environment. The role of earth observation from space.* London: Taylor & Francis.
- 6584 Stevens, A., Van Wesemael, B., Bartholomeus, H., Rosillon, D., Tychon, B. & Ben-Dor, E.
6585 2008. Laboratory, field and airborne spectroscopy for monitoring organic carbon
6586 content in agricultural soils. *Geoderma*, 144, 395-404.
- 6587 Stevens, A., Wesemael, B. V., Vandenschrick, G., Touré, S. & Tychon, B. 2006. Detection of
6588 carbon stock change in agricultural soils using spectroscopic techniques. *Soil Science*
6589 *Society of America Journal*, 70, 844.
- 6590 Stuffer, T., Förster, K., Hofer, S., Leipold, M., Sang, B., Kaufmann, H., Penné, B., Mueller, A.
6591 & Chlebek, C. 2009. Hyperspectral imaging—an advanced instrument concept for the
6592 ENMAP mission (environmental mapping and analysis programme). *Acta*
6593 *Astronautica*, 65, 1107-1112.
- 6594 Stuffer, T., Kaufmann, C., Hofer, S., Förster, K. P., Schreier, G., Mueller, A., Eckardt, A., Bach,
6595 H., Penné, B., Benz, U. & Haydn, R. 2007. The ENMAP hyperspectral imager—an
6596 advanced optical payload for future applications in earth observation programmes.
6597 *Acta Astronautica*, 61, 115-120.
- 6598 Swayze, G. A., Smith, K. S., Clark, R. N., Sutley, S. J., Pearson, R. M., Vance, J. S., Hageman, P.
6599 L., Briggs, P. H., Meier, A. L., Singleton, M. J. & Roth, S. 2000. Using imaging
6600 spectroscopy to map acidic mine waste. *Environmental Science & Technology*, 34, 47-
6601 54.
- 6602 Tainton, N. 1999. *Veld management in South Africa*, Pietermaritzburg, University of Natal
6603 Press.
- 6604 Tainton, N. M. & Mentis, M. T. 1984. Fire in grassland. *In: BOOYSEN, P. D. V. & TANTON, N.*
6605 *M. (eds.) Ecological effects of fire in South African ecosystems: Ecological studies 48.*
6606 Berlin: Springer-Verlag.
- 6607 Terhoeven-Urselmans, T., Vagen, T.-G., Spaargaren, O. & Shepherd, K. D. 2010. Prediction of
6608 soil fertility properties from a globally distributed soil mid-infrared spectral library.
6609 *Soil Sci. Soc. Am. J.*, 74, 1792-1799.
- 6610 Tongway, D. J. & Hindley, N. L. 2004. *Landscape function analysis: Procedures for monitoring*
6611 *and assessing landscapes with special reference to minesites and rangelands*, CSIRO
6612 Sustainable Ecosystems.
- 6613 Tongway, D. J. & Ludwig, J. A. 1996. Rehabilitation of semiarid landscapes in Australia. I.
6614 Restoring productive soil patches. *Restoration Ecology*, 4, 388-397.
- 6615 Tucker, C. J. 1977. Spectral estimation of grass canopy variables. *Remote Sensing of*
6616 *Environment*, 6, 11-26.
- 6617 Tucker, C. J. 1978. Post senescent grass canopy remote sensing. *Remote Sensing of*
6618 *Environment*, 7, 203-210.
- 6619 Tucker, C. J. 1979. Red and photographic infrared linear combinations for monitoring
6620 vegetation. *Remote Sensing of Environment*, 8, 127-150.
- 6621 Tutu, H., McCarthy, T. S. & Cukrowska, E. 2008. The chemical characteristics of acid mine
6622 drainage with particular reference to sources, distribution and remediation: The
6623 Witwatersrand basin, South Africa as a case study. *Applied Geochemistry*, 23, 3666–
6624 3684.

- 6625 Ungar, S. G. 2001. Overview of EO-1, the first 120 days. *In: Geoscience and Remote Sensing*
6626 *Symposium, IGARSS '01*, 2001. 43-45.
- 6627 Ungar, S. G., Pearlman, J. S., Mendenhall, J. A. & Reuter, D. 2003. Overview of the earth
6628 observing one (EO-1) mission. *Geoscience and Remote Sensing, IEEE Transactions on*,
6629 41, 1149-1159.
- 6630 Ustin, S. L., Hart, Q. J., Scheer, G. & Duan, L. 1994. Estimating dry grass biomass residues
6631 using AVIRIS image analysis. *In: Geoscience and Remote Sensing Symposium, 1994.*
6632 *IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis*
6633 *and Interpretation*, 1994. 1211-1212.
- 6634 Van Der Meer, F. D. 2010. Basic physics of spectrometry. *Imaging spectroscopy: Basic*
6635 *principles and prospective applications*. ITC: International Institute for Geo-
6636 Information Science and Earth Observation.
- 6637 Van Der Voet, H. 1994. Comparing the predictive accuracy of models using a simple
6638 randomization test. *Chemometrics and Intelligent Laboratory Systems*, 25, 313-323.
- 6639 Van Der Willigen, C., Pammenter, N. W., Mundree, S. & Farrant, J. 2001. Some physiological
6640 comparisons between the resurrection grass, *Eragrostis nindensis*, and the related
6641 desiccation-sensitive species, *E. curvula*. *Plant Growth Regulation*, 35, 121-129.
- 6642 Van Oudtshoorn, F. 1999. *Guide to grasses of Southern Africa*, Pretoria, South Africa, Briza
6643 Publications.
- 6644 Van Wyk, E. & Van Oudtshoorn, F. 1999. *Guide to grasses of South Africa*, Pretoria, South
6645 Africa, Briza Publications.
- 6646 Verchot, L. V., Groffman, P. M. & Frank, D. A. 2002. Landscape versus ungulate control of
6647 gross mineralization and gross nitrification in semi-arid grasslands of Yellowstone
6648 National Park. *Soil Biology and Biochemistry*, 34, 1691-1699.
- 6649 Viscarra Rossel, R. A. & Behrens, T. 2010. Using data mining to model and interpret soil
6650 diffuse reflectance spectra. *Geoderma*, 158, 46-54.
- 6651 Viscarra Rossel, R. A., Bui, E. N., De Caritat, P. & McKenzie, N. J. 2010. Mapping iron oxides
6652 and the color of Australian soil using visible and near-infrared reflectance spectra. *J.*
6653 *Geophys. Res.*, 115, F04031.
- 6654 Viscarra Rossel, R. A., Walvoort, D. J. J., Mcbratney, A. B., Janik, L. J. & Skjemstad, J. O. 2006.
6655 Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for
6656 simultaneous assessment of various soil properties. *Geoderma*, 131, 59-75.
- 6657 Vitousek, P. M., Gosz, J. R., Grier, C. C., Melillo, J. M., Reiners, W. A. & Todd, R. L. 1979.
6658 Nitrate losses from disturbed ecosystems. *Science*, 204, 469-474.
- 6659 Vossbrinck, C. R., Coleman, D. C. & Woolley, T. A. 1979. Abiotic and biotic factors in litter
6660 decomposition in a semiarid grassland. *Ecology*, 60, 265-271.
- 6661 Wang, J.-J., Zhang, Y. & Bussink, C. 2014. Unsupervised multiple endmember spectral
6662 mixture analysis-based detection of opium poppy fields from an EO-1 Hyperion
6663 image in Helmand, Afghanistan. *Science of The Total Environment*, 476-477, 1-6.
- 6664 Wang, J., Fu, B., Qiu, Y. & Chen, L. 2001. Soil nutrients in relation to land use and landscape
6665 position in the semi-arid small catchment on the loess plateau in China. *Journal of*
6666 *Arid Environments*, 48, 537-550.
- 6667 Watt, A. S. 1947. Pattern and process in the plant community. *Journal of Ecology*, 35, 1-22.
- 6668 Weber, B., Olehowski, C., Knerr, T., Hill, J., Deutschewitz, K., Wessels, D. C. J., Eitel, B. &
6669 Büdel, B. 2008. A new approach for mapping of biological soil crusts in semidesert
6670 areas with hyperspectral imagery. *Remote Sensing of Environment*, 112, 2187-2201.

- 6671 Wedin, D. A. 1995. Species, nitrogen, and grassland dynamics: The constraints of stuff. *In:*
 6672 JONES, C. G. & LAWTON, J. H. (eds.) *Linking species and ecosystems*. New York:
 6673 Chapman & Hall, Inc.
- 6674 Weiersbye, I. M. & Witkowski, E. T. F. 2002. Seed fate and practical germination methods for
 6675 46 perennial species that colonize gold mine tailings and acid mine drainage-polluted
 6676 soils in the grassland biome. *In:* J, J., BESTER, SEYDACK, A. H. W., VORSTER, T., VAN
 6677 DER MERWE, I. J. & DZIVHANI, S. (eds.) *Multiple use management of natural forests
 6678 & woodlands: Policy refinements and scientific progress III.*: Department of Water
 6679 Affairs and Forestry of South Africa.
- 6680 Weiersbye, I. M. & Witkowski, E. T. F. 2003. Acid rock drainage (ARD) from gold tailings
 6681 dams on the Witwatersrand basin impacts on tree seed fate, viability, inorganic
 6682 content and seedling morphology. *In:* ARMSTRONG, D., VILLIERS, A. B. D.,
 6683 KLEINMANN, R. L. P., MCCARTHY, T. S. & NORTON, P. J., eds. *Proceedings of the 8th
 6684 International Congress on Mine Water & the Environment, 2003 Johannesburg, South
 6685 Africa*.
- 6686 Weiersbye, I. M. & Witkowski, E. T. F. 2007. Impacts of acid mine drainage on the
 6687 regeneration potential of highveld phreatophytes. *In:* J, J., BESTER, SEYDACK, A. H.
 6688 W., VORSTER, T., VAN DER MERWE, I. J. & DZIVHANI, S. (eds.) *Multiple use
 6689 management of natural forests and woodlands: policy refinements and scientific
 6690 progress IV*. Department of Water Affairs and Forestry of South Africa.
- 6691 Weiersbye, I. M., Witkowski, E. T. F. & Reichardt, M. 2006. Floristic composition of gold and
 6692 uranium tailings dams, and adjacent polluted areas, on South Africa's deep-level
 6693 mines. *Bothalia*, 36, 101-127.
- 6694 Weng, Y., Gong, P. & Zhu, Z. 2008. Soil salt content estimation in the Yellow River delta with
 6695 satellite hyperspectral data. *Canadian Journal of Remote Sensing*, 34, 259-270.
- 6696 Wessman, C. A. 1991. Remote sensing of soil processes. *Agriculture, Ecosystems &
 6697 Environment*, 34, 479-493.
- 6698 Wetterlind, J., Stenberg, B. & Söderström, M. 2010. Increased sample point density in farm
 6699 soil mapping by local calibration of visible and near infrared prediction models.
 6700 *Geoderma*, 156, 152-160.
- 6701 White, D. 2011. Hyperion tools 2.0.
- 6702 White, H. P. & Abuelgasim, A. 2010. Monitoring environmental remediation: Hyperspectral
 6703 mapping of re-vegetated areas affected by smelting operations in Sudbury, Canada.
 6704 *In:* *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing
 6705 (WHISPERS)*, 2010 2nd Workshop on, 2010. 1-4.
- 6706 White, R. E. 2006. *Principles and practice of soil science : The soil as a natural resource, 4th
 6707 ed.*, Malden, MA, Blackwell Publishing Company.
- 6708 Wiens, J. A. 1976. Population responses to patchy environments. *Annual Review of Ecology
 6709 & Systematics*, 7, 81-120.
- 6710 Wight, J. P., Ashworth, A. J. & Allen, F. L. 2016. Organic substrate, clay type, texture, and
 6711 water influence on nir carbon measurements. *Geoderma*, 261, 36-43.
- 6712 Wilcox, B. P., Wood, M. K. & Tromble, J. M. 1988. Factors influencing infiltrability of semiarid
 6713 mountain slopes. *Journal of Range Management*, 41, 197-206.
- 6714 Wilke, B.-M. 2005. Determination of chemical and physical soil properties. *In:* MARGESIN, R.
 6715 & SCHINNER, F. (eds.) *Manual for soil analysis – monitoring and assessing soil
 6716 bioremediation*. Berlin: Springer-Verlag.

- 6717 Williamson, T. N., Graham, R. C. & Shouse, P. J. 2004. Effects of a chaparral-to-grass
6718 conversion on soil physical and hydrologic properties after four decades. *Geoderma*,
6719 123, 99-114.
- 6720 Wilson, M. D., Ustin, S. L. & Rocke, D. M. 2004. Classification of contamination in salt marsh
6721 plants using hyperspectral reflectance. *Geoscience and Remote Sensing, IEEE*
6722 *Transactions on*, 42, 1088-1095.
- 6723 Wiltshire, G. H. 1973. Response of grasses to nitrogen source. *Journal of Applied Ecology*, 10,
6724 429-435.
- 6725 Winde, F. & Sandham, L. 2004. Uranium pollution of South African streams – an overview of
6726 the situation in gold mining areas of the Witwatersrand. *GeoJournal*, 61, 131-149.
- 6727 Wold, S., Sjöström, M. & Eriksson, L. 2001. PLS-regression: A basic tool of chemometrics.
6728 *Chemometrics and Intelligent Laboratory Systems*, 58, 109-130.
- 6729 Woodmansee, R. G., Dodd, J. L., Bowman, R. A., Clark, F. E. & Dickinson, C. E. 1978. Nitrogen
6730 budget of a shortgrass prairie ecosystem. *Oecologia*, 34, 363-376.
- 6731 Woodwell, G. M. & Whittaker, R. H. 1968. Primary production in terrestrial ecosystems.
6732 *American Zoologist*, 8, 19-30.
- 6733 Worrall, G. A. 1960. Patchiness in vegetation in the northern Sudan. *Journal of Ecology*, 48,
6734 107-115.
- 6735 Wu, C., Han, X., Niu, Z. & Dong, J. 2010. An evaluation of EO-1 hyperspectral Hyperion data
6736 for chlorophyll content and leaf area index estimation. *International Journal of*
6737 *Remote Sensing*, 31, 1079 - 1086.
- 6738 Xin, Q., Broich, M., Zhu, P. & Gong, P. 2015. Modeling grassland spring onset across the
6739 western United States using climate variables and MODIS-derived phenology
6740 metrics. *Remote Sensing of Environment*, 161, 63-77.
- 6741 Yang, Y., Mohammat, A., Feng, J., Zhou, R. & Fang, J. 2007. Storage, patterns and
6742 environmental controls of soil organic carbon in China. *Biogeochemistry*, 84, 131-
6743 141.
- 6744 Yokoya, N., Miyamura, N. & Iwasaki, A. 2010. Preprocessing of hyperspectral imagery with
6745 consideration of smile and keystone properties. In: Proc. SPIE 7857, Multispectral,
6746 Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques, and
6747 Applications III, 2010. 78570B-78570B-9.
- 6748 Zedler, P. H. 2007. Fire effects on grasslands. In: MIYANISHI, E. A. J. (ed.) *Plant disturbance*
6749 *ecology*. Burlington: Academic Press.
- 6750 Zhang, X., Liao, C., Li, J. & Sun, Q. 2013. Fractional vegetation cover estimation in arid and
6751 semi-arid environments using HJ-1 satellite hyperspectral data. *International Journal*
6752 *of Applied Earth Observation and Geoinformation*, 21, 506-512.
- 6753 Zhou, X., Lin, H. S. & White, E. A. 2008. Surface soil hydraulic properties in four soil series
6754 under different land uses and their temporal changes. *CATENA*, 73, 180-188.
- 6755 Zribi, M., Paris Anguela, T., Duchemin, B., Lili, Z., Wagner, W., Hasenauer, S. & Chehbouni, A.
6756 2010. Relationship between soil moisture and vegetation in the Kairouan plain
6757 region of Tunisia using low spatial resolution satellite data. *WATER RESOURCES*
6758 *RESEARCH*, 46, W06508.
- 6759 Zumft, W. G. 1997. Cell biology and molecular basis of denitrification. *Microbiology and*
6760 *Molecular Biology Reviews*, 61, 533-616.