Statistical modelling of the shrinkage behaviour of South African concretes

Petra Cornelia Gaylard

A research report submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in partial fulfillment of the requirements for the degree of Master of Science.

Johannesburg, 2011

Declaration

I declare that this research report is my own, unaided work. It is being submitted for the Degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

____

Petra Cornelia Gaylard

___________________ day of ___________________ 2011.

Abstract

A hierarchical non-linear model was developed for the time-dependent shrinkage behaviour of South African concretes, from historical laboratory data.

The fit of fifteen growth curve models to the shrinkage-time profiles was evaluated and MCDA was used to identify the best model. The three parameters of the chosen growth curve model were modelled in terms of covariates (relating to concrete raw materials, concrete composition and shrinkage testing conditions) by multivariate multiple regression to produce the WITS model. The model largely conformed to existing knowledge about the factors affecting concrete shrinkage.

Published models for concrete shrinkage were compared to the WITS model regarding their predictive ability with respect to the South African data set. The WITS model performed the best across a variety of graphical and numerical goodness-of-fit measures.

The importance of the study is two-fold:

- The concept of hierarchical non-linear modelling has been applied for the first time to the modelling of the time-dependent properties of concrete.
- This is the first comprehensive model to bring together laboratory data on the shrinkage of concrete generated in South Africa over a span of thirty years.

Acknowledgements

I gratefully acknowledge:

My supervisors, Professors Paul Fatti and Yunus Ballim, for their encouragement, guidance and support

Professor Yunus Ballim for financial assistance for travel

Professor Mark Alexander and Dr Hans Beushausen (Department of Civil Engineering, University of Cape Town) for contributing data to this project and for helpful discussions

Dr Peter Graham (School of Civil and Environmental Engineering, University of the Witwatersrand) for helping me understand the science of concrete better

Mr John Goodman (Cement and Concrete Institute) for contributing data to this project

The staff at the Cement and Concrete Institute's Information Centre for their efficient assistance in obtaining literature references

Contents

List of Figures

List of Tables

List of Symbols

α	ultimate shrinkage parameter in growth curve models
α_1	cement type coefficient in RILEM B3 model
α_2	curing type coefficient in RILEM B3 model
α_3	cement type coefficient in GL2000 model
β	growth rate parameter in growth curve models
γ	shape parameter in growth curve models
$\varepsilon_{sh}(t-t_0)$	mean shrinkage strain in the cross-section ($x10^{-6}$)
$\varepsilon_{sh\infty}$	ultimate shrinkage strain ($x10^{-6}$)
$\varepsilon_{s\infty}$	ultimate shrinkage strain without elastic modulus correction in RILEM B3 model ($x10^{-6}$)
$\varepsilon_{sd}(t-t_0)$	drying shrinkage strain in CEB MC90-99 model
$\varepsilon_{sd,0}$	basic drying shrinkage strain in CEB MC90-99 model
$\varepsilon_{sa}(t-t_0)$	autogenous shrinkage strain in CEB MC90-99 model
$\varepsilon_{sa,0}$	basic autogenous shrinkage strain in CEB MC90-99 model
τ_{sh}	shrinkage half-time (days) = time for shrinkage to reach half of ultimate value
Ψ	mass ratio of sand to total aggregate content
a	cross-sectional area of specimen (mm^2)
A_C	air content per volume of concrete (%)
c	cement content of concrete mix in kg/m^3
D	effective cross-sectional thickness (mm) in RILEM B3 model
E(t)	Young's elastic modulus at age t (GPa)
f_c	28-day standard cylinder compressive strength (MPa)
h	relative humidity (= humidity / 100%)
H	humidity (%)
k_h	humidity dependence term
k_t	coefficient for shrinkage half-time in RILEM B3 model
k_s	cross-sectional shape factor in RILEM B3 model
l	length of specimen (mm)
p	exposed perimeter of specimen (mm)
S	surface area of specimen (mm^2)
$S(t-t_0)$, $S'(t-t_0)$, $S^*(t)$	time dependence terms

S_C	slump of concrete (mm)
t	age of concrete (days) (includes curing time)
t_0	age at first drying = curing time (days)
$t - t_0$	drying time = time of shrinkage test measurement (days)
u	effective section thickness (mm)
V	volume of specimen (mm^3)
w	water content of concrete mix in kg/m^3

Nomenclature

ACI	American Concrete Institute
AIC	Aikaike’s Information Criterion
AR	Autoregressive
ASTM	American Society of Testing Materials
BIC	Bayesian Information Criterion
BS	British Standard
CAIC	Corrected Aikaike’s Information Criterion
CCC	Concordance Correlation Coefficient
CEB	Comité European du Beton
CHQ	Corrected Hannan and Quinn (Information Criterion)
FA	Fly Ash
GGBS	Ground Granulated Blast-furnace Slag
GGCS	Ground Granulated Corex Slag
GGFS	Ground Granulated Ferro-manganese Slag
GLS	Generalised Least Squares
HPC	High-Performance Cement
HQ	Hannan and Quinn (Information Criterion)
MAD	Mean Absolute Deviation
MCDA	Multiple Criteria Decision Analysis
MFPE	Multivariate Final Prediction Error
MSE	Mean Square Error
MSPE	Mean Squared Prediction Error
OLS	Ordinary Least Squares
OPC	Ordinary Portland Cement
PVF	Paste Volume Fraction
RILEM	Reunion Internationale des Laboratoires et Experts des Materiaux, Systemes de Construction et Ouvrages (French: International Union of Laboratories and Experts in Construction Materials, Systems, and Structures)
RMS	Root Mean Square
SABS	South African Bureau of Standards
SANS	South African National Standard
SS	Sum of Squares

SSE	Sum of Squares of Errors
UCT	University of Cape Town
WITS	University of the Witwatersrand, Johannesburg
WLS	Weighted Least Squares

1 Introduction

Concrete is one of the oldest and most widely used construction materials in the world and is used to make foundations, buildings, roads, bridges and many other load-bearing structures. Concrete shrinks over time as it dries under normal atmospheric conditions. Tensile stresses, and ultimately cracks, develop within the concrete if it is prevented from shrinking freely. Thus, **shrinkage** is an important property of concrete as it influences the durability, aesthetics and long-term serviceability of the concrete as well as its load-bearing capacity[1,2]. Thus **the accurate prediction of shrinkage is important in the design stage of any concrete structure[3]**.

Concrete is composed of the following **raw materials:**

- Portland **cement**,
- **supplementary cementitious materials** such as fly ash, blast furnace slag, condensed silica fume and limestone,
- **aggregate:** a *coarse* aggregate (stone) such as granite or quartzite, as well as a *fine* aggregate (sand) such as crushed stone or natural sand,
- **water**, and
- in some cases, chemical **admixtures**.

When these materials are mixed, the water reacts with the cement, forming hydration products that bond the other components together. The mixture eventually hardens to form a stone-like material.

The **composition of the concrete** determines its workability (after mixing), its hardening (curing) behaviour as well as its engineering properties. These engineering properties include strength, elasticity, **shrinkage**, cracking and creep (the inelastic deformation of concrete under load)[1,2]. Shrinkage results mainly from loss of moisture to the environment, but also from chemical reactions within the concrete, as well as reaction with carbon dioxide in the atmosphere[4,5].

The **prediction** of both the long-term shrinkage as well as the rate of shrinkage of concrete, as a function of its composition and environmental factors (such as temperature and humidity), is very important. The shrinkage of concrete leads to a number of structural effects, such as the widening of movement joints in structures, failures at the interface between the concrete and other materials (e.g. ceramic tiles) and a reduction in prestressing force in prestressed concrete. If the shrinkage of a concrete structure is restrained, cracking may occur[(1)]. With a knowledge of the predicted shrinkage behaviour of a particular concrete formulation, structural designers can design concrete structures appropriately to reduce or minimise these effects.

There already exist a number of theoretical and empirical models to predict shrinkage, but most do not take into account the effect of concrete raw materials, such as *different supplementary cementitious materials and aggregate types*. Furthermore, these models were all developed using data derived from non-South African concretes and thus do not take into account effects of cements, supplementary cementitious materials and aggregates *produced and used in South Africa* which may differ substantially from those used elsewhere. Studies on the effects of South African aggregates on concrete shrinkage have been carried out in the past but these have not resulted in a reliable prediction model.

The **aim of this study** was to **consider the possibility that a general approach towards modelling the shrinkage profiles of concrete could be found and to apply it to historical laboratory data on South African concretes.** The different steps in the research were:

- compiling a database of all the available results of laboratory studies on concrete shrinkage carried out in South Africa,
- identifying the most appropriate growth curve model for individual shrinkage profiles and
- modelling the parameters of this growth curve model, fitted to each shrinkage profile individually, in terms of suitable covariates, viz. the composition of the concrete, its other engineering properties as well as shrinkage test conditions.

A final aim of the study was to compare the model developed in this study to several models for shrinkage already in use in the concrete industry, with respect to the data in the South African database.

The **data for this study** were taken from 25 individual studies performed at the University of the Witwatersrand and University of Cape Town laboratories between 1979 and 2009. Each study comprised a number of experiments, with the raw materials, concrete composition and, in some cases, the shrinkage testing conditions being varied over the experiments. Data from a total of 291 experiments was available.

Shrinkage is defined as the decrease in volume of hardened concrete with time, when it is exposed to air of humidity less than 100% and is not subjected to loading or restraint[2,5,9]. Shrinkage is defined as zero at the beginning of the drying time, and is expected to reach an asymptote at large values of time (usually of the order of several years up to 30 years).

Shrinkage data are obtained as follows in a laboratory experiment: Once the concrete mixture has been made, it is poured into moulds. When the concrete has set, the specimens are removed from the moulds and placed into a water bath at 22 ± 1 ºC to cure for 7-49 days. When the specimens are removed from the water bath (at time t_0), pairs of targets to measure shrinkage are affixed to the specimens, as illustrated in Figure 1.1. Six sets of shrinkage measurements (for any given drying time) are made possible by affixing three sets of targets to each of two specimens, or two sets of targets to each of three specimens. The specimens are placed in a laboratory kept at constant temperature and humidity. Shrinkage measurements are taken at various time intervals (time t) by measuring the decrease in the distance between the targets mounted on the specimens. This is done by measuring the strain, normally measured over a gauge length of 100 mm, using a Demec strain gauge measuring to a precision of approximately 15 x10^{-6}. **Shrinkage at a given time *t*** is then calculated as follows:

$$Shrinkage\ at\ time\ t = \frac{change\ in\ length\ of\ sample\ at\ time\ t}{length\ of\ sample\ at\ start\ of\ test} = -\frac{l(t)-l(t_0)}{l(t_0)}$$

Since the length of the sample decreases with time, the negative sign is introduced to represent shrinkage as a positive quantity. Shrinkage is typically expressed as microstrain (x10^{-6} m/m). A mean value for shrinkage at a given drying time is calculated from the six individual measurements, which is then used as the quoted shrinkage value. Shrinkage measurements in the overall data set were terminated between 28 and 992 days after the start of drying. A typical **shrinkage profile** for a concrete sample is shown in Figure 1.2.

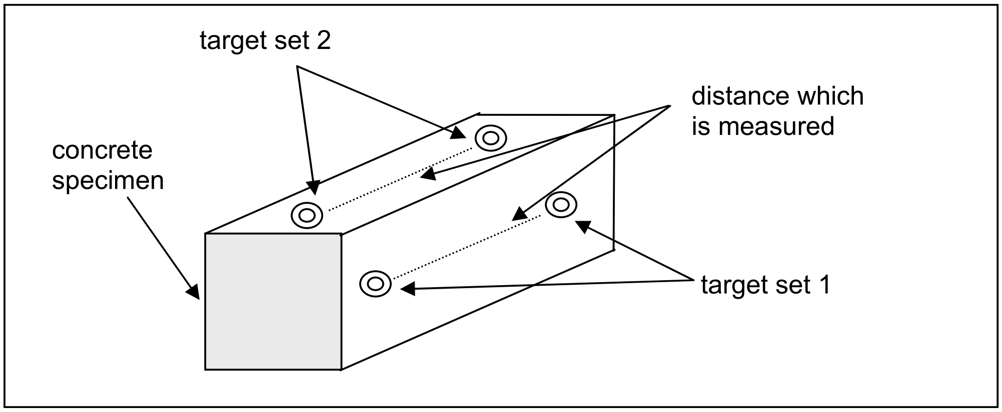


Figure 1.1 Illustration of a concrete specimen with targets for the measurement of shrinkage.

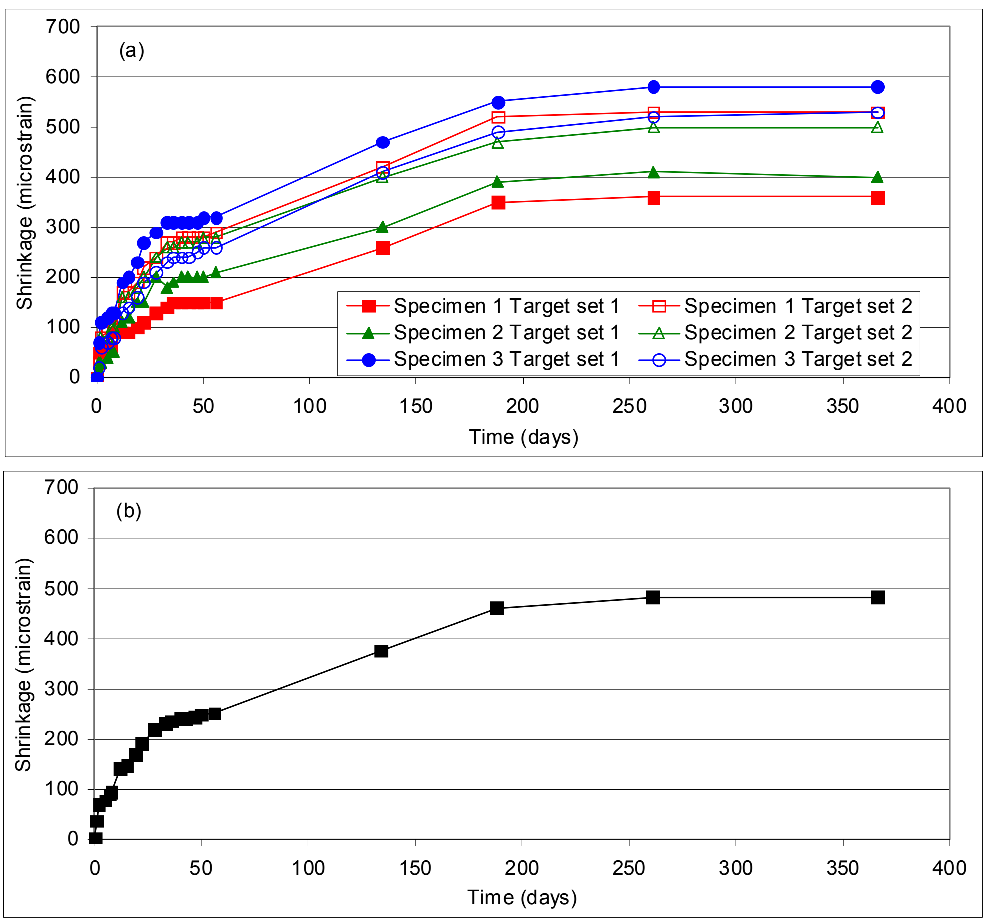


Figure 1.2 A typical shrinkage profile for a concrete sample: (a) six individual measurements and (b) mean shrinkage measurement.

It should be noted that the drying conditions under which the specimens are kept are those specified by the American Society of Testing Materials (ASTM)[(6)] (natural drying conditions) and not those specified by the relevant South African National Standard (SANS) method[(7)].

In the **first stage in the hierarchical modelling process** the most appropriate growth curve for modelling the nonlinear shrinkage-time relationship (shrinkage profile) was chosen. Fifteen candidate growth curve models, thought to be suitable for modelling the shrinkage-time relationship, were identified by means of a literature survey. Ten “key” shrinkage profiles for the evaluation of fifteen candidate growth curve models were selected and used to evaluate these models. Multiple Criteria Decision Analysis (MCDA) was used to identify four models with the best fit for the “key” shrinkage profiles according to a number of goodness-of-fit criteria. The four models were then fitted to all the shrinkage profiles in order to identify the best growth curve model, chosen on the basis of practical considerations as well as the further application of MCDA.

In the **second stage of the hierarchical modelling process**, the parameters of this nonlinear growth curve, fitted to each shrinkage profile, became the dependent variables in a multivariate multiple regression model to form a link between the covariates and the shrinkage profiles. MCDA was once again used to identify the best regression model, using appropriate goodness-of-fit measures.

The **importance of the study** is two-fold:

- The concept of hierarchical nonlinear modelling[(8)], as outlined above, has been applied for the first time to the modelling of concrete shrinkage. This approach could prove useful to other researchers seeking to model concrete shrinkage and related time-dependent properties such as creep.
- The model derived in this study is the first comprehensive model to bring together laboratory data on the shrinkage of concrete generated in South Africa over a span of 30 years, identifying which covariates are the most important contributors to both the magnitude and rate of concrete shrinkage. Recognising the limitations of the study (outlined below), the model provides

a starting point for further, statistically designed, experiments to explore the effect of key variables more fully.

Several **limitations** are inherent in such a study:

- The data are obtained from a number of unrelated studies which were carried out over many years, by different people, in different laboratories, using different raw materials and different equipment. Such data are likely to be inconsistent to some degree.
- The experiments in the database do not meet all the requirements of a well-designed set of experiments: all factors varied over sensible ranges; orthogonal experimental design; randomisation and blocking.
- The database is compiled from research which was not necessarily performed with the specific aims of this study in mind. As a result, the database contains some missing data, i.e. information which was not recorded or was unknown at the time of publication.

However, as this research report will show, the historical data nonetheless provide a good starting point for the development of a model for the shrinkage of South African concretes.

This research report is organised as follows:

- Chapter 2 provides an overview of models for concrete shrinkage already in use in the design of concrete elements, as well as attempts by South African research groups to 'calibrate' some of these models to predict local data more accurately. Variables which are generally known to affect shrinkage are also briefly discussed.
- The nature of the data set used in this study is described in detail in Chapter 3, along with the results of analyses of the internal consistency of the data set. The variance heterogeneity and autocorrelation of the shrinkage data is also analysed.
- Chapter 4 describes, in general terms, the hierarchical nonlinear modelling approach used in this study.
- The application of the hierarchical nonlinear modelling approach to the data set is then dealt with in Chapters 5 and 6: the identification of an appropriate growth curve model for the shrinkage-time relationship is described in Chapter 5 and the modelling of the parameters of this growth curve, for each

shrinkage profile, in terms of the independent variables, by means of multivariate regression, is covered in Chapter 6. For both stages of the hierarchical model, criteria for evaluating and comparing different models and multicriteria methods for combining the different criteria towards selecting the best model(s) are discussed.

- Chapter 7 discusses the comparison of the model developed in this study to some of the models already in use in the concrete industry, with reference to the South African database.
- A summary of the key findings as well as recommendations for further work are contained in Chapter 8.

2 Background to the study and literature review

This chapter begins with a definition of the measurement of shrinkage. Existing models for the prediction of shrinkage are introduced, and their shortcomings in the prediction of the shrinkage of South African data are outlined. Efforts by South African researchers to adapt these models to local data are described. The objective of the present study is then placed into the context of this previous research. Finally, a brief survey of the variables known to affect the shrinkage of concrete is presented.

2.1 Existing models for the prediction of shrinkage

Many models to predict concrete shrinkage have been published over the years(9,10,11,12,13,14,15,16,17,18,19,20,21,22,23). This study is limited to the following **five models which are widely accepted and used by designers or researchers**:

- The ACI 209R-92 model developed by the American Concrete Institute(9)
- The RILEM B3 model developed by Bažant and co-workers(10,11)
- The CEB MC90-99 model developed by the Comité European du Beton(12)
- The GL2000 model developed by Gardner and co-workers(13)
- The SANS 10100-1 model adopted by the South African Bureau of Standards(14).

A recent guide for modelling shrinkage and creep in hardened concrete published by the American Concrete Institute (ACI)(3) focuses on the first four models in the above list.

The **ACI 209R-92 model** was developed in 1971 and revised in 1982 with minor modifications. The model is empirical (based on American concrete data) and does not model chemical or physical shrinkage phenomena as such(9).

The **RILEM B3** model, on the other hand, is based on a mathematical description of more than ten physical phenomena affecting creep and shrinkage, including known fundamental asymptotic properties which, in the opinion of its authors, ought to be satisfied by a shrinkage model(10). The model was published in 2000 and was calibrated against the RILEM database. The RILEM database[24] consists of a collection of 426 shrinkage profiles with associated covariate data, representing concretes made in North America, Europe, New Zealand and Australia.

The **CEB MC90-99 model**, published in 1999 by the Comité European du Beton (CEB) was optimised based primarily on European concretes[12]. The model is known to under-estimate the shrinkage of North American concretes and to over-estimate the shrinkage of some concretes used in Australia and New Zealand[3].

The **GL2000 model** was developed by Canadian researchers Gardner and Lockman in 2004[13]. This model was validated with reference to a subset of the RILEM database.

These models are described in detail in Chapter 7. For the present discussion, it is sufficient to note that the models all have the **general form**

$$\varepsilon_{sh}(t-t_0) = -\varepsilon_{sh\infty} S(t-t_0)$$

where

- $\varepsilon_{sh}(t-t_0)$ is the mean shrinkage strain of the specimen in cross-section at time $(t-t_0)$,
- $\varepsilon_{sh\infty}$ is the **ultimate shrinkage** which is determined by the covariates such as concrete composition, engineering properties (e.g. compressive strength) and drying conditions (e.g. humidity), and
- $S(t-t_0)$ is the **time dependence term** which determines the rate at which the ultimate shrinkage is reached. This term depends on drying time $(t-t_0)$ but may also include some covariates of the types mentioned above.

2.2 Modification of existing shrinkage prediction models for use with South African concretes

From the discussion above, it can be anticipated that the four existing models for concrete shrinkage might not perform well when used to predict the shrinkage of South African concretes, since unique raw materials used to make South African concretes were not taken into account when these models were developed. These unique raw materials encompass

- aggregate (stone and sand) types and morphologies unique to South Africa, and
- supplementary cementitious materials, such as fly ash and blast furnace slag types unique to South Africa.

South African researchers have long recognised this and have attempted to **'calibrate' some of the published models** to obtain a closer fit to South African concrete shrinkage data.

Marques[25] published a study in 1992 based on 151 shrinkage profiles representing studies carried out at the University of the Witwatersrand, Johannesburg (WITS) and the University of Cape Town (UCT) from 1970 to 1992. For concretes containing cement without supplementary cementitious materials, the shrinkage profiles were predicted by means of the BP-KX[18] and CEB-FIP 78[23] models (predecessors of the RILEM B3 and CEB MC90-99 models discussed above, respectively) as well as by the British Standard BS 8110[26] method for the prediction of shrinkage (which is based on simplifications of the ACI 209R-92 model). The following relationship was then derived by regression of the actual shrinkage data against the predicted shrinkage data, using indicator variables for aggregate type:

$$S_{corrected,j}(t-t_0) = a_j\left[S_{predicted}(t-t_0)\right]^{(b+c_j)}(t-t_0)^d$$

where

- $S_{predicted}(t-t_0)$ is the shrinkage at time t-t_0 predicted by a particular model,
- $S_{corrected,j}(t-t_0)$ is the corrected shrinkage at time t-t_0,
- j refers to the aggregate type and

- *a, b, c* and *d* are the coefficients determined from the regression: *b* and *d* depend only on the particular model, while *a* and *c* also depend on the aggregate type.

Correction factors for supplementary cementitious materials were obtained by comparing the shrinkage profile data of concretes containing cement blends to those of exactly matched controls. These were then used as multiplicative 'correction coefficients' for the corrected shrinkage predictions described above.

This study represents a notable contribution to understanding the relationship between aggregate type and shrinkage. The derived coefficients for the different aggregates describe the influence of the aggregates on both the ultimate shrinkage and the rate of shrinkage. However, the method did not allow for the prediction of the performance of any aggregates not represented in the database, since the aggregates were represented by qualitative type and not by a quantitative property (e.g. aggregate stiffness or water demand). However, the derivation of 'correction coefficients' for supplementary cementitious material type necessitated the use of closely matched experiments, which resulted in a good deal of potentially useful data being discarded.

A study carried out in 1993 by **Alexander**(27) on 23 South African aggregates types (each tested by means of two standard concrete mixes), confirmed the importance of the effect of different aggregate types on the shrinkage of concrete. A relative shrinkage value for each aggregate type was calculated based on the average ultimate shrinkage value for the data set. These values were then used as multiplicative 'correction coefficients' for the prediction of shrinkage by the British Standard BS 8110(26) method for the prediction of shrinkage (which is based on simplifications of the ACI209R-92 model). These correction coefficients were subsequently published for use in adjusting the SANS 10100-1 model(14) for the prediction of shrinkage in concrete. The SANS 10100-1 model is closely derived from the BS 8110 model.

While a large variety of aggregates was covered in this study, which enabled a comprehensive comparison of the relative shrinkage of concretes made with different South African aggregates, the derived 'correction coefficient' only addressed long-term shrinkage and not the rate of shrinkage, and also did not make provision for aggregates not represented in the data set.

In 1999, **Ballim**[28,29] assessed the RILEM B3 model, CEB Model Code 1990[22] (a predecessor of the CEB MC90-99 model) and the SANS 10100-1 method[14] for their suitability for the prediction of shrinkage for concretes made in the Gauteng region of South Africa. This was done by comparing the model predictions to the shrinkage data for ten laboratory mixes containing shale and quartzite aggregate. The RILEM B3 model predicted the experimental data most closely, followed by the CEB 1990 model and then the SANS 10100-1 model. None of the models were able to account for the presence of shale in the concrete mix, which led to the development of a 'correction coefficient' for aggregate type. An earlier study had shown that the only significant difference between shale and quartzite aggregate thought to affect shrinkage was the water absorption of the aggregate. Thus the model predictions were adjusted by a multiplicative factor:

$$\textit{Correction factor} = 1 + \left[\frac{0.23}{w_a} \cdot \frac{m_a}{m_t} \right]$$

where

- 0.23 is the water absorption of quartzite,
- w_a is the water absorption of that component of the aggregate which has the highest water absorption,
- m_a is the mass (per unit volume of concrete) of that component of the aggregate which has the highest water absorption, and
- m_t is the total mass of aggregate per unit volume of concrete.

This correction resulted in substantially improved predictive ability of the models, for the data in the study.

This approach had the advantage of relating the effect on shrinkage of different aggregates to a quantitative physical property of the aggregates, which could conceivably be extended to aggregates not represented in the data set. Unfortunately the study was limited to two aggregate types so it was not possible to determine if the relationship between aggregate water absorption and shrinkage held over a wide range of aggregates. Lastly, as was the case in the study performed by Alexander described above, the 'correction coefficient' only affected the ultimate shrinkage and not the rate of shrinkage.

Finally, a study published by **Badenhorst**[30] in 2003 considered eight types of South African aggregates, tested by means of a number of concrete mix proportions, including varying cement types with different types and levels of supplementary cementitious materials. The laboratory shrinkage data were compared to predictions from the RILEM B3 and SANS 10100-1 models. Two 'correction coefficients' were derived:

- Correction for aggregate type: The model predictions were adjusted by a multiplicative factor:

$$\textit{Correction factor} = \frac{\textit{measured shrinkage at 250 days}}{\textit{predicted shrinkage at 250 days}}$$

- Correction for cement type: The study found that cements containing fly ash (FA) or ground granulated blast furnace slag (GGBS) exhibited similar shrinkage to cement which did not contain supplementary cementitious materials. As a result, it was recommended that, for the RILEM B3 model, the factor which accounts for cement types containing supplementary cementitious materials at the levels tested in this study, be set equal to the factor for cement which does not contain these materials. This amounts to including a multiplicative 'correction coefficient' for cement type when using the RILEM B3 model.

The 'correction coefficients' proposed in this study suffer from the same drawbacks as those in the earlier studies by Alexander and Ballim described above. Only the ultimate shrinkage prediction is modified, without consideration for the rate of shrinkage, and no provision was made for aggregates not tested in the study since the aggregates were represented qualitatively (by type) in the 'correction coefficients'.

It is worth mentioning that **researchers in Chile**, where local concretes are also not well-modelled by the existing models, have also tried to determine correction coefficients for their local cement and aggregate types(31). In this work, published in 2004 and based on 24 experiments, modifications were made to both the ultimate shrinkage term as well as the time dependence term for the prediction of shrinkage. The time function from the CEB Model Code 1990(22),

$$S(t-t_0)=\left[\frac{t-t_0}{0.035\cdot\left(2\dfrac{V}{S}\right)^2+t-t_0}\right]^{0.5}$$

where V/S is the ratio of the volume to the surface area of the specimen (in mm), was modified as follows to achieve an improved fit to the experimental data:

$$S'(t-t_0)=\left[\frac{t-t_0}{\mathbf{K_C}\cdot\left(2\dfrac{V}{S}\right)^2+t-t_0}\right]^{\mathbf{b}}$$

The factors K_C and b depend on the cement type. In addition, the ultimate shrinkage term was modified by a multiplicative 'correction coefficient' also representing cement type. These modifications to the model allowed adjustment of the ultimate shrinkage, as well as the rate of shrinkage, predictions for Chilean concretes.

2.3 Present state of the problem

Until now, isolated studies (described earlier), each using a relatively small amount of data, have derived various correction coefficients for existing models to improve the predictive ability of these models for the shrinkage of South African concretes. Over the last 30 years, however, a comparatively large amount of shrinkage data from laboratory tests has been gathered from various studies (not all of which had the modelling of shrinkage in mind but which included the measurement of shrinkage as part of the characterisation of the concrete samples produced). This presented the opportunity to combine all these data sets into one coherent database and to attempt to derive a model for the shrinkage of South African concretes directly from these data, as opposed to determining correction coefficients for existing models. This is the objective of the present study.

2.4 Factors which affect the shrinkage of concrete

It is appropriate to give a brief review of the variables which are known to affect the shrinkage of concrete[1,2,28,29,32,33,34,35]. Firstly, it is important to note that shrinkage occurs exclusively within the paste component of the concrete; the paste comprises the cement, supplementary cementitious materials and water in the concrete mix. The aggregate (stone and sand) component do not shrink (with the exception of so-called shrinking aggregates, such as Karoo sandstones, which are not included in this study).

It is known that **ultimate shrinkage increases with**:

- Decreasing aggregate (sand and stone) content in the cement and thus increasing paste volume fraction (PVF). (PVF is the volume fraction of cement, supplementary cementitious materials and water in the concrete mixture.) The aggregate has a diluting effect on the concrete, thereby reducing the shrinkage capacity of the concrete.
- Decreasing aggregate stiffness. Stiffer aggregates cause increased restraint of paste movement. This trend may only become important in concretes containing aggregates with low stiffness, e.g. low density or weathered aggregates. This effect may also be related to the surface area or porosity of the aggregate[33].
- Increased water demand of the aggregate. (The water demand, or water requirement, of an aggregate is the amount of water required in a standard concrete mix to achieve a given level of workability of the concrete.) The water demand influences the amount of cement required in the concrete mix for a given concrete strength requirement and thus determines the PVF which in turn determines the aggregate content of the concrete (see above). The water demand of an aggregate is not only

influenced by the type of aggregate, but also by its size distribution (average size as well as range of sizes), shape and texture as well as by the presence of impurities in the aggregate(27).

- Decreased bulk density of aggregate (at a given relative density). Decreased aggregate bulk density generally leads to increased water requirement which in turn results in increased shrinkage.
- Increasing water/binder ratio. (Binder refers to cement plus any supplementary cementitious materials.)
- Increasing water content. For a wide range of aggregate/binder and water/binder ratios, the shrinkage is related to original water content.
- Increasing surface area to volume ratio of the test specimen. (Drying occurs only from the exposed surface.) The effect of the shape of the specimen appears to be secondary.
- Moist curing. While moist curing delays the onset of shrinkage, the effect of curing on shrinkage is small. Prolonged curing (greater than one month) seems to somewhat reduce shrinkage. Steam or autoclave curing decreases the magnitude of the shrinkage.
- Decreasing relative humidity.
- Increasing temperature.

The last four factors are extrinsic factors and may be distinguished from the former factors which are all intrinsic to the composition of the concrete.

Shrinkage may also be affected by:

- Cement type: the addition of supplementary cementitious materials such as fly ash and silica fume can affect the shrinkage properties of the concrete. The fineness and alkali content of the cement may also play a role. Cement type is believed to play a less important role than aggregate type(30).
- Cement source: Cements from different sources have different levels of tricalcium aluminate, gypsum and sulphate, all of which appear to have an effect on shrinkage(2).
- The size, size distribution, shape and texture of the aggregate(2,36,37): These factors affect the water requirement of the concrete which is known to affect shrinkage (see above).

- The nature and level of any admixtures used in the concrete: The effect of admixtures on shrinkage is variable and depends on the specific admixture and cement, and the drying environment[2].

Further, **the rate of shrinkage increases with**:

- Increasing water/binder ratio.
- Decreasing specimen size (which increases the rate of water loss).

Previous studies of South African concretes[28,29,33,35] have shown that there were large differences in shrinkage even within a generic group of aggregates and that there was little or no correlation between shrinkage and the water content of the concrete or the stiffness of its aggregates. This suggests that the above-mentioned trends may or may not be applicable to the data used for this study.

This chapter has placed the current study into context by introducing existing models for the prediction of concrete shrinkage, their inherent shortcomings when applied to South African data, and attempts by South African researchers to adapt these models to improve their predictive ability for local data. A brief overview of the factors known to affect concrete shrinkage has been given. In the next chapter, the nature of the data used in the present study will be described in detail.

3 Description of the data set and data preparation

This chapter begins with a description of the sources of the data and a discussion of the issues associated with the analysis of historical data. Next, a detailed description of the covariate and shrinkage data is given, along with the steps taken to prepare the data for analysis. Shortcomings of the data set in common with those of the RILEM database(24) are discussed. Finally, the results of an analysis of comparable experiments within and between individual studies making up the overall data set are presented.

3.1 Sources of data

The data for this study were provided by Ballim (School of Civil and Environmental Engineering, WITS) and Alexander (Department of Civil Engineering, UCT). The data arise out of a number of laboratory studies carried out between 1979 and 2009 in both the WITS and UCT laboratories. No shrinkage measurements from in-service structures (such as bridges or dam walls) which are exposed to diurnally and seasonally varying environmental conditions (such as temperature and humidity) were included in this study.

The data were collated into a database in preparation for analysis. All shrinkage profiles were plotted to check for data entry errors and for internal consistency within the individual studies making up the data set. The data were cross-checked against published journal articles or reports based on the raw data, to confirm accuracy and to obtain any missing data which these sources might contain.

The entire data set is given in the spreadsheet “Data – concrete.xls” (sheets ‘Data_1’, ‘Data_2’ and ‘Data_3’) on the enclosed CD, which also includes plots of all the shrinkage profiles. A subset of the data set is listed in Appendix A. The data set consisted of 291 experiments from 25 individual studies. An experiment is defined as an individual set of shrinkage measurements (as a function of time) together with information on concrete raw materials and composition as well as shrinkage testing conditions. The data from 46 of these experiments could not be used since either the concrete composition (covariate) data or the shrinkage (dependent variable) data were completely missing, effectively reducing the number of experiments from which a model could be derived to 245. Further discussion in this research report is limited to these 245 experiments.

3.2 Issues associated with the analysis of historical data

The database consisted of individual data sets generated over a period of 30 years. Use of such data for statistical modelling has a number of issues associated with it.

Data gathered over such a long period of time are likely to be inconsistent, at least to some degree. The concrete mixes were made by different people, in several different laboratories and using different equipment. The concrete mixes were made using different batches of raw materials whose properties may not be consistent over time. For example, in 1996 (approximately in the middle of the period spanned by the data) a local standard for cement(38) was replaced by a European Standard(39) which resulted in changes in the performance of the cements and may have affected the shrinkage of the concrete in ways not adequately captured by other independent variables. The shrinkage measurements were carried out in two laboratories (with different specimen shapes and under slightly different conditions), by different people and using different equipment. In a planned experimental design, on the other hand, such variation would be controlled by means of blocking, and factors not in the design would be held constant.

However, although the data used for this study were generated by different people, in different laboratories and with different equipment, the concrete preparation and shrinkage testing procedures are relatively simple and were carried out according to well-established standard procedures. The results could thus be expected to be fairly reproducible across laboratories, equipment and researchers preparing specimens. As will be discussed later in this chapter, it was not possible to find directly comparable experiments in the data set to check the inter-laboratory reproducibility of the data, so it is recommended that a study be carried out to test this assumption. Further, it is recommended that a control concrete mix be prepared and tested as part of each future study, so that the long-term reproducibility of these experiments can be assessed.

The individual studies making up the data set were not necessarily carried out with the development of a model for shrinkage as the main aim. This has two consequences. Firstly, **not all important factors affecting shrinkage were varied over sufficiently wide ranges**. In fact, some factors were kept constant because they were known to have a significant effect on shrinkage, e.g. the humidity during the shrinkage test. This is because a standard test for shrinkage was used in most of the experiments, in which all the environmental parameters are kept as constant as possible, in order to compare the shrinkage behaviour of different concretes. However, a model can still be developed with this limitation in mind, and can then be enhanced by further, designed, experiments to include such factors. Secondly, **not all data required for this study was recorded in some of the studies making up the data set**. This missing data was regarded as Missing Completely At Random (MCAR). Except for the properties of the aggregates (see later) no imputation of missing data was carried out.

Collinearity may exist between variables whose levels were not changed independently (as would be the case in an orthogonal experimental design). However, given the different studies making up the data set, there was sufficient variation in the concrete mixes represented so that this was not problematic.

Lurking variables and **serially correlated errors** which, in a designed experiment, would have been accounted for by randomisation and blocking, could be problematic in the analysis of historical data. However, in this case, the data set is made up of a number of unrelated studies (subsets of data) which provides some insurance against possible bias in the data resulting from the lack of blocking and randomisation.

Not all levels of all variables are equally represented in the data set. For example, sixty experiments were carried out using dolerite as the stone type, which means that results for this stone type can be interpreted with greater confidence compared to those for other stone types represented by fewer experiments.

3.3 Covariate data

The data **as provided** contained 42 **time-independent covariates**. There were no time-dependent covariates. An overview of these covariates is given in Table 3.1.

The following steps were taken to prepare the covariate data for analysis:

Inference of cement type from concrete composition / inference of concrete composition from cement type: In cases where supplementary cementitious materials were added separately to the concrete mix (at the concrete mixing stage), the cement type was updated to reflect this, according to South African National Standard SANS 50197-1[40]. Conversely, in cases where the cement type indicated that supplementary cementitious materials were included at the cement factory, the concrete mix variables (e.g. fly ash content) were updated to reflect this. The midpoints of the specification ranges for the relevant supplementary cementitious materials were used since their exact levels were not known. These steps were taken to standardise the use of the cement type and concrete composition variables across the data set.

Table 3.1 Covariates in the data set

Variable	Interpretation	Variable type	Range / Number of categories
Data Set	The data set to which the experiment belongs.	Nominal	25
Test Year	The year in which the experiment was carried out.	Ordinal	1979 - 2009
Cement Type	The type of cement used to make the concrete sample. This also indicates which supplementary cementitious materials are present in the sample and their approximate level.	Nominal	13
Cement Strength Class	The strength class (defined in terms of mortar strength, setting time and expansion (soundness)) to which the cement belongs. This was not defined for cements produced prior to 1996 but for OPC/CEM I lay between a 32,5 and a 42,5.	Ordinal	4
Cement Early Strength Gain Class	The classification of the rate of early strength gain (N = normal, R = rapid) of the cement. This was not defined for cements produced prior to 1996 (which comprises the third category of data).	Nominal	3
Cement Producer	The manufacturer of the cement.	Nominal	4
Cement Milling / Blending Unit	The location at which the cement was milled and / or blended.	Nominal	3
Cement Production Unit	The location at which the cement was produced.	Nominal	12
Stone Type	The type of stone (coarse aggregate) used in the concrete mix.	Nominal	15
Stone Source	The supplier of the stone.	Nominal	34
Stone Size	The nominal maximum size of the stones in mm.	Continuous	13-33 mm

Table 3.1 Covariates in the data set (continued)

Variable	**Interpretation**	**Variable type**	**Range / Number of categories**
Sand Type	The type of sand (fine aggregate) used in the concrete mix.	Nominal	27
Sand Source	The supplier of the sand.	Nominal	34
Admix Type	The type of admixture (if any) used in the concrete mix.	Nominal	9
Cement Content	The number of kg of cement required for 1 m^3 of the concrete mix.	Continuous	113-700 kg/m^3
CSF Content	The number of kg of Condensed Silica Fume required for 1 m^3 of the concrete mix.	Continuous	0-67 kg/m^3
FA Content	The number of kg of Fly Ash required for 1 m^3 of the concrete mix.	Continuous	0-150 kg/m^3
GGBS Content	Kilograms of Ground Granulated Blast furnace Slag required for 1 m^3 of the concrete mix.	Continuous	0-262 kg/m^3
GGCS Content	Kilograms of Ground Granulated Corex Slag required for 1 m^3 of the concrete mix.	Continuous	0-350 kg/m^3
GGFS Content	Kilograms of Ground Granulated Ferro-manganese Slag required for 1 m^3 of the concrete mix.	Continuous	0-244 kg/m^3
Limestone Content	Kilograms of limestone required for 1 m^3 of the concrete mix.	Continuous	0-68 kg/m^3
Water Content	Kilograms of water required for 1 m^3 of the concrete mix.	Continuous	153-225 kg/m^3
Stone Content	Kilograms of stone required for 1 m^3 of the concrete mix.	Continuous	900-1400 kg/m^3
Sand Content	Kilograms of sand required for 1 m^3 of the concrete mix.	Continuous	323-1249 kg/m^3
Admix Content	Kilograms of admixture required for 1 m^3 of the concrete mix.	Continuous	0-3.33 kg/m^3
Paste Volume Fraction	The volume fraction of (cement + supplementary cementitious materials + water) in the concrete mix.	Continuous	0.24-0.44

Table 3.1 Covariates in the data set (continued)

Variable	**Interpretation**	**Variable type**	**Range / Number of categories**
CSF Fraction	The mass fraction of Condensed Silica Fume in (cement + all supplementary cementitious materials) in the concrete mix.	Continuous	0-0.10
FA Fraction	The mass fraction of Fly Ash in (cement + all supplementary cementitious materials) in the concrete mix.	Continuous	0-0.30
GGBS Fraction	The mass fraction of Ground Granulated Blast furnace Slag in (cement + all supplementary cementitious materials) in the concrete mix.	Continuous	0-0.51
GGCS Fraction	The mass fraction of Ground Granulated Corex Slag in (cement + all supplementary cementitious materials) in the concrete mix.	Continuous	0-0.50
GGFS Fraction	The mass fraction of Ground Granulated Ferro-manganese Slag in (cement + all supplementary cementitious materials) in the concrete mix.	Continuous	0-0.50
Limestone Fraction	The mass fraction of Limestone in (cement + all supplementary cementitious materials) in the concrete mix.	Continuous	0-0.13
Water / binder ratio	The mass ratio of water to (cement + all supplementary cementitious materials) in the concrete mix.	Continuous	0.30-0.83
Laboratory	The laboratory (WITS / UCT) in which the shrinkage measurements were conducted.	Nominal	2
Age at first drying	The age of the concrete in days when it is first taken out of the curing bath and starts to dry (i.e. starts to shrink). This is t=0 in the shrinkage experiment.	Continuous	7-49 days
Specimen shape	The shape of the concrete specimens used for the shrinkage test (e.g. cylinder, prism).	Nominal	2 (cylinder, square prism)

Table 3.1 Covariates in the data set (continued)

Variable	**Interpretation**	**Variable type**	**Range / Number of categories**
Specimen dimensions	The dimensions (in mm) of the concrete specimens used for the shrinkage test.	Continuous	Cylinder: diameter 100-300 mm, height 150-305 mm Prism: (51-102) x (51-102) x (200-300) mm
Temperature	Average temperature in ºC for storage of the specimens during the shrinkage test.	Continuous	21-25 ºC
Humidity	Humidity in % for storage of the specimen during the shrinkage test.	Continuous	43-72%
Slump	The slump in mm of the wet concrete mix before pouring into moulds. The higher the slump, the greater the 'workability' of the wet concrete.	Continuous	10-115 mm
28-day cube compressive strength	The compressive strength in MPa of the hardened concrete measured 28 days after pouring into cube moulds. Compressive strength is the amount of compressive force the concrete can withstand before it fails (crushes).	Continuous	20-79 MPa
28-day elastic modulus	The elastic modulus in GPa of the hardened concrete measured 28 days after pouring into moulds. Elastic modulus (stiffness) is the amount of bending force the concrete can withstand before it fails (crushes).	Continuous	15-52 GPa

Properties of the aggregates: In a model, it would be desirable to describe the aggregates (stone and sand) by means of one or more continuous properties, as opposed to simply indicating their type by using indicator variables. Not only would this reduce the number of covariates required to describe the aggregates during model development, but it might also allow the prediction of the shrinkage profile of concretes containing aggregates not represented in the data set from which the model was developed, if the continuous property (or properties) of the new aggregates was known. Hence, data on the physical properties of the aggregates as listed in Table 3.2 were collated, generating a further fifteen time-independent covariates. In cases where information on these characteristics for the specific batch of aggregate used in an experiment was not available (which was quite often the case), the published values for that aggregate type were used(36,41). Where such values were given as ranges, the midpoints of the ranges were used.

Definition of new covariates: Three further covariates were added to the data set since it was thought that these could be important explanatory variables for shrinkage. The ratio of aggregate to binder in the concrete mix was defined analogously to a criterion used in the RILEM B3 model(10,11). The ratio of stone to sand in the concrete mix was defined in order to complete the description of the concrete mix using only ratios of the raw materials. Lastly, the ratio of twice the volume to the exposed surface area of the test specimen was calculated for each experiment. This definition is in line with that used in other models(9,10,11,12,13) and standardises the different shapes and dimensions of the specimens in the data set. These variables are listed in Table 3.3.

Calculation of temperature and humidity: Typically, shrinkage specimens are kept in a laboratory at constant temperature and humidity. However, in some of the earlier studies, adequate control of these variables for the duration of the shrinkage test was not always possible. Fortunately, in these cases (56 experiments), temperature and humidity readings were taken at the same time as shrinkage measurements. In order to calculate (as closely as possible from the available data) the average temperature and humidity to which the specimens were exposed over the duration of the shrinkage test,

Table 3.2 Additional covariates representing the physical properties of the aggregates

Variable	**Interpretation**	**Variable type**	**Range**
	Stone		
Relative Density	The density of the stone divided by the density of water (1000 kg/m^3).	Continuous	2.6-3.1
Loose Bulk Density	The mass of stone particles occupying a certain volume. This depends on the packing of the particles which in turn depends on particle size, shape and surface texture.	Continuous	1285-1580 kg/m^3
Consolidated Bulk Density	The definition is similar to that of the loose bulk density, except that the sample is 'consolidated' by vibrating or rodding to give a measure of the practical upper bound of the bulk density.	Continuous	1430-1760 kg/m^3
Water absorption	The % water absorption capacity of the stone.	Continuous	0.21-1.64 %
Coefficient of thermal expansion	The relative increase in size of the stone per 1ºC increase in temperature.	Continuous	6.0-12.9 x10^{-6}
Unconfined compressive strength	The strength of the stone (in MPa) when crushed uniaxially without lateral restraint.	Continuous	113-527 MPa
10% FACT	The 10% Fines Aggregate Crushing Value: The load required to crush an aggregate sample to give 10% material passing a specified sieve after crushing. This represents the crushing strength of the stone.	Continuous	110-460 kN
Elastic Modulus	The elastic modulus in GPa of the stone. Elastic modulus is the slope of the stress-strain function in the linear, elastic range.	Continuous	11-114 GPa
Flakiness Index	The proportion of flaky particles in the stone.	Continuous	11-26

Table 3.2 Additional covariates representing the physical properties of the aggregates (continued)

Variable	Interpretation	Variable type	Range
Sand			
Relative density	As defined above for stone.	Continuous	2.4-2.9
Loose Bulk Density	As defined above for stone.	Continuous	1370-2020 kg/m^3
Consolidated Bulk Density	As defined above for stone.	Continuous	1530-2130 kg/m^3
Fineness Modulus	A measure of the broadness of the particle size distribution of the sand.	Continuous	1.8-3.6
Fines	The percentage of sand particles smaller than 75 µm.	Continuous	0.1-16%
Elastic Modulus	As defined above for stone.	Continuous	67-114 GPa

Table 3.3 Further covariates defined for the data set

Variable	Interpretation	Variable type	Range
Aggregate / binder ratio	The mass ratio of (stone + sand) to (cement + all supplementary cementitious materials) in the concrete mix.	Continuous	2.1-8.7
Stone / sand ratio	The mass ratio of stone to sand in the concrete mix.	Continuous	0.8-3.6
2 * Volume / surface area of specimen in contact with atmosphere	Twice the ratio of the volume to the surface area of the concrete specimen in contact with the atmosphere, in mm. Note: Specimens tested at WITS have their ends sealed, unless otherwise noted, whereas specimens tested at UCT do not have their ends sealed unless otherwise noted. The 'ends' refer to the smallest faces of square prisms (shaded area in Figure 1.1) and the circular faces of a cylindrical prism. While WITS generally uses square prisms and UCT generally uses cylindrical prisms, this is not always the case in the data set and there is no trend regarding specimen shape with time. Internationally, both specimen geometries are used[24].	Continuous	33-150 mm

time-weighted average temperature and humidity values were calculated for these experiments. For each drying time at which shrinkage, temperature and humidity measurements were recorded, it was assumed that the specimen had been exposed to the recorded temperature and humidity conditions since the day after the previous set of measurements, as illustrated in Figure 3.1 for one experiment. Thus, the average temperature and humidity for the entire duration of the shrinkage test could be calculated.

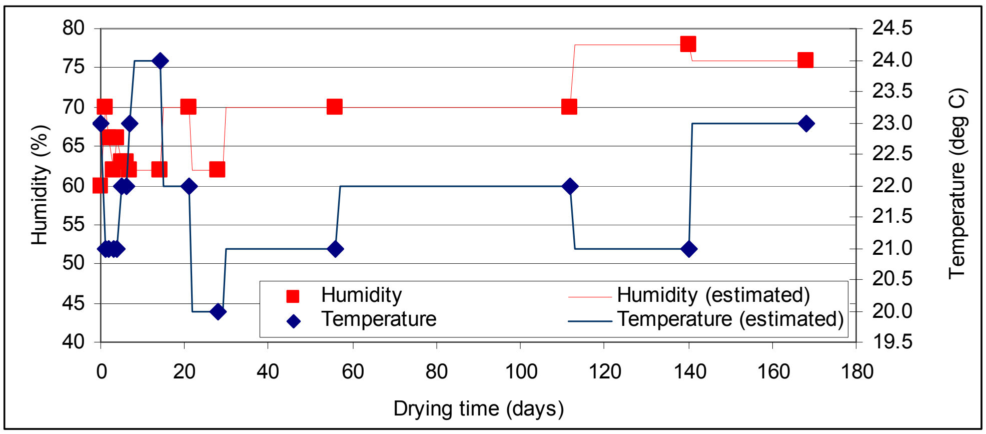


Figure 3.1 Temperature and humidity readings and estimated values for an experiment during which the environmental conditions could not be kept constant.

3.4 Shrinkage data

Given the laboratory procedure described in Chapter 2 for the preparation and testing of laboratory specimens, the data represent drying shrinkage only (not carbonation or autogenous shrinkage) determined under laboratory drying conditions (a temperature of 22 ± 2ºC and a relative humidity of 65 ± 5%)[42].

The data for one experiment comprise repeated measures on the same sample. The times at which the measurements are taken are not separated by constant intervals, nor are the times necessarily the same from experiment to experiment. The shape of the shrinkage profile resembles a growth curve.

The shrinkage data from each experiment consisted of one to 55 data points, i.e. shrinkage measurements taken at different times. (There were 18 experiments for which only one shrinkage data point was recorded. These experiments could obviously not be used for fitting growth curves – see Section 6.2 – but were included in the data set for completeness and were used in the evaluation of the different models, described in Chapter 7.) The ranges covered by the shrinkage measurements were 0 microstrain (at a drying time of zero days) to a final value of 158-829 microstrain at 28-992 days. Some shrinkage experiments were conducted for much longer periods (~ 900 days) than others (~ 28 days). Since shrinkage is a long-term characteristic of concrete, it raised the question as to whether all experiments should have equal weight in model development and whether some form of weighted regression might be needed. This issue will be addressed in Chapter 6.

Some shrinkage experiments ended before the shrinkage started to approach an asymptote, as illustrated in Figure 3.2. This raised the question whether such experiments should have equal weight in model development. This issue, too, will be dealt with in Chapter 6.

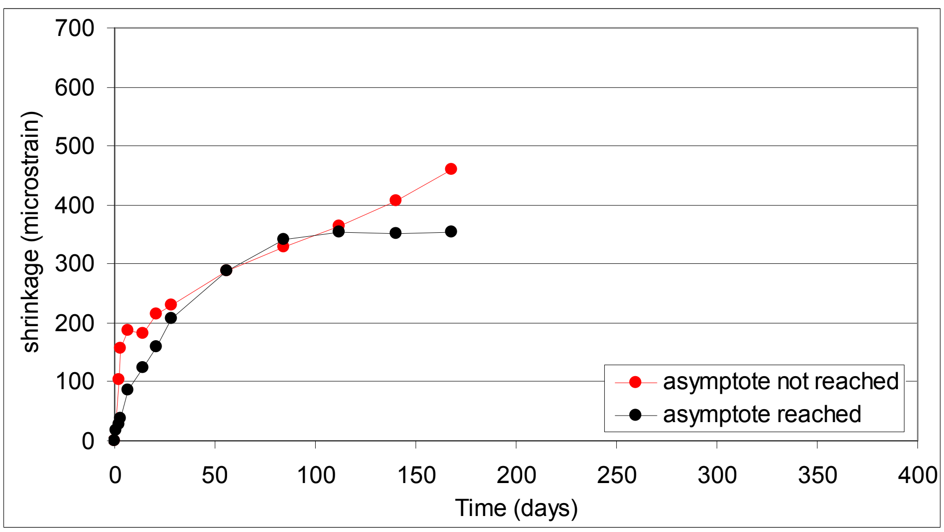


Figure 3.2 Shrinkage profiles for a concrete sample which had approached an asymptote and a sample which had not approached an asymptote by the time shrinkage testing ended

It was found that the **variance of the shrinkage measurements increased as the shrinkage increased**, as illustrated in Figure 3.3 for a single experiment. The variance of the shrinkage measurements could be determined since, for a single experiment, six shrinkage measurements are averaged to obtain a mean shrinkage value at each drying time. This is discussed in more detail in Chapter 5.

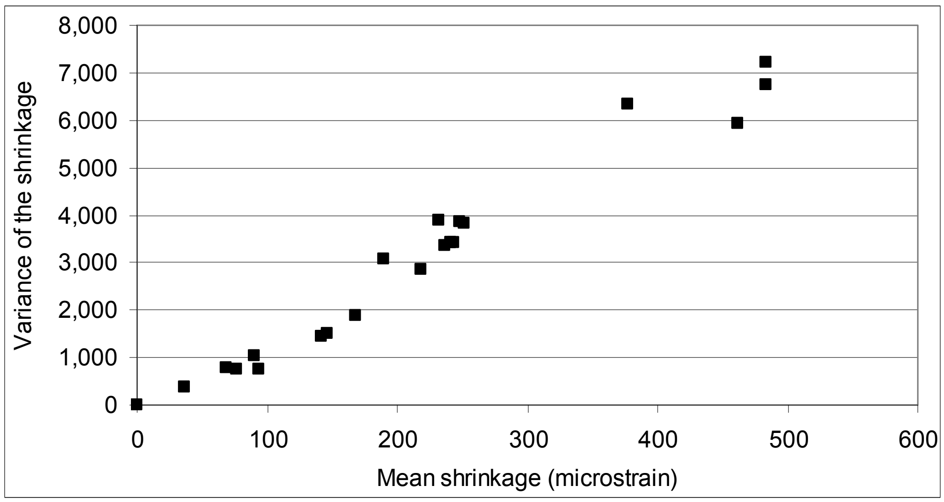


Figure 3.3 An illustration of the relationship between the variance of the shrinkage and shrinkage for a single experiment

Further, it was found that there was **serial correlation among measurements within an experiment**, as illustrated in Figure 3.4. The correlogram, based on the method of Weiss[43], shows the correlation coefficients between shrinkage values at different times (effectively a graphical representation of the correlation matrix). The shrinkage values at different drying times are expected to be correlated and the correlation coefficients and correlogram confirm this. The correlation between shrinkage values decreases with increasing lag, as might be expected. The correlations between later values of shrinkage are higher than between earlier values of shrinkage, which reflects the reduced rate of change in the shrinkage-time curve (refer back to Figure 1.2).

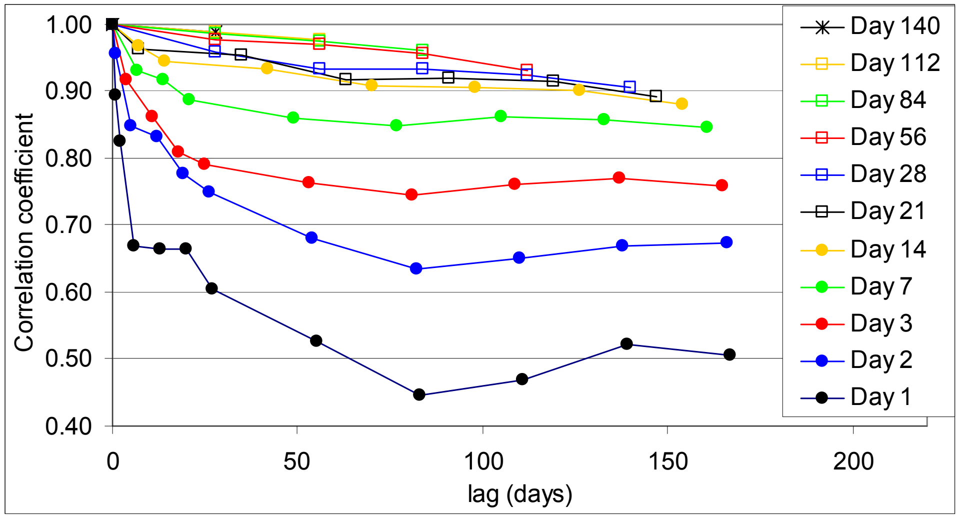


Figure 3.4 Correlogram for the data set showing autocorrelation of the shrinkage data

3.5 Shortcomings of the data set

The data set shares two shortcomings with the RILEM database[(2424)] as outlined by the American Concrete Institute[(3)]. Firstly, the experiments were performed with relatively small specimens compared with more regularly sized structural elements. It is questionable whether the curing environment and resulting mechanical properties of concrete in the interior of large elements are adequately represented by experiments carried out on small specimens. Secondly, many of the shrinkage experiments were carried out over relatively short durations, which reduces the usefulness of the data in predicting long-term effects. This is illustrated in Figure 3.5 which shows the distribution of shrinkage measurement times as a function of drying time and in Figure 3.6. which shows the distribution of final drying times for each experiment.

However, in spite of these limitations, it is agreed that databases such as the one used in this study and the RILEM database provide an important source of data and a basis for comparing prediction models[(3)].

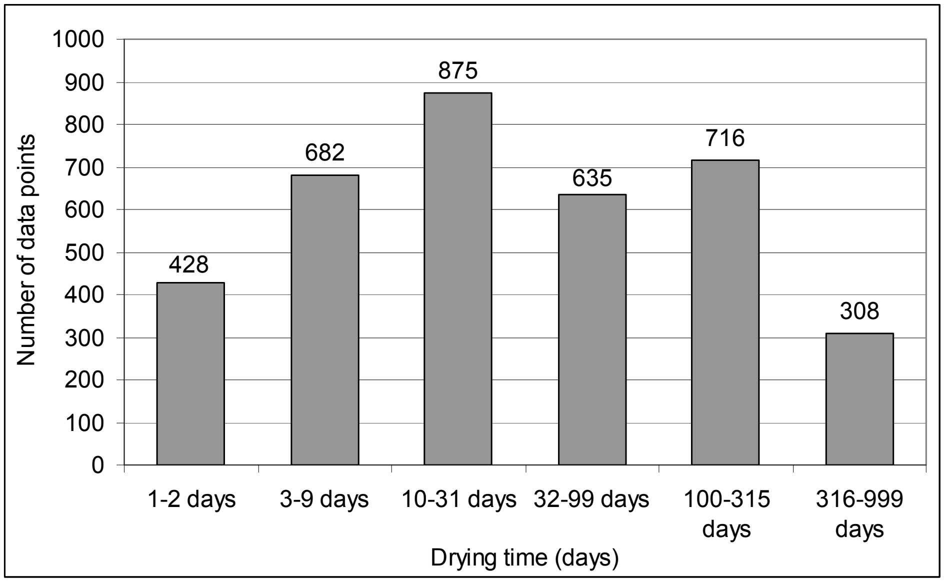


Figure 3.5 Distribution of shrinkage measurement times as a function of drying time (in half-log decades)

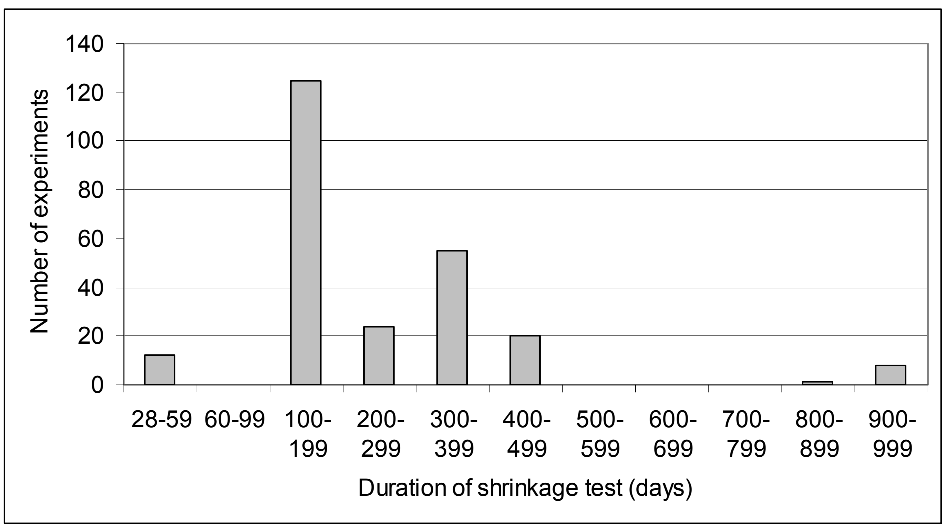


Figure 3.6 Distribution of shrinkage test durations

3.6 Comparison of experiments within and between individual data sets, between test years and between laboratories

Comparable experiments (shrinkage profiles) were identified by matching cement, stone and sand type, followed by identifying experiments with similar water/binder ratio, aggregate/binder ratio, cement content, water content, sand content, stone content, stone/sand ratio, paste volume fraction and supplementary cementitious material fraction, in that order. The comparable experiments thus identified are listed in Appendix B. Full graphical comparisons may be found in the spreadsheet file "Comparable experiments.xls" on the enclosed CD.

Certain individual data sets contained comparable experiments which allowed an estimation of the reproducibility of the experiments. These arose from 'control' experiments repeated for different parts of a study carried out at different times, or from experiments carried out with the same cement types but sourced from different suppliers, or simply from very similar experiments within a study. Fifteen sets of comparable experiments were identified, each consisting of between two and four experiments. Inspection of the shrinkage profiles revealed that in eight instances, the shrinkage profiles matched quite well, whereas the profiles in the other seven instances were quite variable, as illustrated in Figure 3.7 for two extreme instances. For the six comparisons containing more than two experiments, the relative standard deviation of the final shrinkage values averaged 16% (range 5-33%), which indicated that the **variability of comparable experiments even within the same study can be quite high in some cases**.

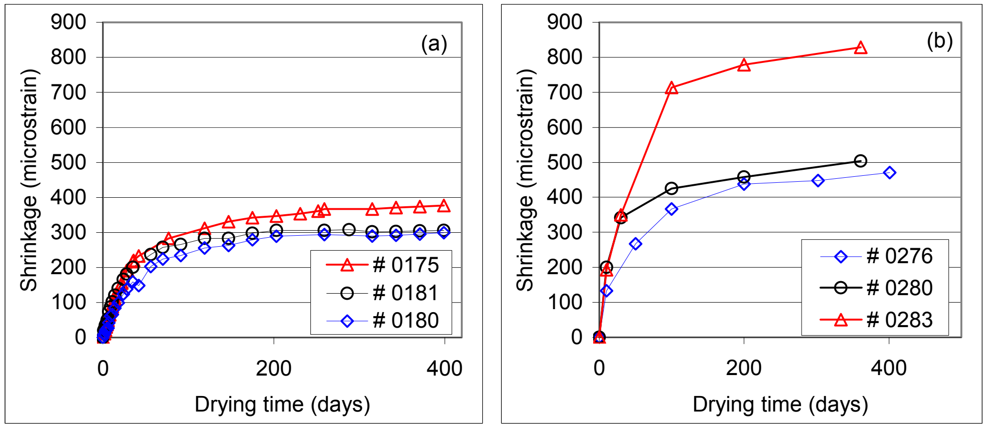


Figure 3.7 Shrinkage profiles of comparable experiments within individual data sets with relative standard deviations of the final shrinkage value of (a) 5% and (b) 33%

Next, twenty sets of comparable experiments between different individual data sets (and thus also between different years in which the experiments were carried out) were identified. The experiments within each set of comparable experiments were performed in the same laboratory. Each set consisted of between two and five experiments. Inspection of the shrinkage profiles, as illustrated in Figure 3.8 for two extreme instances, revealed that in eight instances the shrinkage profiles matched quite well and in seven instances the profiles were quite variable. The other five instances could not be assessed as the experiments were of very short duration. For the six comparisons containing more than two experiments, the relative standard deviation of the final shrinkage values averaged 15% (range 4-25%), which is similar to what was found for the intra-study comparisons. In the three instances where comparison by ANOVA was feasible, the differences were not significant at the 5% level, although it must be noted that these comparisons were limited to data sets generated in 2000 and 2001. Those comparisons spanning a number of years unfortunately contained only two experiments each, but the shrinkage profiles were very similar in each case.

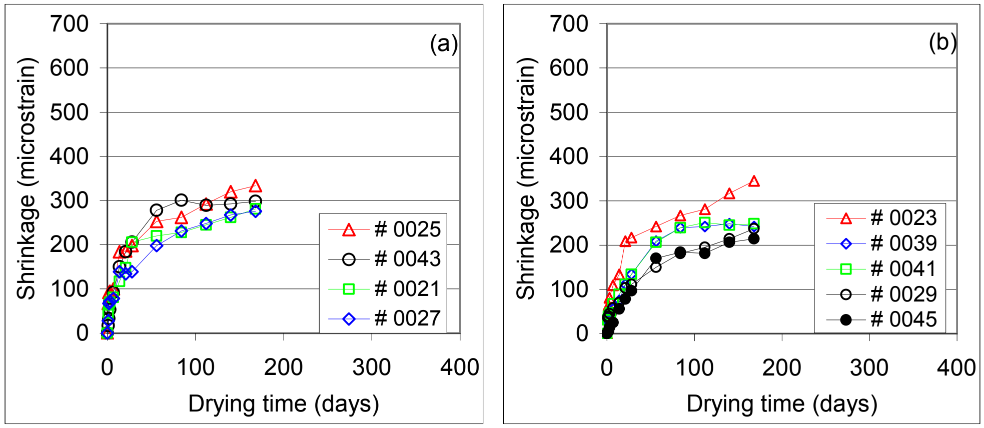


Figure 3.8 Shrinkage profiles of comparable experiments between individual data sets and test years with relative standard deviations of the final shrinkage value of (a) 4% and (b) 20%

In conclusion, the fact that there were good matches between experiments from different individual data sets and from different test years, indicated that **the overall data set was sufficiently consistent for analysis as a whole and that provision need not be made in the covariates for the test year or the individual data set from which a particular experiment originated**.

It was not possible to identify any sets of comparable experiments which allowed the comparison of results across the two laboratories which contributed shrinkage data to the study. Due to their different geographical locations, the research groups in these laboratories used different sources of stone and sand. As a result, in this study, **the laboratory covariate is confounded with the other covariates specifying the concrete raw materials**. As mentioned previously, it is recommended that a laboratory round-robin study be conducted to compare the shrinkage test method as carried out at the WITS and UCT laboratories.

This chapter has described the nature of the data and how it was prepared for analysis. In the next chapter, the overall approach to developing a model for shrinkage from this data set will be outlined.

4 Overall approach to developing a model for shrinkage from the data set

In this chapter, the requirements for a reliable shrinkage model are discussed relative to what is possible to achieve with the data set used in this study. The nonlinear hierarchical modelling approach is presented in detail and it is shown how this technique suits the nature of the data and the modelling problem. The main steps in the development of the model, to be covered in detail in subsequent chapters, are then highlighted.

4.1 Requirements for a reliable shrinkage model

Bažant notes that an acceptable model for shrinkage must(4)

(1) fit the data – individually and as a whole,
(2) cover a broad range of drying time, age at first drying (curing time), relative humidity, specimen thickness, etc.,
(3) conform to the mathematical consequences of the physical phenomena associated with shrinkage, such as moisture diffusion and chemical processes causing autogenous volume change,(44,45) and
(4) allow the extrapolation of short-time tests to longer times, higher ages at first drying, greater thicknesses, etc.

Furthermore, the American Concrete Institute Committee 209, in its guide for modelling and calculating shrinkage and creep in hardened concrete(3), remarks that a shrinkage model should (*numbering continued from above for ease of reference)*

(5) at least include the following information: a description of the concrete either as mixture proportions or by mechanical properties (such as compressive strength or elastic modulus), relative humidity, age at first drying, drying time and specimen size,
(6) allow for the use of measured values of concrete strength and elastic modulus (as opposed to merely using design values for these characteristics),

(7) allow the extrapolation of measured shrinkage data to obtain long-term predictions and
(8) contain mathematical expressions that are not unreasonably sensitive to small changes in input parameters and are 'easy to use'.

The model developed in this research was evaluated for its fit to the data (see Chapters 5 and 6). The model did <u>not</u> cover a broad range of the <u>extrinsic</u> variables listed in (2) and (5) above (with the exception of drying time) since the data came from laboratory experiments carried out under controlled conditions, as was discussed in Chapter 3. The model developed in this study, due to the nature of the data, focuses more on the <u>intrinsic</u> concrete properties such as its composition and mechanical properties (including measured values of these properties). However, model components, such as those for relative humidity, from models which <u>do</u> cover extrinsic variables more adequately, such as the RILEM B3 model(10,11), could potentially be adapted to supplement the model developed in this study.

The model was empirically derived and does not necessarily conform to all the mathematical consequences of the physical phenomena associated with shrinkage. As will be shown in Chapter 6, the model is not unduly sensitive to small changes in input parameters. Finally, the identification of an appropriate growth curve model (as described in Chapters 4 and 5) allows the extrapolation of short-term shrinkage data to obtain longer-term predictions, although this was not specifically addressed in this study.

The American Concrete Institute Committee 209 goes on to note(3) that the stiffness of the aggregate significantly affects the shrinkage of concrete. Some models assume that the effects of the aggregate relate to aggregate density or to the elastic modulus of the concrete. Models that do not incorporate the mechanical properties of the concrete (relying on mixture proportions alone) may not account for variations in behaviour due to aggregate properties(3). In this study, the aggregate types (or their continuous characteristics) will be used as covariates.

4.2 Approach to modelling the data

The following characteristics of the data and requirements of the model have already been established:

- The model must include the time component (i.e. the growth curve): It is not sufficient to model 'summary' values of the growth curve such as the rate of shrinkage, the final (ultimate) shrinkage, etc.
- The data consist of repeated response (shrinkage) measurements made on a number of different experiments.
- The shrinkage y depends nonlinearly on a set of unknown parameters $\boldsymbol{\beta}$ for each experiment. At the very least, the relationship between shrinkage and time is nonlinear.
- The shrinkage profiles are similarly shaped across experiments, but may have different values of the parameter vector $\boldsymbol{\beta}$ for different experiments.
- The variance depends on the level of shrinkage.
- There is serial correlation among measurements within an experiment.
- The inter-experiment variation between the values of the parameter vector $\boldsymbol{\beta}$ may be considered to be a combination of experiment-specific characteristics (i.e. related to the covariates) and random variability.
- As far as possible, the model's form and parameters should have a practical interpretation.

The approach used in modelling the data was that of **hierarchical nonlinear modelling**[8]. This approach is suitable in situations where the data consists of repeated observations of a continuous response (shrinkage) on each of several subjects (the individual experiments) over a continuous variable (time); where there is variability in the relationship between the response and time across subjects; and where parameters (covariates) exist which vary across experiments and characterise the behaviour of individual subjects, dictating the variation in the patterns of time-response. Hierarchical nonlinear modeling may be used to model the typical behaviour of the phenomena represented by the covariates: the extent to which the covariates vary across experiments and whether some of the variation is systematically associated with individual attributes. Using the model, individual-level prediction may also be done[46]. This approach has been used in a variety of applications[46,47].

4.3 Theoretical basis for hierarchical nonlinear modelling

The outline given by Davidian and Giltinan[8] is followed. Let y_{ij} denote the j^{th} response (shrinkage value), j = 1, …, n_i, for the i^{th} experiment, i = 1, …, m, at a set of conditions summarised by the vector $\mathbf{t_{ij}}$ which in this case contains the values of time at which shrinkage measurements were made. A nonlinear function $f(\mathbf{t},\boldsymbol{\beta})$ models the relationship between the response and $\mathbf{t}$, where $\boldsymbol{\beta}$ is a (p x 1) vector of unknown parameters, in this case the parameters of the nonlinear function used to model the shrinkage-time relationship. The function f has the same basic functional form for all the experiments and depends nonlinearly on the regression parameters $\boldsymbol{\beta}_\mathbf{i}$ specific to the i^{th} experiment as well as on the experiment-specific vectors $\mathbf{t_{ij}}$. Thus

$$E(y_{ij}) = f(\mathbf{t_{ij}},\boldsymbol{\beta}_\mathbf{i}).$$

The **first stage of the model** describes the systematic and random features of the data at the ***intra-experiment*** level. For the i^{th} experiment, the j^{th} response can be written as

$$y_{ij} = f(\mathbf{t_{ij}},\boldsymbol{\beta}_\mathbf{i}) + e_{ij}$$

where e_{ij} is a random error term reflecting the residual variability in the response, over and above that explained by the model, in the i^{th} experiment, with $E(e_{ij}) = 0$.

The data for the i^{th} experiment can thus be summarised as follows:

$$\begin{aligned}\mathbf{y_i} \quad &= \begin{bmatrix} f(\mathbf{t_{i1}},\boldsymbol{\beta}_\mathbf{i}) \\ \dots \\ f(\mathbf{t_{in_i}},\boldsymbol{\beta}_\mathbf{i}) \end{bmatrix} + \mathbf{e_i} \\ &= \mathbf{f_i}(\boldsymbol{\beta}_\mathbf{i}) + \mathbf{e_i}\end{aligned}$$

where $E(\mathbf{e_i}) = 0$.

We assume a comparable pattern of intra-experiment variation across experiments (which is realistic given what is known about the data), and thus define the common intra-experiment variance structure, which allows for variance heterogeneity and correlation within experiments, as

$$\begin{aligned} Cov(\mathbf{e_i}) &= \sigma^2 \mathbf{G_i^{1/2}}(\boldsymbol{\beta}_i, \boldsymbol{\theta}) \boldsymbol{\Gamma}_i(\boldsymbol{\alpha}) \mathbf{G_i^{1/2}}(\boldsymbol{\beta}_i, \boldsymbol{\theta}) \\ &= \mathbf{R_i}(\boldsymbol{\beta}_i, \boldsymbol{\xi}), \quad \boldsymbol{\xi} = [\sigma, \boldsymbol{\theta}', \boldsymbol{\alpha}']' \end{aligned}$$

where

- the scale parameter σ describes the overall level of precision of the response.
- the (n_i x n_i) diagonal matrix $\mathbf{G_i^{1/2}}(\boldsymbol{\beta}_i, \boldsymbol{\theta})$ characterises intra-individual variance.
- the variance parameters **θ** specify the functional forms of the variance functions.
- the (n_i x n_i) matrix $\boldsymbol{\Gamma}_i(\boldsymbol{\alpha})$ describes the correlation pattern within the i^{th} experiment.
- **α** is a vector of correlation parameters common to all experiments.
- the matrix $\mathbf{R_i}$ is a covariance matrix summarising the pattern of random variability associated with the data for experiment *i.*

To complete the description of intra-experiment variation, it is assumed that the intra-experiment response is normally distributed, which follows from the error specification

$$\mathbf{e_i} \mid \boldsymbol{\beta}_i \sim N(\mathbf{0}, \mathbf{R_i}(\boldsymbol{\beta}_i, \boldsymbol{\xi})).$$

Along with an assumption about the distribution of y_i, the first stage of the model describes the systematic and random features of the data at the intra-experiment level.

In the **second stage of the model, inter-experiment variation** is characterised by a model for variation in each of the *p* regression parameters in $\boldsymbol{\beta}$. This is done by deriving a multiple regression model for each regression parameter:

$$\boldsymbol{\beta}_\mathbf{l} = \mathbf{A}_\mathbf{l}\mathbf{b}_\mathbf{l} + \boldsymbol{\delta}_\mathbf{l}$$

where

- *l = 1…p*
- $\boldsymbol{\beta}_\mathbf{l}$ is a (*m* x 1) vector containing the values of the $\hat{\beta}_l$ estimated for each of the *m* experiments.
- $\mathbf{A}_\mathbf{l}$ is an (*m* x (*k*+1)) matrix with containing the values of the *k* independent variables for the *m* experiments, plus a column (to refer to the intercept) containing *m* values equal to 1.
- $\mathbf{b}_\mathbf{l}$ is a ((*k*+1) x 1) vector containing all the model parameters to be estimated on the basis of the data: the intercept and *k* slope coefficients relative to each independent variable.
- $\boldsymbol{\delta}_\mathbf{l}$ is an (*m* x 1) vector containing the error terms.

For a given set of experimental conditions (independent variables), each nonlinear regression parameter may then be estimated on the basis of its multiple regression model

$$\hat{\beta}_l = \hat{b}_{l0} + \hat{b}_{l1}a_{l1} + \hat{b}_{l2}a_{l2} + \ldots + \hat{b}_{lk}a_{lk} \quad l = 1,.., p\ .$$

Once each nonlinear regression parameter has been estimated, these estimates may be substituted into the nonlinear regression model to obtain a prediction of the shrinkage profile for the given set of experimental conditions.

4.4 Main steps in model development

The main steps in developing the model were thus the following:

- Modelling the individual shrinkage profiles by means of growth curves and nonlinear regression,
- choosing the best growth curve model for the data,

- exploratory data analysis and weighting of data in preparation for modelling of the effects of the covariates,
- modelling the parameters of the chosen growth curve in terms of the independent variables and
- assessing the goodness-of-fit of the models to the data and choosing the best model for the data.

These steps, and their results, will be described in detail in Chapters 5 and 6.

This chapter has shown to what extent the nature of the data used in this study matches the criteria for a reliable shrinkage model. The nonlinear hierarchical modelling approach was outlined and its applicability to the data set and the research question were highlighted. Finally, the main steps in model development were summarised. Chapter 5 describes the first step in the development of the model, viz. the selection of an appropriate growth curve to model the shrinkage-time relationship of individual shrinkage profiles.

5 Modelling the individual shrinkage profiles by means of nonlinear regression

The first step in the hierarchical modelling process is to choose the most appropriate growth curve for modelling the nonlinear shrinkage-time relationship (shrinkage profile). In the second step, the parameters of this nonlinear growth curve, fitted to each shrinkage profile, become the dependent variables in a multiple regression model to form a link between the covariates and the shrinkage profiles.

In this chapter, candidate growth curve models, thought to be suitable for modelling the shrinkage-time relationship, are firstly reviewed. Then, the heterogeneous variance of the response variable (shrinkage) is modelled in order to define a weighting scheme for the shrinkage-time data. This weighting scheme is subsequently evaluated to determine the best nonlinear least squares fitting method appropriate to the data. The autocorrelation of the shrinkage-time data is also considered.

Next, "key" shrinkage profiles for the evaluation of the candidate growth curve models are selected and used to evaluate these models. MCDA is used to identify a group of models with the best fit for the "key" shrinkage profiles. This group of models is then fitted to all the shrinkage profiles in order to identify the best growth curve model, chosen on the basis of practical considerations as well as the application of MCDA.

5.1 Selection of candidate growth curves for modelling the shrinkage-time relationship

The shape of the shrinkage profiles resembles a growth curve (as illustrated in Figure 2.1). Growth curves arise when repeated measurements are observed for a number of individuals (in this case, experiments) with an ordered dimension for occasions (in this case, time)[48]. Typically the analysis of growth curves emphasises the exploration of the differences in the within-experiment development of the growth curve between experiments[48].

Of the many growth curve models available[49], growth curves considered for fitting to individual shrinkage profiles were selected according to the following criteria: for a growth curve model, y = f(t),

- the shape of the growth curve should resemble that of the shrinkage profile shown in Figure 2.1,
- $f(t) = 0 \;\; at \;\; t = 0$,
- $\lim_{t \to \infty} f(t)$ = ultimate shrinkage = constant (the constant is a parameter to be modelled in the second stage of the hierarchical modelling process),
- the parameters should have a practical meaning if at all possible, i.e. they should relate to the ultimate shrinkage, rate of shrinkage, etc. and
- the number of parameters should encourage model parsimony[49].

In total, fifteen suitable candidate growth curves were identified, according to the above criteria, from various literature sources. The models and their sources are listed in Table 5.1. Where applicable, the use of a particular growth curve in a published model for concrete shrinkage is identified.

It should be noted that Model 9 reduces to Model 5 when γ=1. Likewise, Model 10 reduces to Model 2, and Models 12, 13 and 15 reduce to Model 4 when γ=1.

Table 5.1 Candidate growth curve models for the shrinkage profiles

Model #	f(t) =	$\lim_{t\to\infty} f(t) =$	Interpretation of the other parameter(s)	Reference
		Two-parameter models		
1	$\alpha(1-\beta^t)$, $\alpha > 0$, $0 < \beta < 1$	α	β: inverse growth rate	49 (eqn. 4.2.5)
2	$\alpha(1-e^{-\beta t})$, $\alpha,\beta > 0$	α	β: growth rate	50
3	$\alpha(1-e^{-t})^{\beta}$, $\alpha,\beta > 0$	α	β: inverse growth rate	51
4	$\frac{\alpha t}{\beta + t}$, $\alpha,\beta > 0$	α	β: inverse growth rate	52 (eqn. 2.22) 53 (Eurocode model)
5	$\frac{\alpha\beta t}{1+\beta t}$, $\alpha,\beta > 0$	α	β: growth rate	49 (eqn. 4.2.16)
6	$\frac{t}{1+\alpha t+\beta\sqrt{t}}$	1/α	β: inverse growth rate	49 (eqn. 4.3.17)
7	$\alpha \tanh\sqrt{\frac{t}{\beta}} = \alpha\frac{e^{\sqrt{t/\beta}} - e^{-\sqrt{t/\beta}}}{e^{\sqrt{t/\beta}} + e^{-\sqrt{t/\beta}}}$	α	β: inverse growth rate	10,11 (RILEM B3 model)
		Three-parameter models		
8	$\alpha t^{\beta t^{-\gamma}}$, $\alpha,\beta,\gamma > 0$	α	Not easily interpreted	49 (eqn. 4.3.15)
9	$\frac{\alpha\beta t^{1-\gamma}}{1+\beta t^{1-\gamma}}$, $\alpha,\beta > 0$, $0 < \gamma < 1$	α	Not easily interpreted	49 (eqn. 4.3.10)

Table 5.1 Candidate growth curve models for the shrinkage profiles (continued)

Model #	f(t) =	$\lim_{t\to\infty} f(t) =$	Interpretation of the other parameter(s)	Reference
10	$\alpha[1-\exp(-\beta t)]^{\gamma}, \qquad \alpha,\beta,\gamma >$	α	β: growth rate γ: shape parameter	49 (eqn. 4.3.35)
11	$\alpha\left[\frac{t^{\beta}}{1+t^{\beta}}\right]^{\gamma}, \quad \alpha,\beta,\gamma>0$	α	β: growth rate γ: shape parameter	51
12	$\alpha\left[\frac{t^{\gamma}}{\beta+t^{\gamma}}\right], \quad \alpha,\beta,\gamma>0$	α	β: growth rate γ: shape parameter	9 (ACI 209 model) 16
13	$\alpha\left[\frac{t}{\beta+t}\right]^{\gamma}, \quad \alpha,\beta,\gamma>0$	α	β: inverse growth rate γ: shape parameter	12 (CEB MC90-99 model) 13 (GL2000 model)
14	$\alpha\left[1-\frac{1}{(t+1)^{\beta}}\right]^{\gamma}, \qquad \alpha,\beta,\gamma >$	α	β, γ: growth rate	51
15	$\alpha\left[1-\frac{1}{1+\left(\frac{t}{\beta}\right)^{\gamma}}\right]^{\gamma}, \qquad \alpha,\beta,\gamma >$	α	β: growth rate γ: shape parameter	54

5.2 Selection of "key" shrinkage profiles for the evaluation of all growth curve models

A subset of ten "key" shrinkage profiles was selected from the 245 profiles in the data set, to which all fifteen growth curve models were eventually fitted. These "key" profiles, illustrated in Figure 5.6, were selected according to criteria chosen to identify the best quality shrinkage data spanning a range of ultimate shrinkage and shrinkage rate values:

- Each profile should comprise a relatively large number of data points.
- Each profile should clearly trend towards an asymptotic value.
- The chosen profiles should include both those with relatively low and high levels of ultimate shrinkage.
- The chosen profiles should include both those with relatively low and high rates of shrinkage development with time.

These criteria were chosen with the objective that any suitable growth curve model should, in the first instance, provide a good fit to the best quality data in the data set.

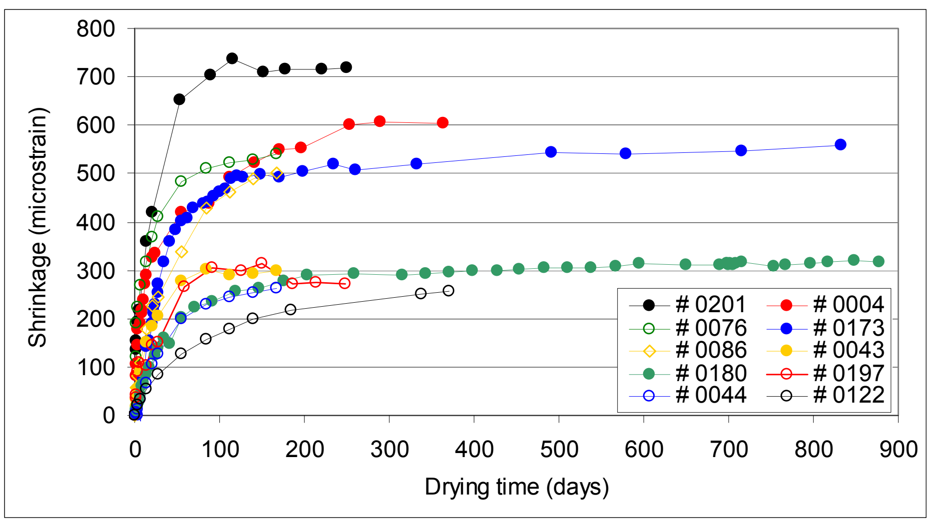


Figure 5.1 The ten "key" shrinkage profiles.

5.3 Variance heterogeneity and autocorrelation in the shrinkage-time data

Before the growth curves identified in Section 5.1 could be fitted to the individual shrinkage-time profiles, provision had to made in the hierarchical nonlinear model for possible autocorrelation between the shrinkage measurements in any given profile as well as the heterogeneous response variance (heteroscedasticity).

5.3.1 Theoretical basis for the modelling of intra-experiment autocorrelation and variance heterogeneity

The approach outlined here is based on that of Davidian and Giltinan[8] and is a continuation of the hierarchical nonlinear model outlined in Chapter 4. For simplicity, the index *i* for the individual experiment is omitted in this section. The basic nonlinear regression model for the j^{th} observation (j = 1, …, n) may be written as follows:

$$y_j = f(t_j, \boldsymbol{\beta}) + e_j$$

where

- t_j is the j^{th} value of time, which is taken to be a fixed (non-random) quantity,
- $\boldsymbol{\beta}$ is a (p x 1) vector of regression parameters (from the chosen growth curve equation) and
- the random errors e_j represent the variability in the measured shrinkage around its predicted value.

In order to accommodate **intra-experiment correlation** among observations, we first assume a correlation pattern among the elements of $\mathbf{e} = [e_1, \ldots, e_n]'$:

$$Corr(\mathbf{e}) = \boldsymbol{\Gamma}(\boldsymbol{\alpha})$$

where the correlation matrix $\boldsymbol{\Gamma}(\boldsymbol{\alpha})$ may be a function of an (s x 1) vector of correlation parameters, $\boldsymbol{\alpha}$. If the variance is constant across the response range and equal to some value σ^2, then

$$Cov(\mathbf{e}) = \sigma^2 \mathbf{\Gamma}(\boldsymbol{\alpha}).$$

Since the repeated measurements were taken over time, it makes sense to consider an autoregressive model of order one (AR(1)) for the serial correlation pattern. In the simplest case where the measurements are equally spaced in time (which is not the case for our data) and the correlation between any two adjacent observations is some value α, $|\alpha| < 1$, the correlation between the j_1-th and the j_2-th observation is given by

$$Corr(e_{j_1}, e_{j_2}) = \alpha^{|j_1 - j_2|}.$$

This may be further expressed by the correlation matrix

$$\mathbf{\Gamma}(\boldsymbol{\alpha}) = \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ & 1 & \alpha & \cdots & \alpha^{n-2} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \alpha \\ & & & & 1 \end{bmatrix}.$$

Where the observations are not equally spaced in time (as in this study), the (j_1,j_2) and (j_2,j_1) elements ($j_1 \neq j_2$) of $\mathbf{\Gamma}(\boldsymbol{\alpha})$ can be replaced by one of the following[8]:

(i) $\quad \alpha^{|t_{j1} - t_{j2}|}$

(ii) $\quad exp\left[-\alpha_1 \left|t_{j_1} - t_{j_2}\right|^{\alpha_2}\right], \ \alpha_2 = 1 \ or \ 2$

(iii) $\quad 1, \ (j_1 = j_2) \qquad \alpha, \ (j_1 = j_2 - 1) \qquad 0, otherwise$

Davidian and Giltinan[8] note that, without a large number of observations covering a broad range of the response, practical implementation of this concept may be difficult and cause complicated interplay among the components of the overall response model.

In order to accommodate **intra-experiment variance heterogeneity**, use is made of a variance function g such that

$$E(y_j) = \mu_j = f(t_j, \beta), \ Var(y_j) = \sigma^2 g^2(\mu_j, z_j, \theta),$$

i.e.
$$\begin{aligned} y_j &= f(t_j, \boldsymbol{\beta}) + e_j & &= \mu_j + e_j \\ where \ \ E[e_j] &= 0 & and \quad & var[e_j] = \sigma^2 g^2(\mu_j, z_j, \boldsymbol{\theta}) \end{aligned}$$

where the variance function may depend on one or more of

- the mean response,
- constants z_j which may include some or all of the components of t_j and
- a q-dimensional parameter vector **θ**.

Common variance functions include the following[8]:

$$
\begin{array}{lll}
(i) & g(\mu_j, z_j, \theta) = \mu_j^{\theta}, & \theta > 0 \\
(ii) & g(\mu_j, z_j, \theta) = exp[\mu_j \theta], & \theta > 0 \\
(iii) & g(\mu_j, z_j, \theta) = exp[t_j \theta], & \theta > 0 \\
(iv) & g^2(\mu_j, z_j, \theta) = \theta_1 + \mu_j^{2\theta_2}, & \theta_1, \theta_2 > 0
\end{array}
$$

Since one cannot specify values *a priori* for **θ**, a data-driven approach to the estimation of (an) appropriate value(s) is used. In general, it is desirable to achieve a parsimonious description of the main features of heterogeneous variance, with one or two variance parameters.

In order to accommodate **BOTH correlation among measurements and heteroscedasticity**, the correlation structure and variance function must be specified simultaneously. For the variance function g, we define the diagonal variance matrix

$$G(\boldsymbol{\beta}, \boldsymbol{\theta}) = diag\left[g^2(\mu_1, z_1, \boldsymbol{\theta}), \ldots, g^2(\mu_n, z_n, \boldsymbol{\theta})\right]$$

If the model for the correlation pattern is given by the matrix $\boldsymbol{\Gamma}(\boldsymbol{\alpha})$, then we can specify

$$
\begin{aligned}
Cov(\mathbf{e}) &= \sigma^2 \mathbf{G}^{1/2}(\boldsymbol{\beta}, \boldsymbol{\theta}) \boldsymbol{\Gamma}(\boldsymbol{\alpha}) \mathbf{G}^{1/2}(\boldsymbol{\beta}, \boldsymbol{\theta}) \\
&= \mathbf{R}(\boldsymbol{\beta}, \boldsymbol{\xi}), \qquad \boldsymbol{\xi} = [\sigma, \boldsymbol{\theta}', \boldsymbol{\alpha}']
\end{aligned}
$$

which implies that

$$Var(y_i) = \sigma^2 g^2(\mu_j, z_j, \boldsymbol{\theta}), \qquad Corr(y_{j1}, y_{j2}) = \boldsymbol{\Gamma}_{(j1, j2)}(\boldsymbol{\alpha}).$$

This general covariance structure allows separate consideration of heterogeneity of variance and the pattern of correlation and is thus very flexible.

Davidian and Giltinan[8] advise against the estimation of <u>all</u> the covariance parameters - i.e. variance heterogeneity <u>and</u> correlation between responses - at the individual experiment level due to sparseness of data. The authors suggest combining information across experiments by pooling residuals across individuals. In the case of the data used in this study, however, this is not possible since data from replicated experiments is not available. It was thus decided initially to **assume an uncorrelated intra-experiment error structure** and to assess the correlation of the residuals after fitting the growth curve models to determine whether this assumption was reasonable in terms of model fitting.

5.3.2 Modelling of the intra-experiment variance heterogeneity

The relationship between the shrinkage variance and the mean shrinkage was modelled to derive a form and parameters for the variance function. Of the 245 experiments in the data set, raw data for the six sets of shrinkage measurements (three from each of two specimens) from which the final shrinkage values were calculated, were available for 77 experiments. (For the other experiments, only the mean shrinkage values at a given drying time were available.) These 77 experiments included

- a variety of cement, stone, sand and slag types,
- a range of water / binder ratios and
- experiments with fairly long drying times (up to 428 days),

thus covering a variety of experimental conditions.

A plots of the shrinkage variance against the mean shrinkage (at a given drying time) is shown in Figure 5.2 for some of the experiments. A full set of plots, as well as the raw data, may be found in the spreadsheet file “Variance heterogeneity.xls” on the accompanying CD.

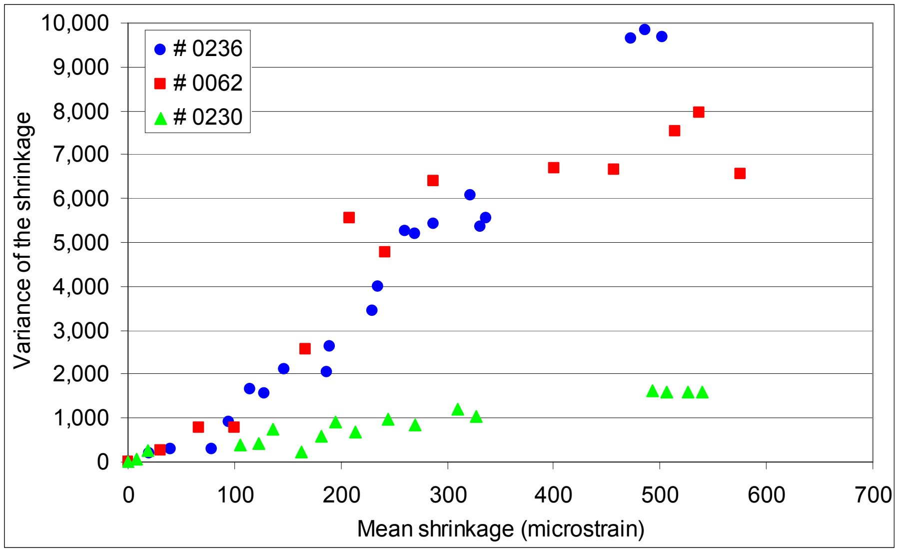


Figure 5.2 Plot of the shrinkage variance against the mean shrinkage for some of the experiments

The plots showed the following:

- The variance increased with shrinkage, with the exception of one experiment (# 0173, the only experiment in data set 13) where the variance remained constant at high levels over the whole range of shrinkage values covered by the experiment. This experiment was excluded from further analysis in this section of work, since it was one of the earliest in the data set (1979) and may have been subject to measurement errors not representative of the bulk of the data set.
- The slope and shape of the variance-shrinkage relationship varied from experiment to experiment.
- There were no obvious trends in the variance-shrinkage relationship which could be related to the covariates mentioned above.

The error of the reading of the Demec strain gauge, used to determine the length of the specimens, remains constant and therefore becomes proportionally smaller with increasing shrinkage. The data however showed that other sources of variation, e.g. variation between specimens and between measurement targets on the same specimen, were more important and led to increasing variance of the shrinkage with increasing mean shrinkage.

An exponential model for the variance-shrinkage relationship was chosen as this appeared to suit most of the variance-shrinkage plots well. For one experiment, this relationship may be expressed as

$$s_j^2 = a\mu_j^b \varsigma$$

or $$\ln(s_j^2) = \theta_1 + \theta_2 \ln(\mu_j) + e$$

where

- s_j^2 is the variance at time j = 1..m,
- μ_j is average shrinkage at drying time j,
- θ_1 = $ln(a)$ and θ_2 = b are constants and
- ς is a multiplicative and e an additive error term.

When there are many experiments, i = 1..n, the relationship may be expressed more generally as

$$\ln(s_{ij}^2) = \theta_{1,0} + \sum_{i=2}^{n} x_i \theta_{1,i} + \theta_2 \cdot \ln(\mu_{ij}) + e_{ij}$$

where

- s_{ij} is standard deviation of the j^{th} shrinkage measurement in experiment i,
- μ_{ij} is the average of the j^{th} shrinkage measurement in experiment i, estimated by $\hat{\mu}_{ij} = \bar{y}_{ij}$ (the average of the six shrinkage values at (i,j)),
- $\theta_{1,i}$ is the fixed effect representing experiment i (allowing a different slope for each experiment),
- x_i is an indicator variable: x_i =1 for experiment i; x_i =0 otherwise; and
- $\theta_{1,0}$ is the intercept term,
- the first experiment (i = 1) is arbitrarily chosen as the base category (x_1 =0) to allow estimation of all the other parameters.

This model was fitted to the data from the 76 experiments by multiple regression. The results are summarised in Table 5.2.

Table 5.2 Results of the regression to estimate the parameters of the variance function

Number of observations and independent variables	
Number of observations	1 111
Number of independent variables ($\theta_{1,0}$, θ_2 and $\theta_{1,i}$, i = 2..76)	77
Ratio of number of observations to number of independent variables	~ 14
Results of regression	
Overall significance of model	$F_{(76,1033)} = 81.67$ ($p < 0.0001$)
R^2	0.86
Adjusted R^2	0.85
Estimated parameters (± 95% confidence interval)	
$\theta_{1,0}$	2.15 (± 0.25) ($p < 0.00001$)
$\theta_{1,i}$ (76 estimates)	Range -1.96 (± 0.33) ($p < 0.00001$) to 1.64 (± 0.33) ($p < 0.00001$) with parameters between -0.31 and 0.31 not significant at the 5% level
θ_2	1.02 (± 0.03) ($p < 0.00001$)

Given the large number of observations, small effects would be regarded as statistically significant, even though they may have no practical significance. The generalisability of the regression (defined as the ratio of the number of observations to the number of independent variables) was close to the acceptable range of 15-20[55]. The distribution of the dependent variable, $\ln(s_{ij}^2)$, shown in Figure 5.3, was reasonably normal.

Calculations for skewness and kurtosis and the Kolmogorov-Smirnov test for normality were not carried out since they do not apply in situations with such a large number of observations[55]. The assumption of normality is not directly on the dependent variable, but rather conditionally on the mean function of the regression, which was checked after the regression analysis by examining the distribution of the residuals.

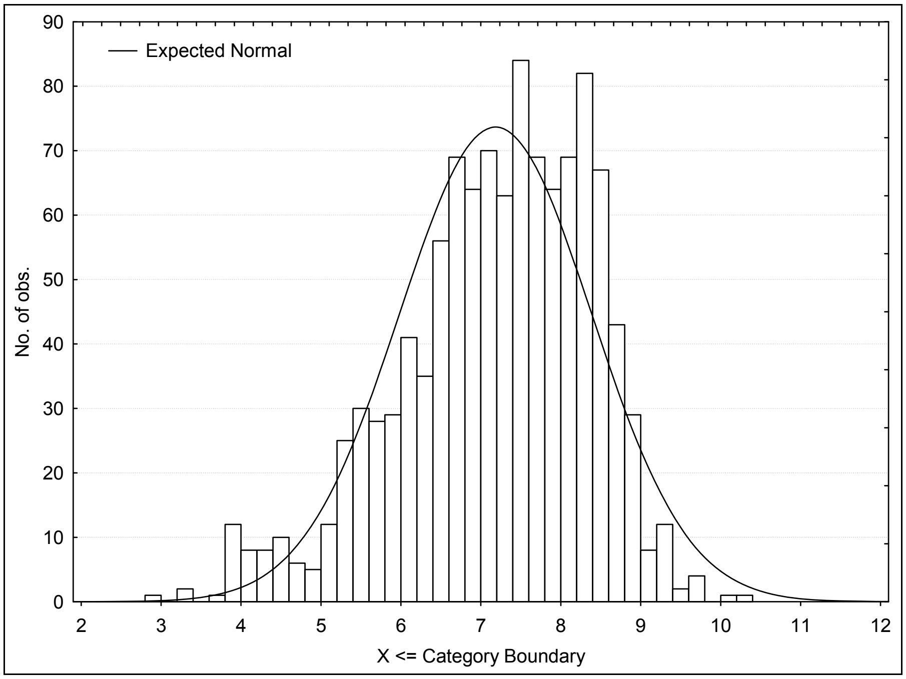


Figure 5.3 Frequency distribution of the dependent variable, ln(variance)

The regression model was highly significant and accounted for 86% of the variation in the data, which was deemed to be acceptable. Plots of the residuals versus the observed and predicted values as well plots of the standardised residuals versus the independent variable ln(μ_j) were all acceptable, as shown in Figure 5.4 (a)-(c). The plot of the residuals versus the observed values has a slope of 1-R^2, as expected from theory. The plot of the residuals versus the predicted values showed no evidence of heteroscedasticity. The partial residual plot for ln(μ_j), shown in Figure 5.4 (d), showed no evidence of outliers or a nonlinear relationship. Similar plots for the indicator independent variables (not shown) were also satisfactory. The normality of the error term distribution was also acceptable, as shown in Figure 5.5.

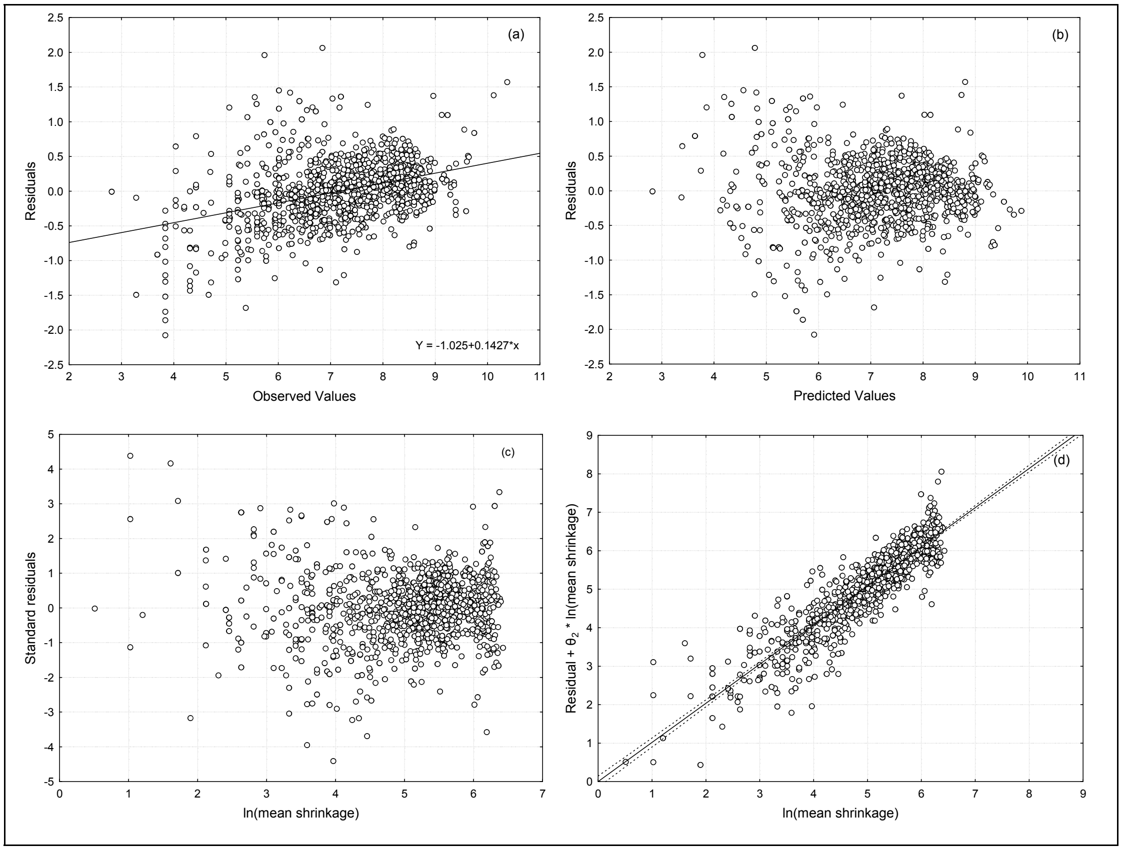


Figure 5.4 Plots of the residuals of the ln(s_{ij}^2) model versus the (a) observed and (b) predicted values (c) Plot of the standardised residuals versus the independent variable ln(μ_j) (d) Partial residual plot for ln(μ_j)

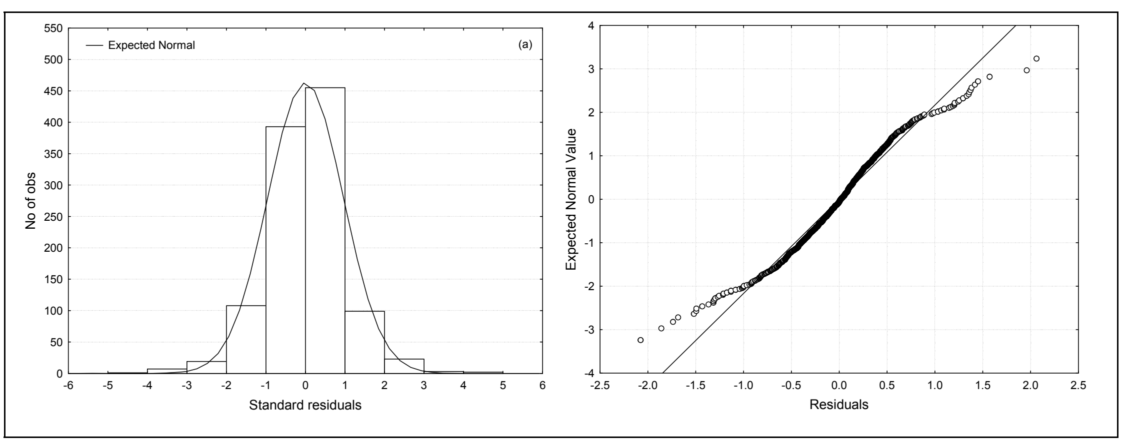


Figure 5.5 (a) Frequency distribution of the standard residuals and (b) normal probability plot of the residuals of the ln(s_{ij}^2) model

The intercept ($\theta_{1,0}$) and θ_2 parameters were highly significant, while 41 of the 75 slope parameters ($\theta_{1,i}$) were significant. The overall slope of the ln(variance):ln(average shrinkage) relationship varied from 0.19 to 3.79 depending on the experiment. The variance function could thus be written as

$$\ln(s_{ij}^2) = 2.15 + \hat{\theta}_{1,i} + 1.02 \cdot \ln(\hat{\mu}_{ij})$$

for experiments where the $\theta_{1,i}$ term was significant, and

$$\ln(s_{ij}^2) = 2.15 + 1.02 \cdot \ln(\hat{\mu}_{ij})$$

for experiments where either the $\theta_{1,i}$ term was not significant or where variance data was not available. The model could have been simplified further by setting θ_2 = 1 (since the estimate for this parameter was not significantly different from 1.0) and then re-estimating the other parameters. However, the variance function as given above was retained.

5.4 Weighting of the shrinkage-time data for fitting the growth curve models

Inferential methods for nonlinear regression models[8] are based on least squares. The ordinary least squares (OLS) estimator for $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}_{OLS}$, minimises

$$\sum_{j=1}^{n} \{y_j - f(t_j, \boldsymbol{\beta})\}^2$$

which gives

$$\hat{\sigma}_{OLS}^2 = \frac{1}{n-p} \sum_{j=1}^{n} \{y_j - f(t_j, \hat{\boldsymbol{\beta}}_{OLS})\}^2$$

where *n* represents the number of data points and *p* the number of parameters in the regression. Thus all deviations receive equal weight in the minimisation in accordance with the assumption of constant variance, which implies that all responses are subject to the same degree of variability. However, when heterogeneity of variance is evident (as in the case of this study), OLS estimation may be inefficient relative to methods which do take heteroscedasticity into account. Also, OLS assumes that the y_i are uncorrelated which, in this case, they may not be.

The weighted least squares (WLS) estimator for $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}_{WLS}$, minimises

$$\sum_{j=1}^{n} w_j \left\{y_j - f(t_j, \boldsymbol{\beta})\right\}^2$$

which gives

$$\hat{\sigma}^2_{WLS} = \frac{1}{n-p} \sum_{j=1}^{n} w_j \left\{y_j - f\left(t_j, \hat{\boldsymbol{\beta}}_{WLS}\right)\right\}^2$$

In this approach, each deviation is weighted (by the weight w_j). The weights can be defined to be inversely proportional to the variance of the associated response. However, this requires knowledge of the variance for each y_j which is often not the case. More generally (as in this case) the weights can be defined as a systematic function of the mean or other factors, as described above.

Fitting of nonlinear models is done numerically by means of an iterative algorithm. Typically, software packages implement the Gauss-Newton, Newton-Raphson and Levenberg-Marquardt algorithms, the latter being preferred(52,56).

The goodness-of-fit of different estimators (OLS, WLS) for a given model may be evaluated as follows(8):

- A plot of the studentised residuals versus the predicted values should show a random distribution of the residuals about zero and no trend with respect to the predicted values.
- Plots of the residuals, lagged by k = 1, 2, 3, … observations, against the residuals themselves, should not show significant correlation.
- In order to select the best fitting method once the above criteria are met, the Aikaike's Information Criterion (AIC) or the Bayesian Information Criterion (BIC) should be used. These statistics are defined as follows for least squares estimation(57):

$$AIC = \ln(MSE) + 2p$$
$$BIC = n \cdot \ln(MSE) + p \cdot \ln(n)$$

where MSE is the Mean Square Error, p is the number of parameters in the model and n is the number of observations. In both cases, smaller values indicate a better model fit.

To test whether the variance function developed in the previous section would give the desired results and would make a significant difference to the results of the growth curve model fitting, one of the fifteen growth curve models listed in Table 5.1 (Model 1, a two-parameter model whose fit might change with a different weighting scheme) was fitted to one of the ten selected key shrinkage profiles (see the next section) with and without weighting of the data, i.e. by WLS and OLS respectively. Fitting of the nonlinear model was done using PROC NLIN in the SAS software[(58)]. The Gauss-Newton method was used as it is the default method in PROC NLIN and yielded acceptable performance. The program listing, together with the output, is given in Appendix C.

The WLS fit was carried out by applying the variance function derived in the previous section. The values of the weight variable in PROC NLIN are taken as the inverse elements of the diagonal variance-covariance matrix of the shrinkage variable; larger weights thus carry greater importance. The weight variable was defined as follows:

$$weight = \frac{1}{exp\left(2.15 + 1.02 \cdot \ln(\,average\ shrinkage\,)\right)}$$

SAS software does not normalise the weights, making comparison of the Sums of Squares (SS) (and statistics which use the SS) impossible. Thus, the weights for both the OLS and WLS schemes were normalised (so that they summed to 1) beforehand and the weights were included in the raw data file.

The output from the nonlinear regressions (shown in Appendix C) indicated that the WLS weighting scheme weighted more heavily towards the lower shrinkage values as these had lower variance. Whilst addressing the heteroscedasticity in the data over time, the adoption of such a weighting scheme would have the unfortunate consequence of reducing the importance of the data points at higher shrinkage values (longer drying times). There are typically fewer of these points in a given profile but they are more important in defining the ultimate shrinkage and the rate of shrinkage. The WLS approach downweighted these points.

The results of the nonlinear regression by OLS and WLS are compared in Table 5.3 and Figure 5.6. The residual plots (as listed earlier in this section) showed no discernible difference between the two estimation methods. Although the AIC and BIC statistics favoured the WLS estimation, there was no significant difference between the parameter estimates and no discernible difference between the plots of the fitted curves. Given the undesirable weighting of the observations with WLS, as discussed above, it was concluded that OLS was an adequate estimation method for these data. Since the weighting scheme for Generalised Least Squares (GLS) would also downweight the data points at higher shrinkage values (longer drying times), GLS was also not deemed to be a suitable approach. It was thus decided to **use OLS for all fitting of growth curve models to the shrinkage profiles**.

Table 5.3 Results of the OLS and WLS estimation of the fit of growth curve Model 1 to a selected shrinkage profile (# 0173)

	Estimation method	
	OLS	WLS
Number of observations	38	35 (3 non-positive shrinkage observations lost due to weight calculation)
AIC	6.7	6.3
BIC	108	89
Parameter estimates (asymptotic 95% confidence interval) Shrinkage = $\alpha(1-\beta^t),\quad \alpha > 0,\ 0 < \beta < 1$		
α	525.6 (511.5; 539.6)	527.3 (512.3; 542.4)
β	0.977 (0.975; 0.979)	0.977 (0.976; 0.979)

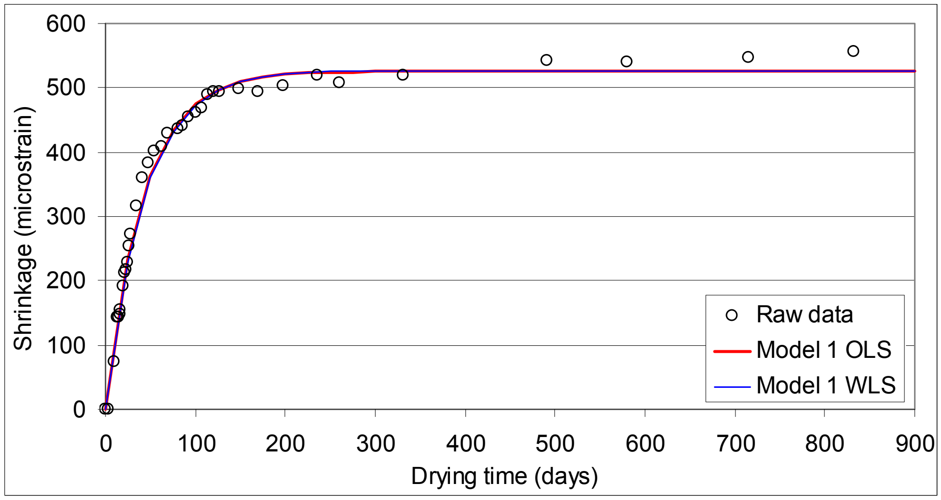


Figure 5.6 Comparison of the OLS and WLS estimation of the fit of growth curve Model 1 to a selected shrinkage profile (# 0173)

5.5 Fitting of all growth curve models to the "key" shrinkage profiles

All fifteen growth curve models were fitted to each of the ten selected shrinkage profiles, using OLS estimation. Fitting of the non-linear models was done using PROC NLIN in the SAS software(58). The program listing is given in Appendix D. Figures 5.7 and 5.8 illustrate the fit of one model (Model 10) to all ten profiles, and the fit of all fifteen models to one shrinkage profile, respectively. Full details may be found in the spreadsheet "Growth curve models – key shrinkage profiles.xls" on the accompanying CD. A summary of the results for one shrinkage profile is given in Appendix E.

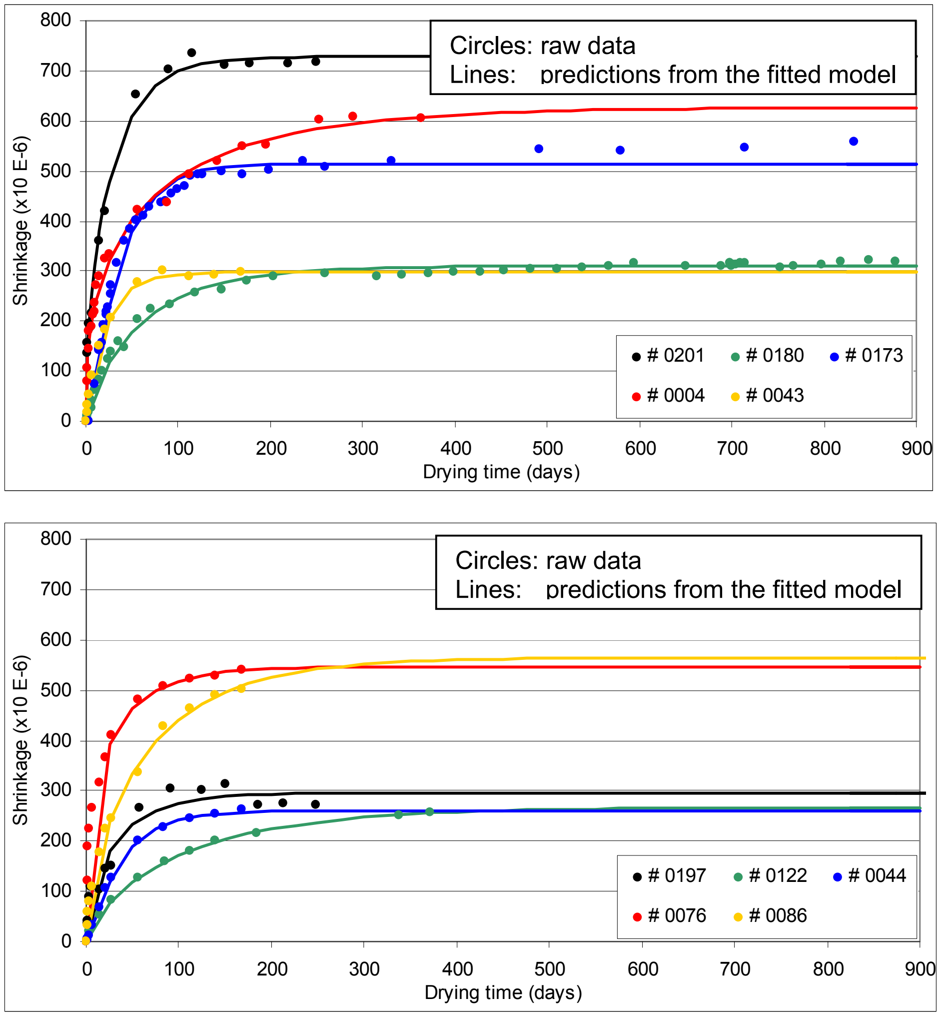


Figure 5.7 Illustration of the fit of one growth curve model (Model 10) to the ten "key" shrinkage profiles

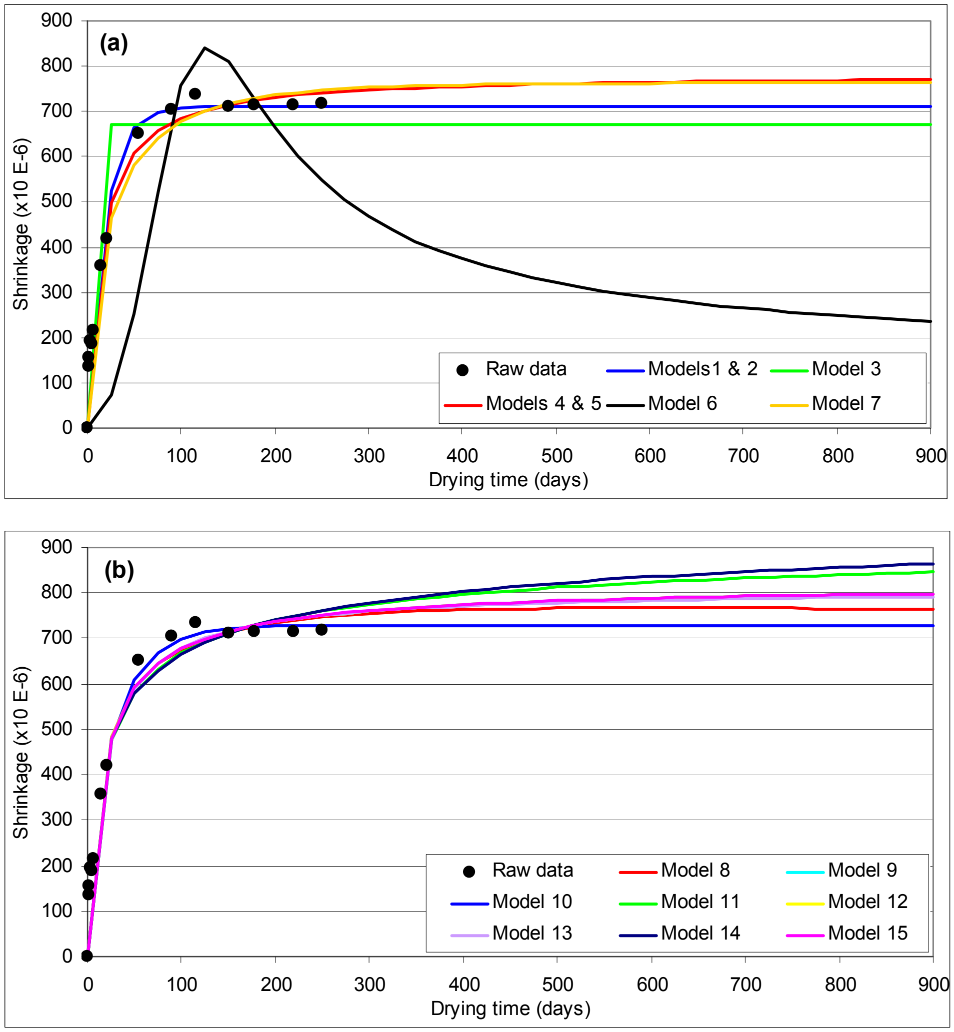


Figure 5.8 Illustration of the fit of the (a) two-parameter and (b) three-parameter growth curve models to one shrinkage profile (# 0201)

Models 3 and 6 proved to be inappropriate to describe the shrinkage profiles, as illustrated in Figure 5.8 (a). Efforts to optimise these models' parameters by finding more suitable starting parameters for the OLS estimation (by means of a grid search), or using different regression algorithms, with different step sizes, failed. These models were eliminated at this point and are excluded from further discussion.

In order to determine which of the remaining thirteen growth curve models gave the best fit to the shrinkage profiles, the following goodness-of-fit diagnostics and measures were used:

Graphical diagnostics:

- Plot of the predicted curve superimposed on the data points: The curve should follow the data points well.
- Plot of observed versus predicted values: This plot should be close to linear.
- Residual plot versus time: This should not show any trend or pattern.
- Residual plot versus predicted values: This should not fan out dramatically (indicating that the variance-shrinkage relationship discussed in section 5.2 ought to be accounted for).
- Normal plot of residuals: This should be close to linear (which may be difficult to achieve in cases where the number of data points is relatively small).

Numerical goodness-of-fit measures:

- The use of R^2, the "proportion of explained variation" does not have meaning in the context of nonlinear regression since it is only in linear models with an intercept in which R^2 genuinely reflects the proportion of variance explained by the model[49]. Instead, a measure closely corresponding to R^2 in the nonlinear case is the Pseudo-R^2 [59]

$$Pseudo\text{-}R^2 = 1 - \frac{SS_{residual}}{SS_{total(corrected)}}\ .$$

  A larger value indicates a better fit.
- The Concordance Correlation Coefficient (CCC) between the observed and predicted values[60], adapted for a single shrinkage profile:

$$r_c = 1 - \frac{(\mathbf{y}_i - \hat{\mathbf{y}}_i)'(\mathbf{y}_i - \hat{\mathbf{y}}_i)}{(\mathbf{y}_i - \bar{y}\mathbf{1}_i)'(\mathbf{y}_i - \bar{y}\mathbf{1}_i) + (\hat{\mathbf{y}}_i - \hat{y}\mathbf{1}_i)'(\mathbf{y}_i - \hat{y}\mathbf{1}_i) + n(\bar{y} - \hat{y})^2}$$

  where

  - $\mathbf{y}_i$ is the vector of observed values in the i^{th} shrinkage profile and $\hat{\mathbf{y}}_i$ is the vector of predicted values,

- $\bar{y}$ is the mean of the observed values and $\hat{y}$ is the mean of the fitted values in the shrinkage profile and
- n is the number of observations in the shrinkage profile.

A larger value indicates a better fit.

- AIC value (as defined in the previous section): This gauges the quality of a model fit by including a penalty term for the number of unknown parameters in the model[52].
- BIC value (as defined in the previous section).
- The Root Mean Square (RMS)[48,61] for which a smaller value indicates a better fit:

$$RMS = \sqrt{SS / df}$$

- The standard errors and significances (p-values) of the parameters: If these are very large, it suggests that that model may be grossly mis-specified or that a local minimum (as opposed to the global minimum) has been found by the loss function optimisation algorithm.
- The correlation between the parameters: A large degree of correlation between parameters suggests that one or more of the parameters may be redundant and that a model with fewer parameters should be sought.

The diagnostics and goodness-of-fit statistics are illustrated in Appendix E for the growth curve models fitted to one of the ten “key” shrinkage profiles.

None of the thirteen growth curve models were found to perform consistently poorly on the graphical diagnostics on all ten shrinkage profiles and thus no further models were eliminated on this basis. Next, for each model, average values (across the ten shrinkage profiles) of the first five goodness-of-fit measures listed above (Pseudo-R^2, CCC, AIC, BIC and RMS) were calculated. These results are presented in Table 5.4 and are illustrated together with their 95% confidence intervals in Figure 5.9 and by means of a radar chart in Figure 5.10.

Table 5.4 Goodness-of-fit scores for each growth curve model

Model	Goodness-of-fit score (95% confidence interval)				
	AIC	BIC	Pseudo-R^2	CCC	RMS
Model 1	10.0 (± 1.3)	128 (± 54)	0.966 (± 0.027)	0.984 (± 0.012)	27.0 (± 14.2)
Model 2	10.0 (± 1.3)	128 (± 54)	0.966 (± 0.027)	0.984 (± 0.012)	27.0 (± 14.2)
Model 4	9.7 (± 1.2)	121 (± 52)	0.978 (± 0.015)	0.989 (± 0.008)	21.9 (± 9.9)
Model 5	9.7 (± 1.2)	121 (± 52)	0.978 (± 0.015)	0.989 (± 0.008)	21.9 (± 9.9)
Model 7	10.1 (± 0.7)	128 (± 62)	0.974 (± 0.016)	0.986 (± 0.009)	23.9 (± 8.4)
Model 8	11.0 (± 1.2)	108 (± 49)	0.988 (± 0.014)	0.993 (± 0.007)	16.4 (± 9.1)
Model 9	11.0 (± 1.4)	111 (± 53)	0.986 (± 0.014)	0.994 (± 0.007)	17.2 (± 9.0)
Model 10	10.6 (± 1.3)	105 (± 56)	0.991 (± 0.010)	0.996 (± 0.005)	13.3 (± 6.5)
Model 11	11.0 (± 1.0)	104 (± 40)	0.988 (± 0.014)	0.994 (± 0.007)	15.9 (± 9.4)
Model 12	10.7 (± 1.3)	100 (± 44)	0.990 (± 0.013)	0.995 (± 0.007)	14.4 (± 8.5)
Model 13	10.7 (± 1.2)	99 (± 43)	0.990 (± 0.012)	0.995 (± 0.006)	13.9 (± 7.8)
Model 14	11.2 (± 1.0)	107 (± 40)	0.986 (± 0.015)	0.993 (± 0.008)	17.5 (± 9.9)
Model 15	10.6 (± 1.3)	99 (± 44)	0.990 (± 0.013)	0.995 (± 0.007)	14.2 (± 8.3)

For each goodness-of-fit statistic, the models with the first, second and third best scores are highlighted in dark orange, light orange and yellow, respectively.

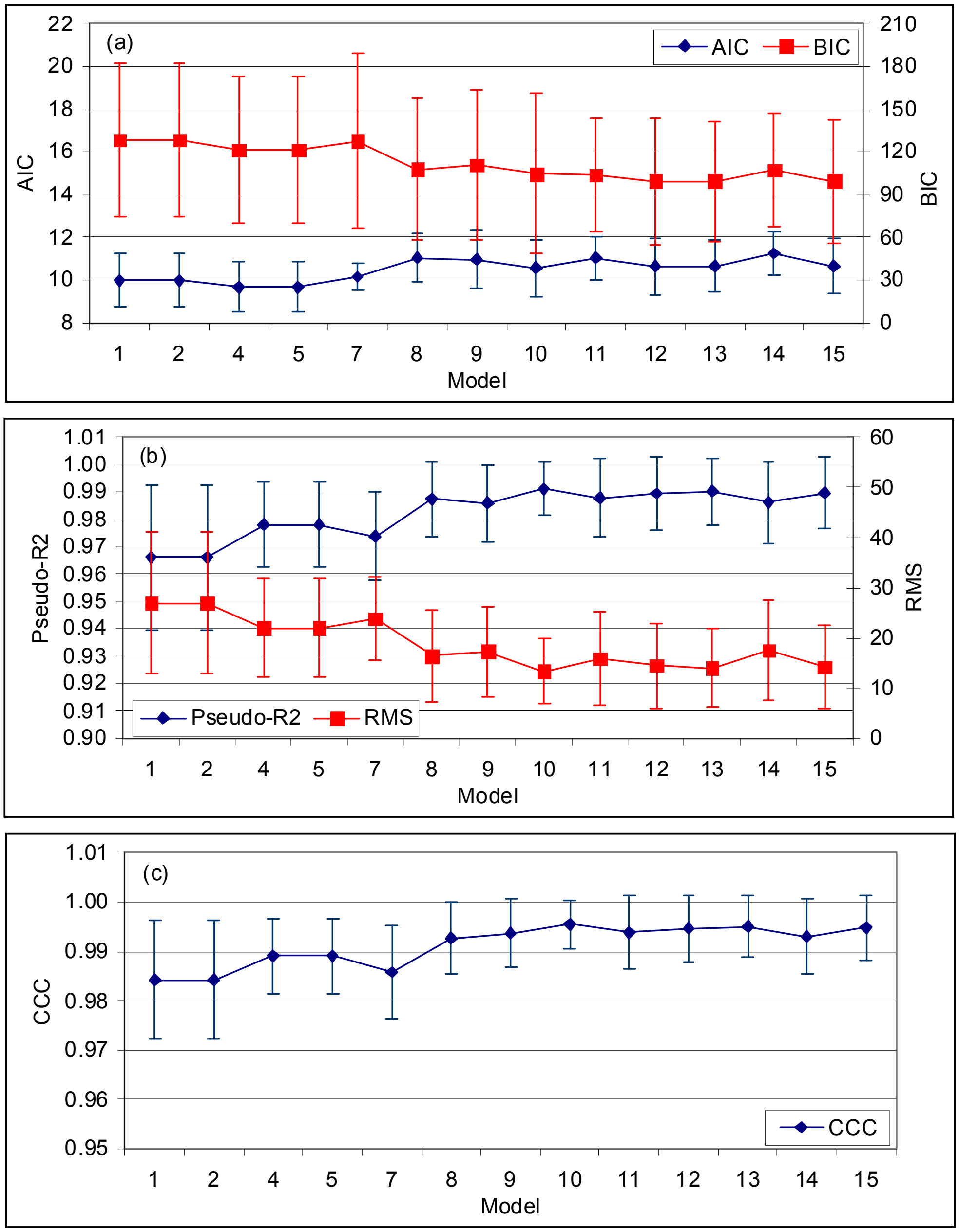


Figure 5.9 Goodness-of-fit scores, together with their 95% confidence intervals, for the growth curve models

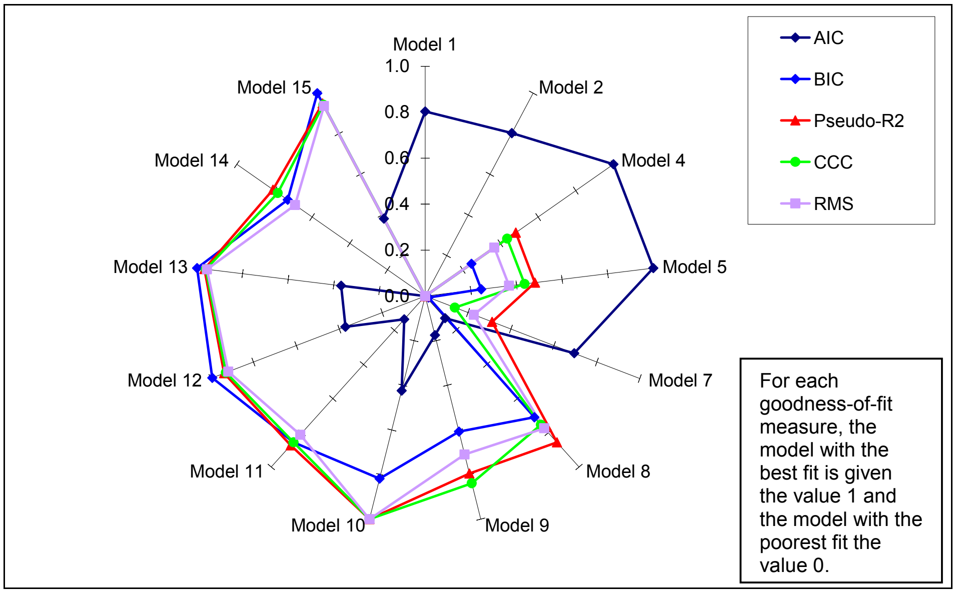


Figure 5.10 Comparison of the goodness-of-fit scores for the growth curve models

For each goodness-of-fit measure, the scores for the different models (averaged over the ten "key' shrinkage profiles) are not significantly different from each other (Figure 5.9). However, considering the means, the penalization on moving from 2-parameter models (models 1-7) to 3-parameter models (models 8-15) can clearly be seen in the AIC values (see Figure 5.10), whereas the other four goodness-of-fit measures show that the three-parameter models give a much better fit than the two-parameter models. Those parameters which give values on an absolute scale (0 to 1), viz. the pseudo-R^2 and the CCC, indicate that all the models fit fairly well (the lowest values are 0.966 and 0.984 for these two measures respectively).

At first glance, models 10, 12, 13 and 15 appear to give the best fit, as determined by the five goodness-of-fit measures (see Figure 5.10), but a 'best model' could not be identified. The following approaches were used to try to select the best model(s):

- For each goodness-of-fit statistic, the models with the first, second and third best scores were highlighted in Table 5.4. A best model could not be identified from this, although Model 10 had the best score in three out of the five goodness-of-fit measures.

- Models were eliminated which, for at least one shrinkage profile, gave a fit to the raw data which appeared unacceptable (by inspection) OR for which the confidence interval for a parameter included zero OR for which the predicted shrinkage continued to increase substantially with time (this was established by comparing the 5- and 10-year predictions). This left only Models 10 and 13 which were not eliminated. Model 15 was eliminated only on one criterion, and Model 12 only on two criteria.
- Models were accepted, for a particular profile, which had Pseudo-R^2 and CCC values in the top 0.01 units of the range of values for these statistics (for each profile) AND AIC and RMS values within 33% and 50%, respectively, of the lowest values for these statistics (for each profile) AND relative standard errors for all the parameters below 40%. In this case, Model 15 performed best, being excluded only on one criterion for one shrinkage profile, followed by Model 12 which was also excluded on one profile only and Model 10 which was excluded on two shrinkage profiles.

It thus appeared that Models 10, 12, 13 and 15 gave the best fit to the ten “key” shrinkage profiles. In order to formalize the model selection process, MCDA was used to identify the best models, based on the five goodness-of-fit measures.

5.6 Selection of the best growth curve models by Multiple Criteria Decision Analysis

Belton and Stewart[62] define MCDA as “an umbrella term to describe a collection of formal approaches which seek to take explicit account of multiple criteria in helping individuals or groups explore decisions that matter”. Two such approaches were selected for the identification of the most suitable growth models for the shrinkage-time profiles, viz.

- the Value Function method and
- outranking.

A third major MCDA approach, satisficing, was excluded on the basis that this technique requires an *ordering* of the evaluation criteria (in this case the goodness-of-fit measures) in order of importance. It was felt that the five goodness-of-fit criteria were of equal importance in this context and hence the satisficing technique was not used.

5.6.1 The Value Function method

The outline of the Value Function method is based on that of Belton and Stewart[62]. In the case of this study, the alternatives are the growth curve models, and the criteria are the five goodness-of fit measures. For each alternative, *a*, a value *V(a)* is associated with *a* such that *a* is judged to be preferred to alternative *b* $(a \succ b)$, taking all criteria into account, if and only if *V(a)* > *V(b)*. Alternatives *a* and *b* are indifferent ($a \sim b$) if *V(a)* = *V(b)*.

Firstly, the partial (marginal) value function $v_i(a)$ for each criterion must be established. Then a weighted sum of the partial value functions is calculated to obtain the overall value function for each alternative:

$$V(a) = \sum_{i=1}^{5} w_i v_i(a) \qquad a = 1..13$$

The partial value function is determined by describing the ‘best’ and ‘worst’ outcomes for a particular criterion and then standardising the results to a scale of 0 to 100 where 0 represents the worst, and 100 the best, outcome respectively. The ‘best’ and ‘worst’ outcomes for numerical criteria such as the ones used in this study may be defined *locally*, as the maximum and minimum values achieved for a particular criterion, or *globally*, as the upper and lower end of a defined scale for a particular criterion. Since only the pseudo-R^2 and CCC criteria had defined ranges (0 to 1) but their actual values were clustered near the upper end of these ranges, it was decided to use local scaling for all five criteria. Equal weights were assigned to all five criteria. The partial and overall value functions for the thirteen growth curve models are given in Table 5.5 and the overall value functions are compared in Figure 5.11.

Table 5.5 Partial and overall value functions for the thirteen growth curve models

Model	Standardised partial value functions					Value function
	AIC	BIC	Pseudo-R^2	CCC	RMS	
Model 1	80	0	0	0	0	16
Model 2	80	0	0	0	0	16
Model 4	100	25	48	44	37	51
Model 5	100	25	48	44	37	51
Model 7	69	2	31	14	23	28
Model 8	13	71	86	75	77	64
Model 9	17	61	80	84	71	63
Model 10	42	82	100	100	100	85
Model 11	14	86	87	86	81	71
Model 12	37	99	94	93	92	83
Model 13	37	100	97	96	96	85
Model 14	0	73	81	78	69	60
Model 15	38	100	95	94	94	84

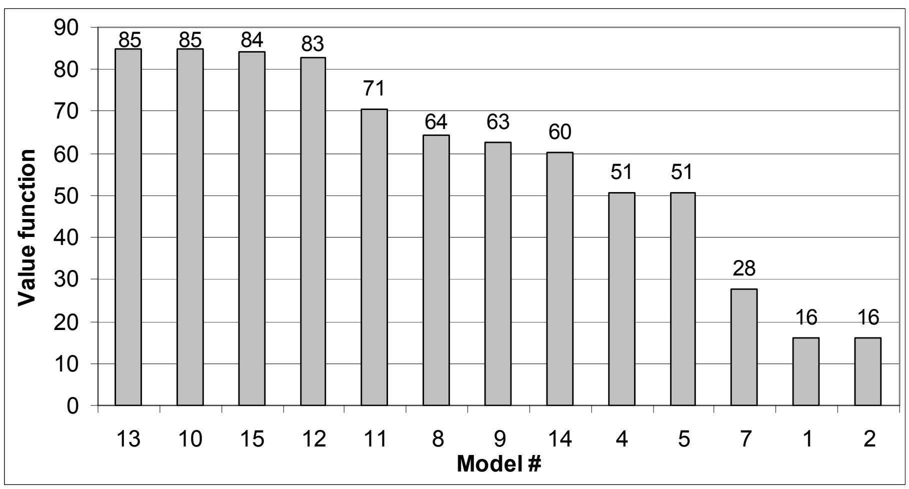


Figure 5.11 Comparison of the value functions for the growth curve models

Models 10 and 13 obtained the highest value function scores, closely followed by models 15 and 12, after which there was a substantial decrease in value function scores for the rest of the models. This confirmed the results of the less formal methods of the previous section. Models 10, 12, 13 and 15 thus gave the best fits to the “key” shrinkage profiles, but a decision for one “best” model could not be made on the basis of the value measurement results since the top two models had identical value function scores.

5.6.2 The Outranking method

Once again, the approach outlined by Belton and Stewart(62) was followed. We continue with the definitions and notation from the previous section. The outranking approach differs from the value function approach in that there is no value function which aggregates the contribution of the various criteria. The output of the analysis is not a value for each alternative, but an outranking relation on the alternatives. An alternative *a* is judged to outrank an alternative *b* if, taking the criteria and the decision maker's preferences into account, there is

- a sufficiently strong argument to support a conclusion that *a* is at least as good as *b*, and
- no strong argument to the contrary.

The particular method by which the outranking relationship was defined and implemented in this study was the ELECTRE I method(63), the earliest and simplest of outranking methods, as presented by Belton and Stewart(62).

The ELECTRE I method is based on the evaluation of two indices, defined for each pair of alternatives *a* and *b*:

- The Concordance Index, *C(a,b)* measures the strength of support within the criteria for the hypothesis that *a* is at least as good as *b*:

$$C(a,b) = \frac{\sum_{i \in Q(a,b)} w_i}{\sum_{i=1}^{5} w_i}$$

  where *Q(a,b)* is the set of criteria for which *a* is equal to or preferred to *b*. The Concordance Index is thus equal to the proportion of criteria weights for which *a* is equal to or better than *b*.

- The Discordance Index, *D(a,b)* measures the strength of the evidence *against* the above hypothesis:

$$D(a,b) = \begin{cases} 1 & \text{if } z_i(b) - z_i(a) > t_i \text{ for any } i \\ 0 & \text{otherwise} \end{cases}$$

where the z_i are individual criterion scores. In other words, a veto threshold, t_i is defined for each criterion, such that *a* cannot outrank *b* if the score for *b* on any criterion exceeds the score for *a* on that criterion by an amount equal to or greater than its veto threshold.

The concordance and discordance indices, calculated for each pair of alternatives, can then be used to construct an outranking relation. In order to do this, concordance and discordance *thresholds*, C* and D*, respectively, must be specified. Alternative *a* is said to outrank alternative *b* if the calculated concordance index is greater than or equal to the concordance threshold AND the calculated discordance index is less than or equal to the discordance threshold. The necessary comparisons and their outcomes are summarised in Table 5.6.

Table 5.6 Comparisons used to construct the outranking relationship

			Comparing *a* with *b*		
			C(a,b) ≤ C*	C(a,b) ≤ C*	
				D(a,b) ≥ D*	D(a,b) < D*
Com-paring *b* with *a*	C(b,a) ≤ C*		incomparable	incomparable	*a* outranks *b*
	C(b,a) > C*	D(b,a) ≥ D*	incomparable	incomparable	*a* outranks *b*
		D(b,a) < D*	*b* outranks *a*	*b* outranks *a*	indifference: *a* outranks *b*, *b* outranks *a*

Some experimentation with the values of C* and D* (and the thresholds t_i) is required to define an outranking relation which does not have too many incomparable or indifferent pairs of alternatives.

In the case of this study, the parameters defined above were set as follows:

- For the concordance index, the five goodness-of-fit criteria were weighted equally, as before.
- For the discordance index, twice the average standard error for each criterion (averaged over the thirteen models) was used as the veto threshold.

- The concordance threshold was set at C* = 0.7, after some experimentation, to give the clearest separation between the models (the values of the concordance index range from 0 to 1).
- The discordance threshold was set at D* = 0.1 (the discordance index can take on values of only 0 or 1).

The tabulated concordance and discordance indices and the outcome of the pairwise outranking comparisons are presented in Tables 5.7 and 5.8 respectively. The outranking relationship between the thirteen growth curve models is depicted in Figure 5.12.

Table 5.7 Concordance and discordance indices for the ELECTRE I outranking method

Concordance Index (a=column, b=row)													
Model	**1**	**2**	**4**	**5**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1		1.0	1.0	1.0	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
2	1.0		1.0	1.0	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
4	0.0	0.0		1.0	0.0	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
5	0.0	0.0	1.0		0.0	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
7	0.4	0.4	1.0	1.0		0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
8	0.2	0.2	0.2	0.2	0.2		0.4	1.0	1.0	1.0	1.0	0.4	1.0
9	0.2	0.2	0.2	0.2	0.2	0.6		1.0	0.8	1.0	1.0	0.4	1.0
10	0.2	0.2	0.2	0.2	0.2	0.0	0.0		0.2	0.2	0.2	0.0	0.2
11	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.8		1.0	1.0	0.0	1.0
12	0.2	0.2	0.2	0.2	0.2	0.0	0.0	0.8	0.0		0.8	0.0	1.0
13	0.2	0.2	0.2	0.2	0.2	0.0	0.0	0.8	0.0	0.2		0.0	0.4
14	0.2	0.2	0.2	0.2	0.2	0.6	0.6	1.0	1.0	1.0	1.0		1.0
15	0.2	0.2	0.2	0.2	0.2	0.0	0.0	0.8	0.0	0.0	0.8	0.0	

Discordance Index (a=column, b=row)													
Model	**1**	**2**	**4**	**5**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1		0	0	0	0	1	0	0	1	0	0	1	0
2	0		0	0	0	1	0	0	1	0	0	1	0
4	0	0		0	0	1	1	0	1	0	0	1	0
5	0	0	0		0	1	1	0	1	0	0	1	0
7	0	0	0	0		0	0	0	0	0	0	1	0
8	1	1	0	0	0		0	0	0	0	0	1	0
9	1	1	0	0	1	0		0	0	0	0	0	0
10	1	1	1	1	1	0	0		0	1	0	1	0
11	1	1	0	0	1	0	0	0		0	0	1	0
12	1	1	0	0	1	0	0	0	0		0	1	0
13	1	1	0	0	1	0	0	0	0	0		1	0
14	1	1	0	0	1	0	0	0	0	0	0		0
15	1	1	0	0	1	0	0	0	0	0	0	1	

Table 5.8 Outranking relation for the pairs of alternatives (growth curve models) according to the ELECTRE I method

	Outranking relation (a=column, b=row)											
Model	**1**	**2**	**4**	**5**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
2	1>2 2>1 indiff											
4	4>1	4>2										
5	5>1	5>2	4>5 5>4 indiff									
7	7>1	7>2	4>7	5>7								
8	inc	inc	inc	inc	inc							
9	9>1	9>2	inc	inc	9>7	inc						
10	10>1	10>2	10>4	10>5	10>7	10>8	10>9					
11	inc	inc	inc	inc	11>7	11>8	11>9	10>11				
12	12>1	12>2	12>4	12>5	12>7	12>8	12>9	10>12	12>11			
13	13>1	13>2	13>4	13>5	13>7	13>8	13>9	10>13	13>11	13>12		
14	inc	inc	inc	inc	inc	inc	inc	10>14	11>14	12>14	13>14	
15	15>1	15>2	15>4	15>5	15>7	15>8	15>9	10>15	15>11	15>12	13>15	15>14

The symbol ">" indicates outranking, "inc" indicates an incomparable relationship while "indiff" indicates an indifferent relationship.

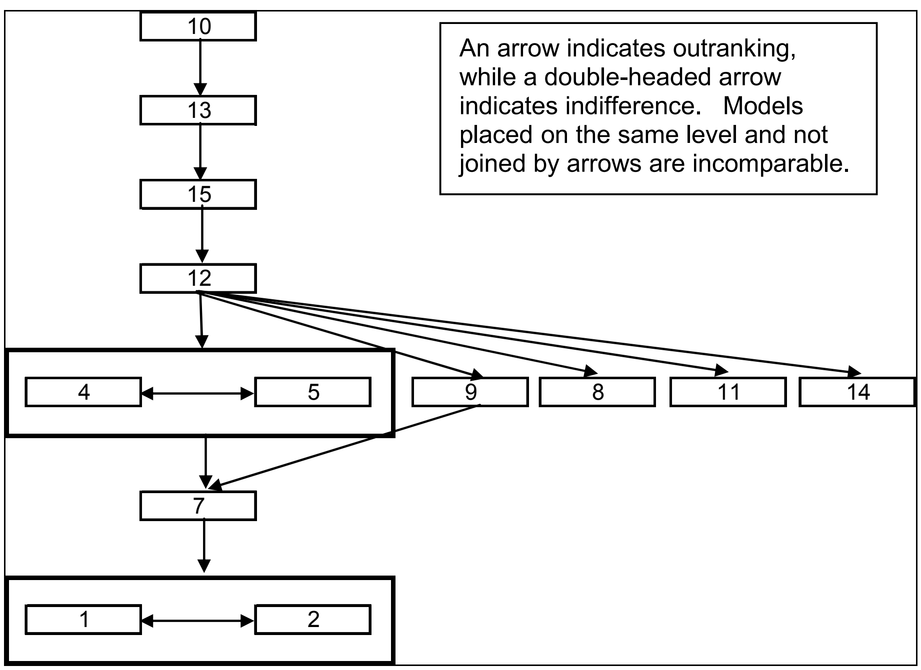


Figure 5.12 The outranking relationship for the thirteen growth curve models according to the ELECTRE I method

The results showed that the same four models, viz. Models 10, 12, 13 and 15, were ranked as the best performing models. In contrast to the value function method, Models 10 and 13 were separated by the outranking method: Model 10 outranked Model 13. Models 15 and 12 followed in the same order as observed in the value function method. There was a large number of incomparable (as well as indifferent) pairwise comparisons lower down in the outranking relationship (see Figure 5.12) but this was not of concern since the aim was to identify only the best methods.

Since the two MCDA techniques had given very similar, but not identical, results, it was decided to fit the four best growth curve models to ALL the shrinkage time profiles before choosing one 'best' model, for two reasons:

- It was necessary to determine whether there were problems with any of these models which were not identified while fitting these models to the ten "key" shrinkage profiles which represented the best quality data in the data set. This might lead to the elimination of one or more of these profiles.
- The goodness-of-fit criteria, averaged across the *entire* data set, could be used to make a final decision as to which growth curve model was most suitable for the data set. The data set is biased to some extent towards shrinkage profiles ending at relatively short drying times (~ 6 months), a good proportion of which have not yet reached their ultimate shrinkage. This could influence the goodness-of-fit statistics for the different growth curve models. (Alternatively, one could compare the results of a MCDA on the set of profiles for which an asymptote has been reached to the results of an MCDA on all profiles, but this approach was not pursued.)

5.7 Fitting of the four best growth curve models to all the shrinkage profiles

Growth curve models 10, 12, 13 and 15 were fitted to all 245 shrinkage-time profiles in the same manner as described previously. Of these 245 profiles, eighteen (from data set 9) had to be excluded from this and subsequent analyses as their profiles were made up of fewer than four data points – too few to fit a three-parameter model. Detailed results may be found in the spreadsheet "Growth curve models-all shrinkage profiles.xls" and the document "Growth curve models-graphs.pdf" on the accompanying CD.

A number of issues were encountered in the fitting of the four growth curve models to the remaining 227 shrinkage profiles, which were relevant both to the selection of the best model as well as to the selection of shrinkage profiles to be included in the second stage of the hierarchical model (as will be discussed in Chapter 6):

- The fit of the models to some shrinkage profiles did not converge (or, in the case of Model 12, gave rise to a singular Hessian matrix which in turn gave unrealistic values for the estimated parameters of the growth curve). The use of different fitting algorithms and step sizes did not make any difference to these outcomes. The number of cases affected for each growth curve model is tabulated in Table 5.9. This problem occurred for shrinkage profiles which had not yet approached their asymptotic value (ultimate shrinkage), as illustrated in Figure 5.13(a) for one shrinkage profile. Given the similarity in the number of profiles affected for each model, this issue did not affect the choice of growth curve model but rather highlighted the fact that further shrinkage profiles would have to be excluded from the second stage of the hierarchical modelling process as growth curve parameters could not be obtained for these profiles.
- The confidence intervals for some parameter estimates (mostly α and / or β) for some models fitted to certain shrinkage profiles were very wide, indicating a low degree of certainty for the estimation of these parameters. This problem tended to occur when there was sufficient information in the raw data for the model fitting algorithm to converge, but insufficient information for accurate prediction of the ultimate shrinkage and / or the rate of shrinkage (parameters α and β in the growth curves: see Table 5.1). An illustrative example is given in Figure 5.13(b) for one shrinkage profile. The number of cases with one or more parameters having relative standard errors exceeding 30%, for each growth curve model, is tabulated in Table 5.9. The number of cases with poor fit was markedly higher for Models 13 and 15 than for Models 10 and 12, suggesting that Models 10 and 12 were more suited to the shrinkage

profile data than Models 13 and 15. This problem clearly also affects the selection of profiles for inclusion in the second step of the hierarchical model, as will be discussed in Chapter 6.

- It was noted that, whereas shrinkage profiles which contained sufficient long-term data resulted in very similar fits to the raw data by all four models, shrinkage profiles with insufficient long-term data generally resulted in a more stable prediction by Model 10, while the other three models tended to predict increasing shrinkage even at very long drying times: compare Figure 5.13 (c) and (d). This is further illustrated in Figure 5.14, where the relative change in the estimated ultimate shrinkage parameter, α, for each growth curve model relative to Model 10 is plotted as a function of the time of the *last* shrinkage measurement for each shrinkage profile. The graphs show that for longer-term data, the estimates of α were much more similar than for profiles where data collection was terminated earlier, which included the cases where the asymptote (ultimate shrinkage) had not yet been reached.

Table 5.9 The performance of Models 10, 12, 13 and 15 with respect to convergence and the relative standard errors of the parameter estimates

Criterion	Model 10	Model 12	Model 13	Model 15
Number of shrinkage profiles for which the model fitting algorithm did not converge (or a singular Hessian matrix was obtained)	17	17	20	19
Number of shrinkage profiles which had one or more parameter estimates with relative standard errors exceeding 30%	58	72	141	118

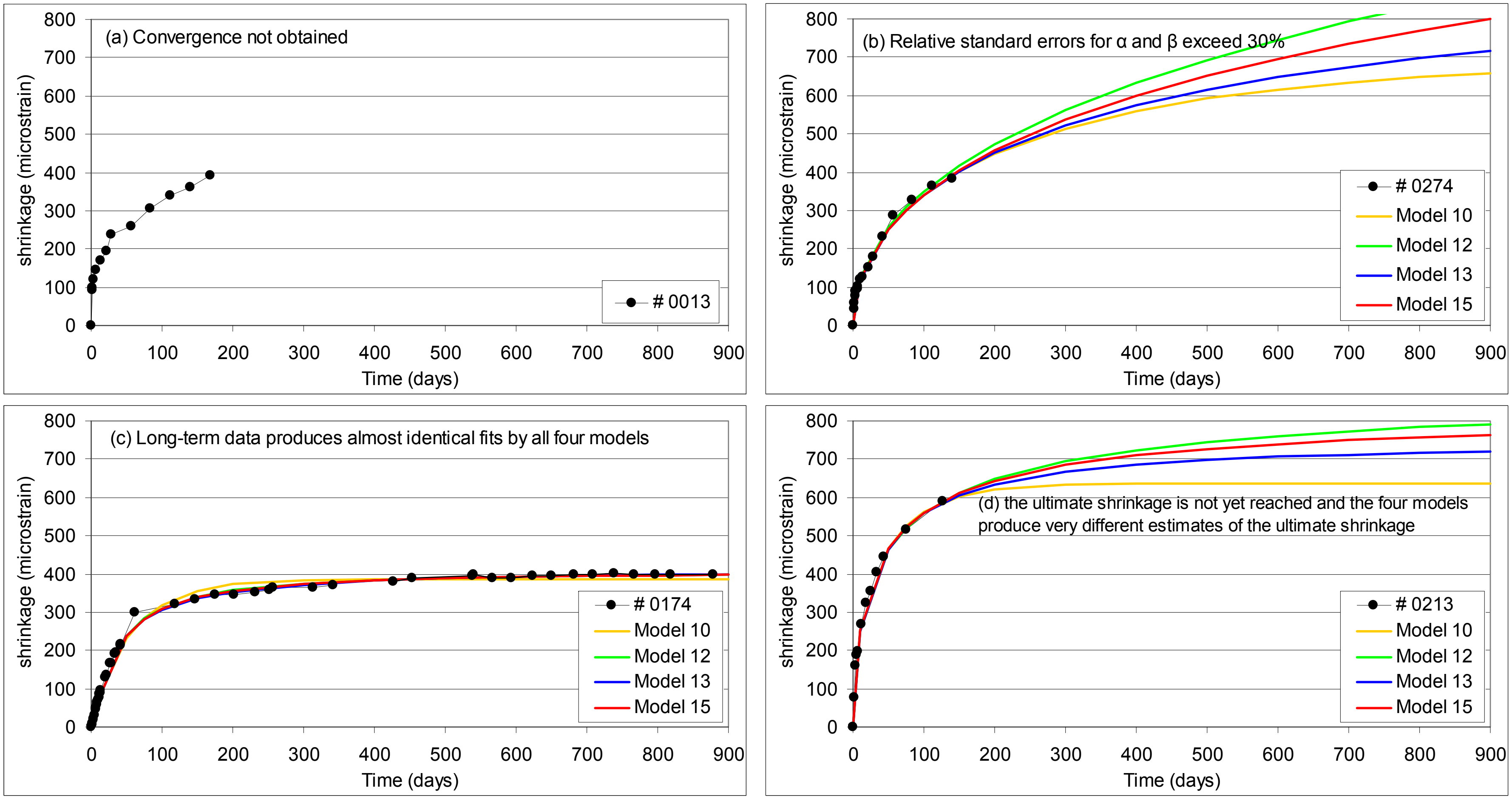


Figure 5.13 Illustration of the fit of the remaining four growth curve models to different types of shrinkage profiles

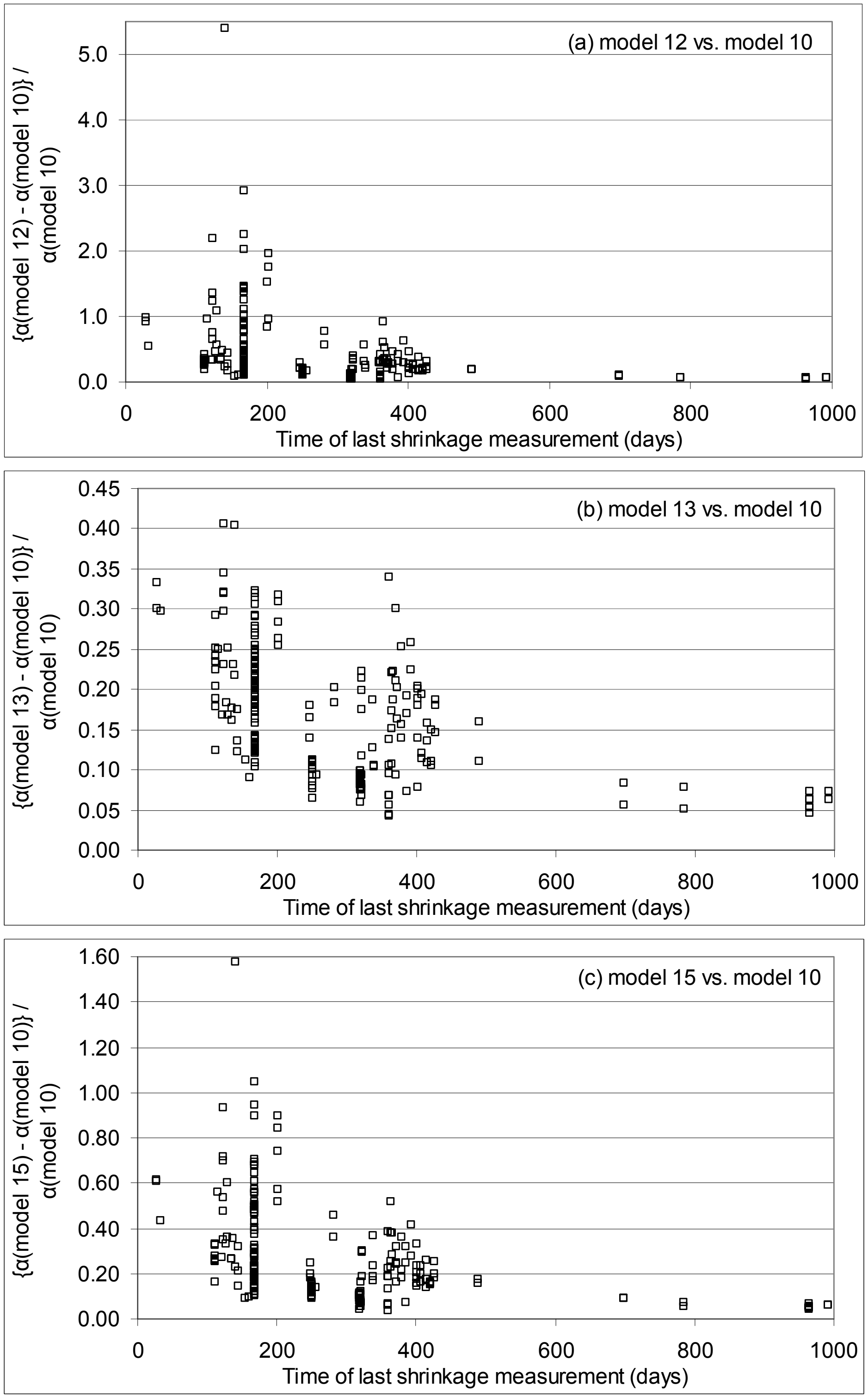


Figure 5.14 The relative change in the estimated value of the ultimate shrinkage parameter, α, as a function of the time of the last shrinkage measurement for a given shrinkage profile

Based on these observations, it was decided that **Model 10 was the most appropriate model to use for the shrinkage-time profiles**. As described in Table 5.1, Model 10 is given by

$$Shrinkage(t - t_0) = \alpha\left[1 - e^{-\beta(t-t_0)}\right]^{\gamma}, \quad \alpha, \beta, \gamma > 0$$

where α represents the ultimate shrinkage, β the growth rate and γ the shape factor. This growth curve model is not used in any of the existing models for concrete shrinkage discussed in Chapter 2.

The same choice of best growth curve model could be made after application of the two MCDA techniques used previously. In this case, the data was taken from the five goodness-of-fit measures for the four remaining growth curve models (Models 10, 12, 13 and 15) fitted to the 206 shrinkage profiles for which a fit for all four models could be obtained. The goodness-of-fit statistics are shown in Table 5.10. The 95% confidence intervals for the RMS include negative values as the distribution of RMS values for each model is skewed towards larger values.

Table 5.10 Goodness-of-fit scores for growth curve Models 10, 12, 13 and 15

Model	Goodness-of-fit score (95% confidence interval)				
	AIC	BIC	Pseudo-R^2	CCC	RMS
Model 10	10.97 (± 2.12)	83.11 (± 85.57)	0.9917 (± 0.0138)	0.9958 (± 0.0069)	13.67 (± 13.28)
Model 12	11.00 (± 2.23)	81.94 (± 69.54)	0.9906 (± 0.0183)	0.9953 (± 0.0092)	14.22 (± 15.76)
Model 13	11.00 (± 2.19)	81.90 (± 69.10)	0.9908 (± 0.0176)	0.9954 (± 0.0089)	14.14 (± 15.33)
Model 15	11.01 (± 2.21)	81.93 (± 69.30)	0.9907 (± 0.0181)	0.9953 (± 0.0091)	14.20 (± 15.60)

For each goodness-of-fit statistic, the models with the first, second and third best scores are highlighted in dark orange, light orange and yellow, respectively.

The Value Function and outranking methods were applied as described previously. The only difference was that, in the outranking method, the veto threshold for the discordance index was defined as the average standard deviation for each criterion (across the four models) since use of the standard error (now very small due to the large number of cases) led to a conclusion of incomparability in almost every pairwise comparison.

The standardised partial value functions as well as the overall value functions for the four growth curve models are given in Table 5.11 and the value functions are plotted in Figure 5.15. Model 10 clearly outperformed the other models when assessed by the Value Function method.

Table 5.11 Partial and overall value functions for growth curve Models 10, 12, 13 and 15

Model	Standardised partial value functions					Value function
	AIC	BIC	Pseudo-R^2	CCC	RMS	
Model 10	20	0	20	20	20	80
Model 12	1	19	0	1	0	22
Model 13	2	20	3	4	3	32
Model 15	0	19	1	0	1	21

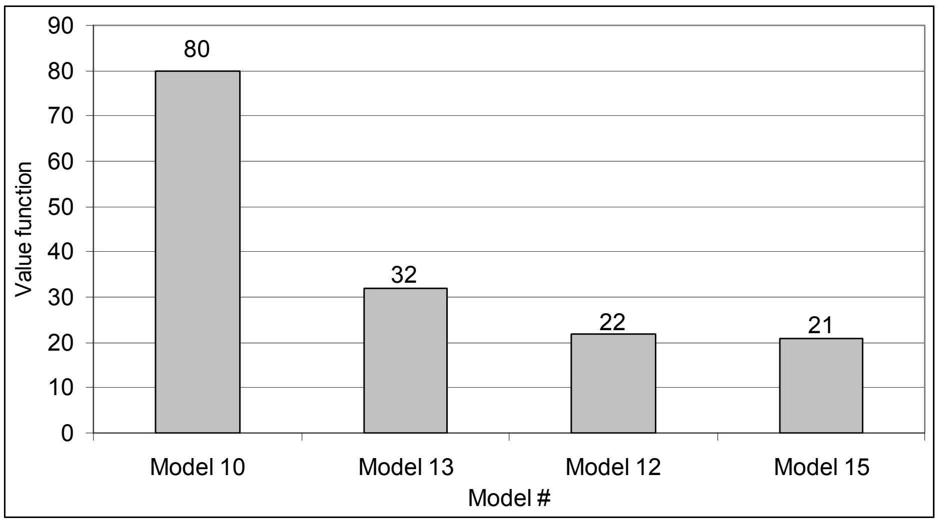


Figure 5.15 Comparison of the value functions for the four growth curve models

The tabulated concordance and discordance indices and the outcomes of the pairwise outranking comparisons are presented in Tables 5.12 and 5.13 respectively. The outranking relationship between the four growth curve models is depicted in Figure 5.16. Once again, Model 10 outranked the other models. The outranking method gave a similar relative order of the four growth curve models as the Value Function method, except that Models 12 and 15, which were narrowly separated by the Value Function method, were classed as incomparable by the outranking method. The ordering of the four models was similar to that found for the evaluation with the ten “key” shrinkage profiles, except that the ranking of Models 12 and 15 was reversed (or classed as incomparable in the case of the outranking method).

Table 5.12 Concordance and discordance indices for the ELECTRE I outranking method for growth curve Models 10, 12, 13 and 15

Concordance Index (a=column, b=row)				
Model	**10**	**12**	**13**	**15**
10		0.2	0.2	0.2
12	0.8		1.0	0.6
13	0.8	0.0		0.2
15	0.8	0.4	0.8	

Discordance Index (a=column, b=row)				
Model	**10**	**12**	**13**	**15**
10		0	0	0
12	0		0	0
13	0	0		0
15	0	0	0	

Table 5.13 Outranking relation for the pairs of alternatives (growth curve models) according to the ELECTRE I method for growth curve Models 10, 12, 13 and 15

Outranking relation (a=column, b=row)				
Model	**10**	**12**	**13**	**15**
10				
12	10 > 12			
13	10 > 13	13 > 12		
15	10 > 15	incomparable	13 > 15	

The symbol ">" indicates outranking.

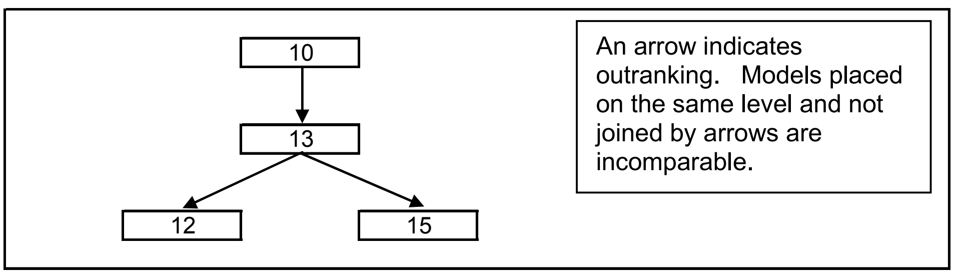
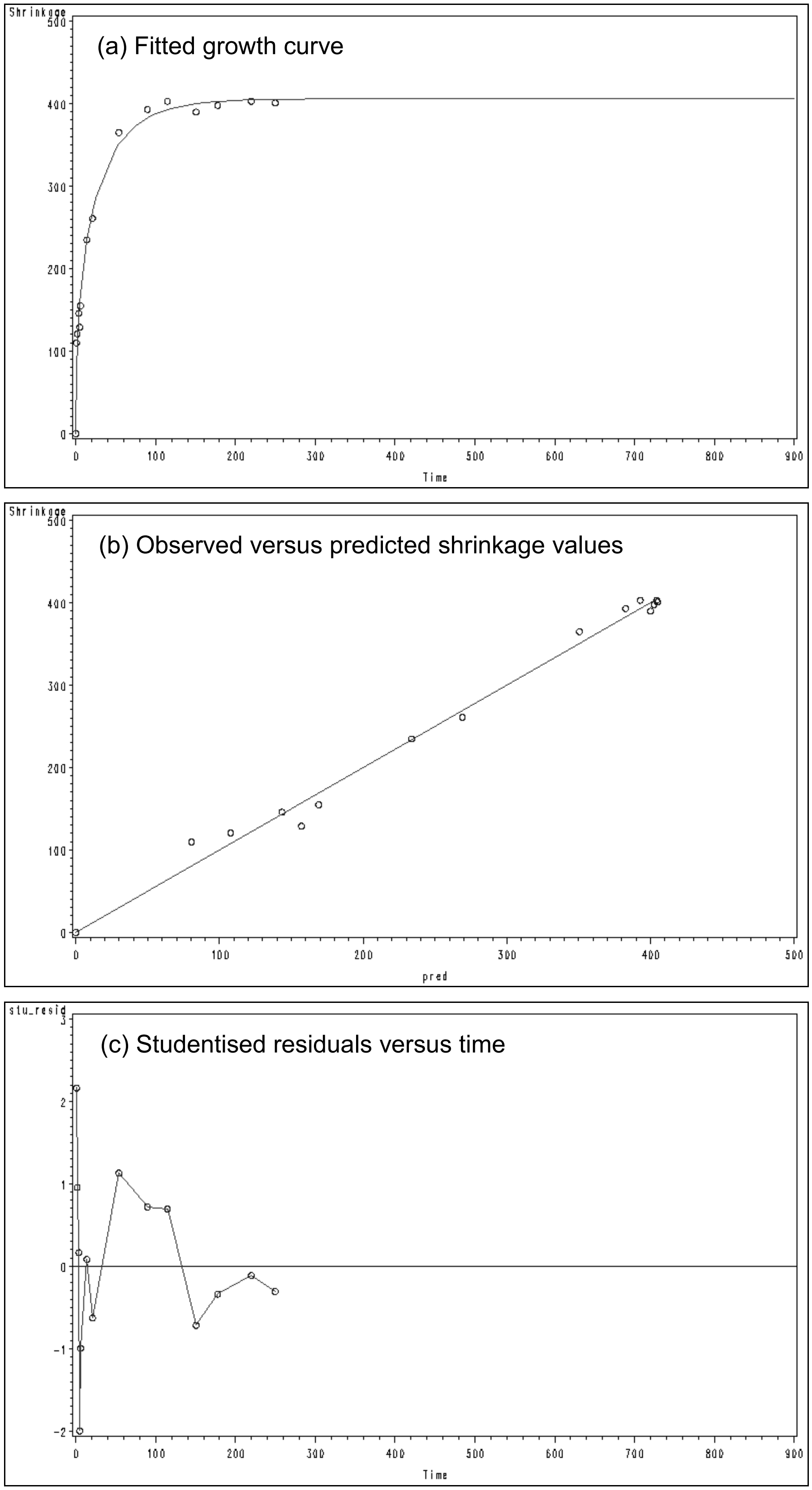


Figure 5.16 The outranking relationship for the four growth curve models according to the ELECTRE I method

Finally, in order to complete the discussion on growth curve model selection, it was noted that the graphical diagnostics for the chosen model were acceptable. The graphical diagnostics (as discussed earlier) for a shrinkage profile which was judged to have near-average values of the five goodness-of-fit measures are presented in Figure 5.17. **Note that the residuals were not appreciably correlated with time or with each other and thus that the assumption of an uncorrelated intra-experiment error structure, made in section 5.3.1, was justified.**



(a) Fitted growth curve
Shrinkage
Time
(b) Observed versus predicted shrinkage values
Shrinkage
pred
(c) Studentised residuals versus time
stu_resid
Time

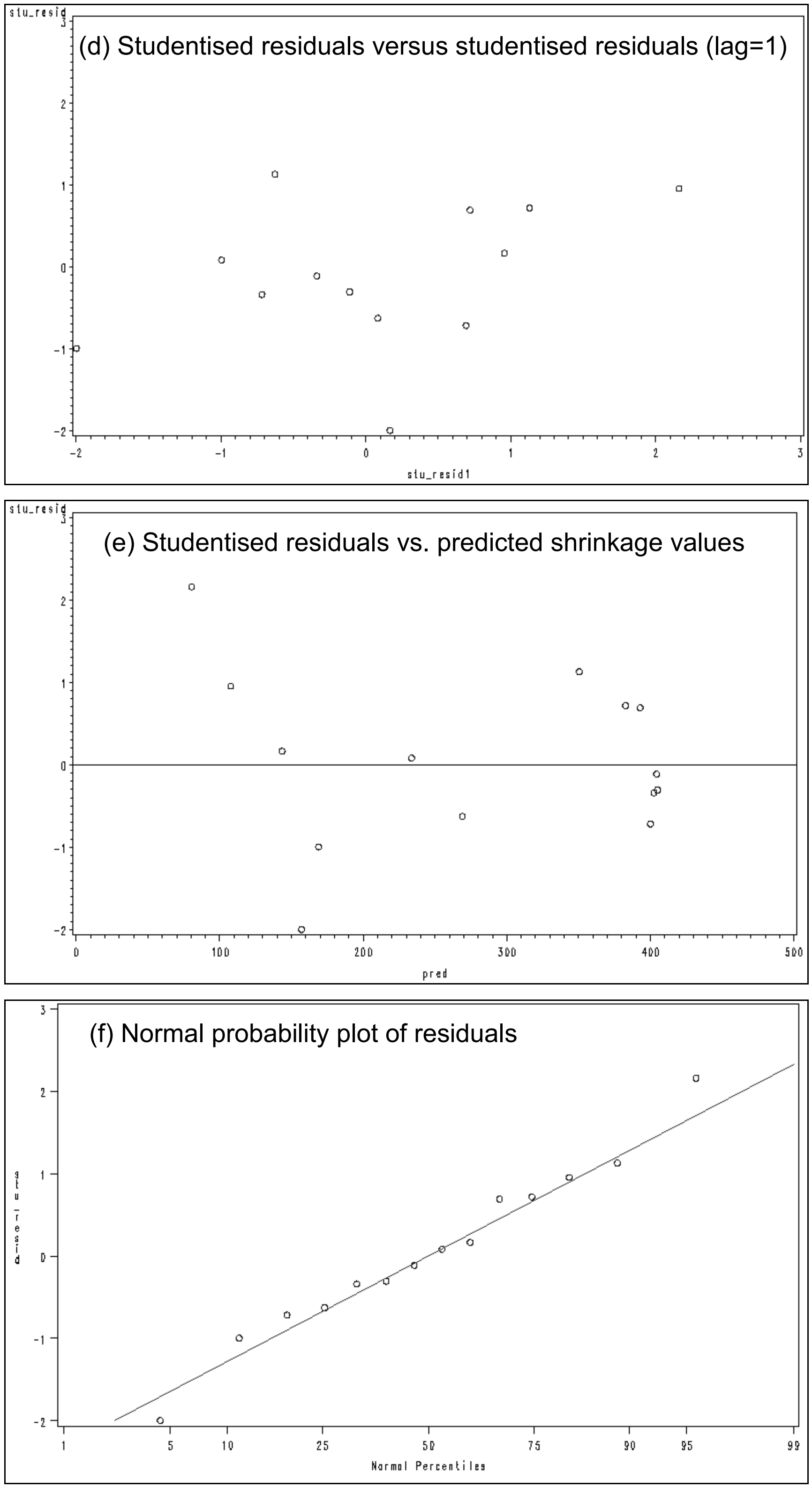


Figure 5.17 Graphical diagnostics for the fit of Model 10 to one shrinkage profile (# 0202)

In this chapter, the first step in the hierarchical non-linear modelling process was completed: a suitable growth curve model for modelling the shrinkage-time relationship for each experiment was identified. The heterogeneous variance of the shrinkage response variable was modelled in order to define a weighting scheme for the shrinkage-time data. However, it was found that such a weighting scheme made no appreciable difference to the model fit and thus it was decided not to weight the data but to fit the non-linear growth curve models by OLS. Candidate growth curve models were fitted to key shrinkage profiles and MCDA was used to select a group of models with the best fit to the key shrinkage profiles. This group of models was then fitted to all the shrinkage profiles in order to identify the best growth curve model, chosen on the basis of practical considerations as well as the application of MCDA.

In the next chapter, the nonlinear regression coefficients which characterise the shrinkage profiles (determined by fitting the chosen growth curve model to each shrinkage profile), will now be used as dependent variables in a multiple regression model to form a link between the covariates and the behaviour of the shrinkage profiles.

6 Modelling the parameters of the growth curve model in terms of the covariates by means of multivariate regression

The second stage in the hierarchical non-linear modelling process is presented in this chapter. The three parameters of the growth curve (α, β and γ) are modelled in terms of the covariates (which describe the concrete raw materials, the concrete composition and the shrinkage testing conditions) by multivariate multiple regression. Firstly, preparation of both the dependent variables and the covariates for regression analysis is outlined. Then, the grouping of the large number of covariates into meaningful groups for variable selection during the regression step, in order to avoid multicollinearity, is described. Thereafter, the results of the multivariate multiple regressions are presented and different weighting schemes for the cases (experiments) in the multivariate multiple regression analysis are explored. MCDA is then used to identify the best regression model, using appropriate goodness-of-fit measures. The best model is described in terms of its range of applicability and significant terms and a 95% confidence interval is derived for the predicted shrinkage profiles. The model is interpreted in terms of what is already known about the factors affecting concrete shrinkage. Finally, the findings of a sensitivity analysis to investigate the effect of changes in the input parameters on the predicted growth curve parameters are presented.

6.1 The second stage of the nonlinear hierarchical model

The dependent variables in the second stage of the hierarchical model are the estimates of the three parameters (α, β and γ) of the growth curve model

$$Shrinkage(t-t_0) = \alpha\left[1-e^{-\beta(t-t_0)}\right]^{\gamma}, \quad \alpha,\beta,\gamma > 0$$

fitted to each shrinkage profile. The results of the fitting of the growth curve model to the individual shrinkage profiles (as described in Chapter 5) showed that the estimates of the three parameters were generally highly correlated, particularly the parameters β and γ. The average correlations between the parameters for the chosen growth curve model over the ten "key" shrinkage profiles is given in Table 6.1.

Table 6.1 Correlations between the three growth curve parameters, averaged over the ten "key" shrinkage profiles

Parameters	(α,β)	(α,γ)	(β,γ)
Average correlation between parameters (range in brackets)	-0.72 (-0.96 to -0.36)	-0.50 (-0.76 to -0.20)	0.88 (0.86 to 0.93)

It was not possible to reduce this correlation by reducing the number of parameters in the growth curve model. It has already been shown that a three-parameter model provided the best fit to the data and that all the three-parameter models which were evaluated had high levels of correlation between their parameters. As a result it was decided to **model the three parameters together by means of multivariate multiple regression** rather than modelling them individually by simple multiple regression.

Multivariate multiple regression[64,65,66] is an extension of simple multiple regression and models the relationship between dependent variables Y_i, $i = 1, 2, \ldots, m$ (in the case of this study, m = 3) and a single set of covariates z_j, $j = 1, 2, \ldots, r$, while accounting for the correlation among the dependent variables. Each dependent variable is assumed to follow its own regression model:

$$\begin{aligned}
Y_1 &= \varphi_{01} + \varphi_{11} z_1 + \ldots + \varphi_{r1} z_r + e_1 \\
Y_2 &= \varphi_{02} + \varphi_{12} z_1 + \ldots + \varphi_{r2} z_r + e_2 \\
&\vdots \\
Y_m &= \varphi_{0m} + \varphi_{1m} z_1 + \ldots + \varphi_{rm} z_r + e_m
\end{aligned}$$

The error term $e = [e_1, e_2, \ldots, e_m]'$ has $E(e) = 0$ and $Var(e) = \Sigma$. Thus, allowance is made for the error terms associated with the different dependent variables, within the same experiment, to be correlated and have different variances.

The overall model may be written as

$$\underset{(n \times m)}{\mathbf{Y}} = \underset{(n \times (r+1))}{\mathbf{Z}} \underset{((r+1) \times m)}{\varphi} + \underset{(n \times m)}{\mathbf{e}}$$

where, k = 1, 2, …, n are the individual cases (experiments) and

$$\mathbf{Y} = \begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1m} \\ Y_{21} & Y_{22} & \cdots & Y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{nm} \end{bmatrix}$$

$$\mathbf{Z} = \begin{bmatrix} z_{10} & z_{11} & \cdots & z_{1r} \\ z_{20} & z_{21} & \cdots & z_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n0} & z_{n1} & \cdots & z_{nr} \end{bmatrix} \qquad where \begin{bmatrix} z_{10} \\ z_{20} \\ \vdots \\ z_{n0} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\varphi = \begin{bmatrix} \varphi_{01} & \varphi_{02} & \cdots & \varphi_{0m} \\ \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{r1} & \varphi_{r2} & \cdots & \varphi_{rm} \end{bmatrix}$$

$$\mathbf{e} = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1m} \\ e_{21} & e_{22} & \cdots & e_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ e_{n1} & e_{n2} & \cdots & e_{nm} \end{bmatrix}$$

The least squares estimates of the parameters of the regression are given by

$$\hat{\varphi} = (\mathbf{Z'Z})^{-1}\mathbf{Z'Y}$$

from which the predicted values, $\hat{\mathbf{Y}} = \mathbf{Z}\hat{\varphi} = \mathbf{Z(Z'Z)}^{-1}\mathbf{Z'Y}$, the matrix of residuals $\hat{\mathbf{e}} = \mathbf{Y} - \hat{\mathbf{Y}} = \left[\mathbf{1} - \mathbf{Z(Z'Z)}^{-1}\mathbf{Z'}\right]\mathbf{Y}$ and the estimate of the variance $\hat{\mathbf{\Sigma}} = \dfrac{\hat{\mathbf{e}}'\hat{\mathbf{e}}}{n-r-1} = \dfrac{(\mathbf{Y} - \mathbf{Z}\hat{\varphi})'(\mathbf{Y} - \mathbf{Z}\hat{\varphi})}{n-r-1}$ may be determined.

The ANOVA table can be formed as for simple multiple regression, but in this case the Sums of Squares are matrices:

Source	Sums of Squares	Degrees of freedom	Mean Squares
Regression	$\mathbf{R} = (\hat{\mathbf{Y}} - \mathbf{1}\bar{y})'(\hat{\mathbf{Y}} - \mathbf{1}\bar{y})$	*r*	**R**/*r*
Residual	$\mathbf{W} = \hat{\mathbf{e}}'\hat{\mathbf{e}}$	*n-r-1*	**W**/*(n-r-1)*
Total	$\mathbf{T} = (\mathbf{Y} - \mathbf{1}\bar{y})'(\mathbf{Y} - \mathbf{1}\bar{y})$	*n-1*	

The usual F-test for the significance of the regression model cannot be performed, since the Sums of Squares are in matrix form. A number of alternative tests are used, based on the eigenvalues of the RW^{-1} matrix including Wilk's lambda, Roy's greatest root, Lawley-Hotelling's trace and Pillai's trace.

The confidence interval for a single parameter, φ_{ij}, is given by

$$\hat{\varphi}_{ij} \pm t_{\alpha/2,n-r-1} s_j \sqrt{g_{ii}}$$

where s_j^2 is the jth diagonal element of $\hat{\Sigma}$ and g_{ii} is the ith diagonal element of $(Z'Z)^{-1}$. Similarly, the hypothesis H_0: $\varphi_{ij} = 0$ can be tested by

$$t = \frac{\hat{\varphi}_{ij}}{s_j \sqrt{g_{ii}}}.$$

The null hypothesis is rejected if $|t| \geq t_{\alpha/2,n-r-1}$.

A prediction, given the covariate set z_0', is given by $\hat{Y}_0' = z_0' \hat{\varphi}$. The standard error for each of the predicted m elements of $\hat{Y}_0'$ is given by $s_j \sqrt{1 + z_0'(Z'Z)^{-1} z_0}$ where j = 1, 2, …, m and s_j^2 is the jth diagonal element of $\hat{\Sigma}$.

6.2 Selection of experiments (cases) for the multivariate multiple regression model

As described in Chapter 5, some shrinkage profiles (experiments) were already eliminated for the following reasons:

- The shrinkage profiles consisted of fewer than four data points, which meant that a 3-parameter growth curve model could not be fitted to the data.
- The fit of the growth curve model did not converge.

In order to derive a reliable second-stage model, it was also necessary to exclude those profiles (experiments) for which data were insufficient to allow a realistic prediction of the ultimate shrinkage, i.e. for which either

- the shrinkage measurements ceased before 60 days, which is the minimum drying time specified for the natural shrinkage test method[6] or
- the shrinkage predictions continued to increase substantially at very long drying times. Such profiles were identified by comparing the estimated value of α (the ultimate shrinkage) with the value of the last shrinkage measurement. If these differed by more than 15% (a limit set by inspection), the experiment was excluded from the regression analysis.

The number of experiments excluded from each of the four causes above is shown in Table 6.2. The total number of experiments retained was 174, representing 71% of the original data set of 245 experiments.

Table 6.2 Causes of exclusion of experiments from further analysis together with the number of experiments affected

Cause of exclusion	Number of experiments excluded
Fewer than four data points in shrinkage profile	18
For remaining profiles, fitting of growth curve model did not converge	17
For remaining profiles, shrinkage measurements ceased before 60 days	7
For remaining profiles, estimate of ultimate shrinkage (α) differs from last shrinkage measurement by more than 15%	29
Total number of experiments excluded	**71**
Number of experiments retained = 174	

6.3 Preparation of the dependent variables

Firstly, the univariate characteristics of the dependent variables were explored. Key results are given in Table 6.3, while the distribution of the dependent variables and their normal probability plots are shown in Figures 6.1 and 6.2 respectively.

Table 6.3 Univariate statistics for the three dependent variables

Statistic	Dependent Variable				
	α	**β**	**γ**	**ln(β)**	**ln(γ)**
Number of observations (n)	174				
Mean (minimum, maximum values)	453 (169, 824)	0.018 (0.003, 0.048)	0.62 (0.27, 1.50)	-4.15 (-5.90, -3.03)	-0.48 (-1.30, 0.41)
Median	465	0.016	0.62	-4.12	-0.48
Standard deviation	125	0.010	0.20	0.58	0.29
z(skewness) = coefficient of skewness / √ (6 / n)	-0.10	3.5 *	5.6 *	-1.72	0.39
z(kurtosis) = coefficient of kurtosis / √ (24 / n)	-0.83	-0.51	5.5 *	-1.76	0.02
p-values for Kolmogorov-Smirnov and Lilliefors tests for non-normality	> 0.20 > 0.20	< 0.05 * < 0.01 *	> 0.20 < 0.05 *	< 0.05 * < 0.01 *	> 0.20 > 0.20
Outliers (beyond 3 standard deviations from the mean)	none	none	2 on high side	none	1 on high side

* Statistic significant at the 0.05 level.

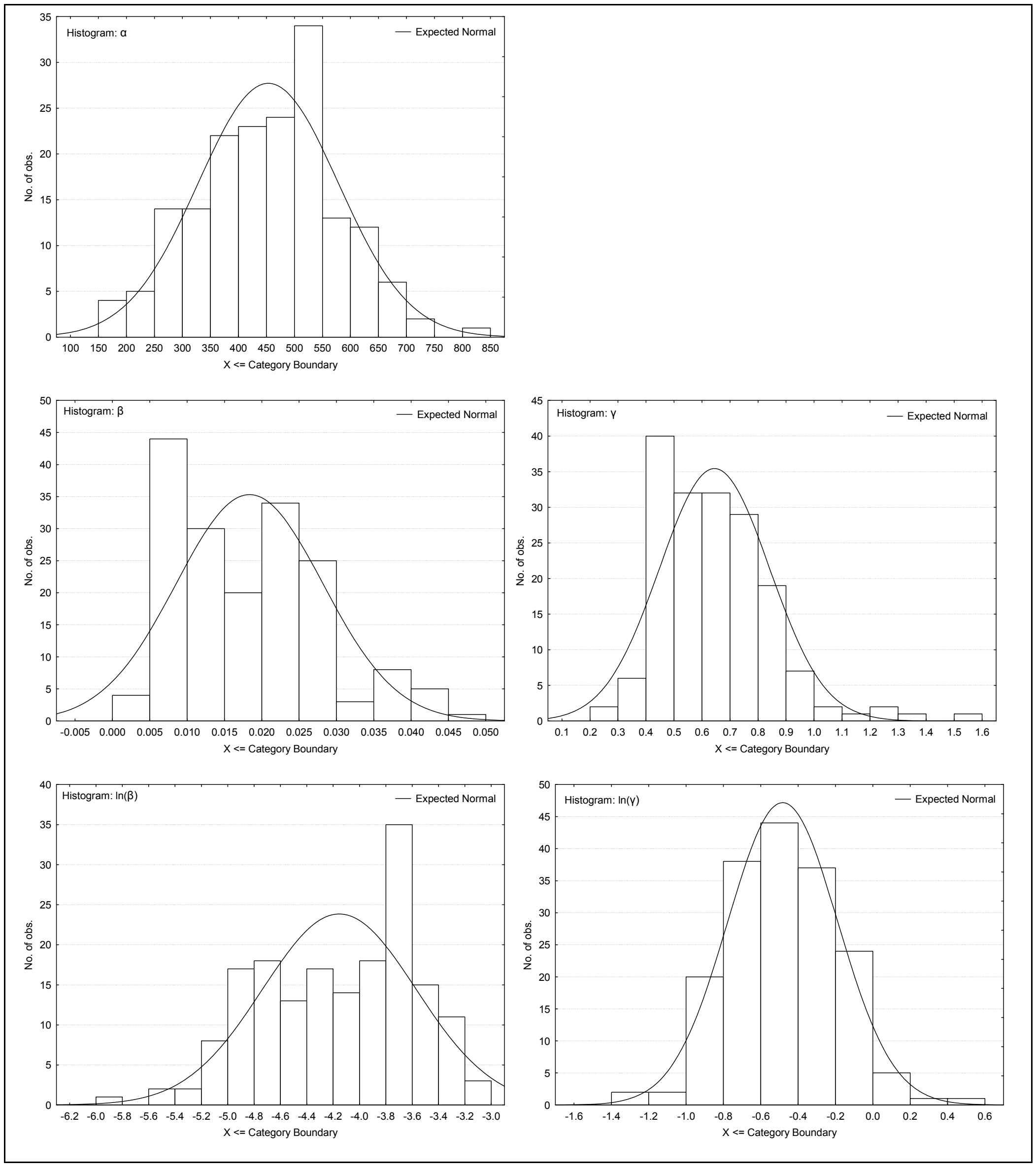


Figure 6.1 The distributions of the dependent variables α, β and γ, and the transformed variables, ln(β) and ln(γ)

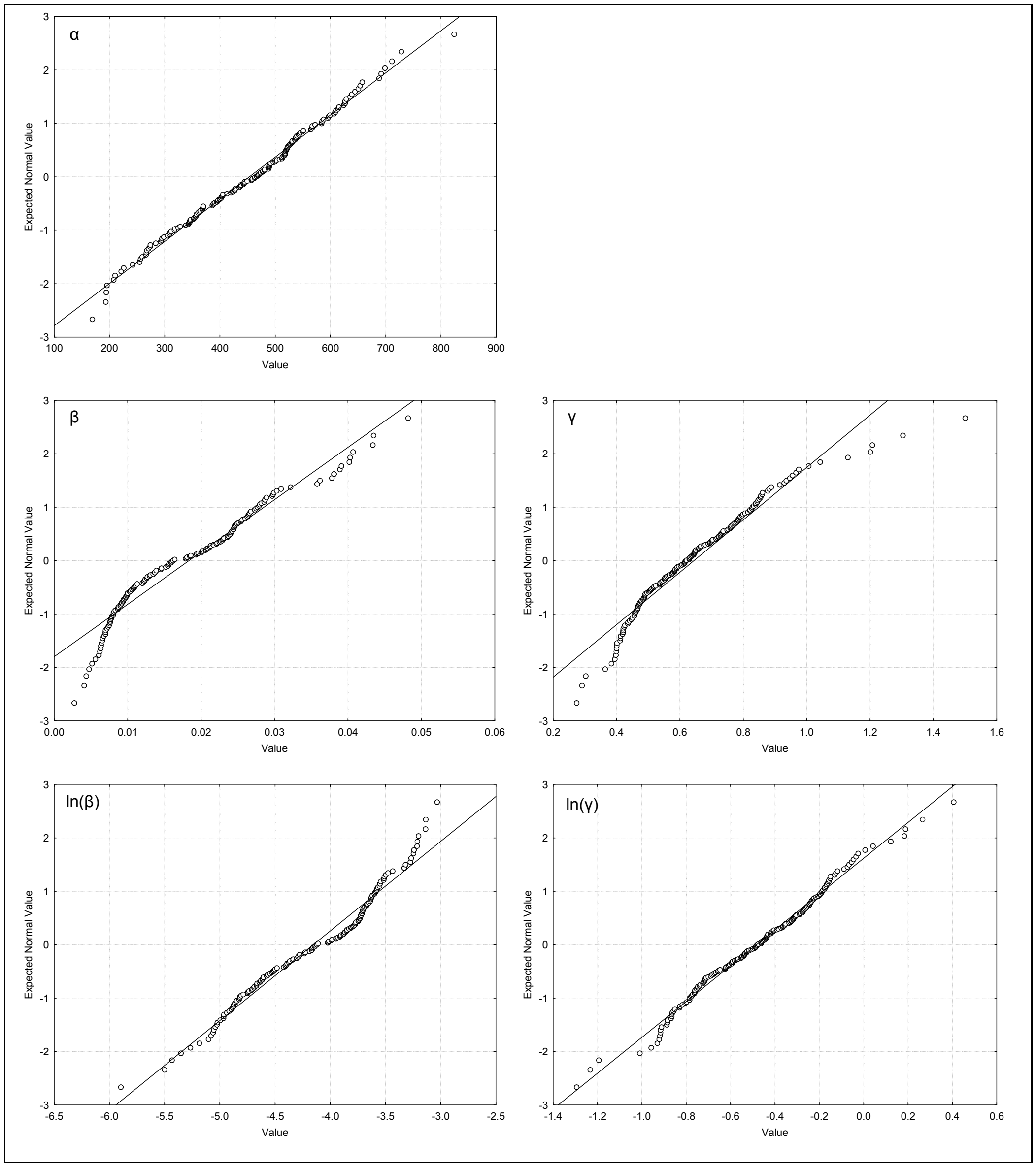


Figure 6.2 The normal probability plots for the dependent variables α, β and γ, and the transformed variables, ln(β) and ln(γ)

While the distribution of the dependent variable α was close to normal, the distribution of β was found to be significantly positively skewed, while that of γ was found to be both positively skewed and leptokurtic. The application of a natural logarithm transformation to both these variables reduced these problems to a large extent. It was thus decided that **the forms of the dependent variables to be used in the multivariate multiple regression would be α, ln(β) and ln(γ)**. As will be shown later, the distribution of the residuals using the transformed variables confirmed that these transformations were indeed appropriate.

The three dependent variables were found to be moderately correlated, as shown in Table 6.4.

Table 6.4 Correlation coefficients for the three dependent variables α, ln(β) and ln(γ)

Variable	**α**	**ln(β)**	**ln(γ)**
α	1.00		
ln(β)	-0.40	1.00	
ln(γ)	-0.40	0.55	1.00

6.4 Preparation of the covariates

The **univariate** characteristics of the covariates, for the 174 experiments now contained in the reduced data set to be used for the second stage of the development of the hierarchical model, are summarised in Table 6.5. Explanations of these variables were given in Tables 3.1-3.3.

At this stage, a number of covariates were eliminated from consideration for the multivariate multiple regression model:

- The variables Data Set and Test Year were excluded since these would be of no use in a predictive model.
- Covariates with more than 30% missing data (see Table 6.5) were excluded. If any one such covariate were included in the regression analysis, casewise exclusion of missing data would reduce the number of experiments (cases) in the model by more than 30%. The following covariates were excluded on this basis:
    - Cement Producer, Cement Milling / Blending Unit and Cement Production Unit: these are of secondary importance to the Cement Type covariate in describing the cement raw material and so their exclusion was not thought to have a detrimental effect on the quality of the covariates available for the regression model.

- Sixty-nine percent of the Cement Strength Class data was associated with data gathered prior to 1998 when such a classification had not yet been introduced. At the time of carrying out this research, it was thought that no inference could be made about the strength class of the cements used in these experiments, and that thus all the data gathered prior to 1998 effectively contained missing data for this covariate. However, it has subsequently been understood that the strength class of OPC produced before 1996 was between 32,5 and 42,5 and that a value of, say, 37,5 could have been used for such cements in the data set, while pre-1996 RHC corresponded to a strength class of 42,5.
- A similar situation existed for the variable Cement Early Strength Gain. In this case, the data prior to 1998 was assigned to the class N (normal), since at the time of carrying out this research it was thought that rapid-hardening cement (class R) had not yet come into mainstream commercial use at that time, and the definition of class R is encompassed by the definition of class N(41). However, it has subsequently been understood that rapid-hardening cement was in fact available before 1998. Since this variable was excluded from later analyses, this finding does not change the validity of the results from this study.
- Sand property variables Loose Bulk Density, Consolidated Bulk Density, Fines and Elastic Modulus: exclusion of these variables leaves only the Relative Density and Fineness Modulus as numeric descriptors of the sand properties. However, Sand Type remains the main descriptive variable in this case, and so again the exclusion of these covariates does not negatively affect the quality of the data set.

Table 6.5 Univariate characteristics of the covariates

Variable (variables excluded from regression analysis are highlighted in grey)	**Percentage of missing data**	**Variable type**	**Mean (range in brackets) / Number of categories**	**Comment**
Data Set	0	Nominal	25 categories	Data fairly well spread across data sets.
Test Year	0	Ordinal	14 categories (1979 – 2009)	71% of data gathered after 2000.

Cement covariates

Variable	**Percentage of missing data**	**Variable type**	**Mean (range in brackets) / Number of categories**	**Comment**
Cement Type	0	Nominal	12 categories	69% of data from 2 categories: 45% CEM I, 24% CEM III.
Cement Strength Class	7	Nominal	4 categories	69% of data from 1 category: strength class not in use.
Cement Early Strength Gain Class	25	Nominal	3 categories	52% of data from 1 category: strength class not in use.
Cement Producer	33	Nominal	4 categories	
Cement Milling / Blending Unit	90	Nominal	3 categories	
Cement Production Unit	50	Nominal	11 categories	

Table 6.5 Univariate characteristics of the covariates (continued)

Stone covariates

Variable	**Percentage of missing data**	**Variable type**	**Mean (range in brackets) / Number of categories**	**Comment**
Stone Type	1	Nominal	11 categories	69% of data from 3 categories: 28% andesite, 22% greywacke, 19% dolerite.
Stone Source	25	Nominal	16 categories	
Stone Size	0	Continuous	19.3 (13.2-32.5) mm	Effectively consists of 3 levels. 97% of data from 1 level (19 mm).
Stone Relative Density	0	Continuous	2.8 (2.6-3.0)	
Stone Loose Bulk Density	6	Continuous	1434 (1285-1578) kg/m^3	
Stone Consolidated Bulk Density	8	Continuous	1603 (1435-1760) kg/m^3	
Stone Water Absorption	2	Continuous	0.41 (0.21-0.90) %	
Stone Coefficient of Thermal Expansion	3	Continuous	9.0 (6.5-11.9) x10^{-6}	
Stone Unconfined Compressive Strength	10	Continuous	347 (135-527) MPa	
Stone 10% FACT	2	Continuous	312 (155-448) kN	
Stone Elastic Modulus	1	Continuous	85 (61-114) GPa	
Stone Flakiness Index	4	Continuous	21 (13-23)	

Table 6.5 Univariate characteristics of the covariates (continued)

Sand covariates

Variable	**Percentage of missing data**	**Variable type**	**Mean (range in brackets) / Number of categories**	**Comment**
Sand Type	1	Nominal	16 categories	50% of data from 3 categories: 18% natural sand, 16% dolomite, 16% Klipheuwel pit sand
Sand Source	10	Nominal	16 categories	54% of data from 3 categories: 19% Leazonia, 18% Klipheuwel, 17% Olifantsfontein (in line with Sand Type)
Sand Relative density	7	Continuous	2.7 (2.4-2.9)	
Sand Loose Bulk Density	64	Continuous	1756 (1510-2020) kg/m^3	
Sand Consolidated Bulk Density	71	Continuous	1986 (1680-2145) kg/m^3	
Sand Fineness Modulus	10	Continuous	2.9 (2.2-3.5)	
Sand Fines	32	Continuous	7.7 (0.1-16) %	
Sand Elastic Modulus	63	Continuous	95 (67-114) GPa	

Admixture covariate

Variable	**Percentage of missing data**	**Variable type**	**Mean (range in brackets) / Number of categories**	**Comment**
Admix Type	0	Nominal	8 categories	74% of data from 1 category (no admixture used).

Table 6.5 Univariate characteristics of the covariates (continued)

Recipe covariates: content values

Variable	**Percentage of missing data**	**Variable type**	**Mean (range in brackets) / Number of categories**	**Comment**
Cement Content	0	Continuous	293 (113-536) kg/m^3	
CSF Content	0	Continuous	0.3 (0-20) kg/m^3	Effectively consists of 4 levels. 98% of data from 1 level (0 kg/m^3).
FA Content	0	Continuous	11 (0-150) kg/m^3	88% of data has the value 0 kg/m^3.
GGBS Content	0	Continuous	44 (0-262) kg/m^3	72% of data has the value 0 kg/m^3.
GGCS Content	0	Continuous	10 (0-244) kg/m^3	93% of data has the value 0 kg/m^3.
GGFS Content	0	Continuous	6 (0-244) kg/m^3	95% of data has the value 0 kg/m^3.
Limestone Content	0	Continuous	1 (0-68) kg/m^3	Effectively consists of 4 levels. 98% of data from 1 level (0 kg/m^3).
Water Content	0	Continuous	191 (160-225) kg/m^3	
Stone Content	0	Continuous	1104 (900-1400) kg/m^3	
Sand Content	0	Continuous	794 (511-1248) kg/m^3	
Admix Content	0	Continuous	0.3 (0-3.3) kg/m^3	

Table 6.5 Univariate characteristics of the covariates (continued)

Recipe covariates: content fractions or ratios

Variable	**Percentage of missing data**	**Variable type**	**Mean (range in brackets) / Number of categories**	**Comment**
Paste Volume Fraction	0	Continuous	0.31 (0.24-0.40)	
CSF Fraction	0	Continuous	0.0 (0.0-0.07)	Effectively consists of 4 levels. 98% of data from 1 level (0).
FA Fraction	0	Continuous	0.03 (0.0-0.30)	88% of data has the value 0.
GGBS Fraction	0	Continuous	0.12 (0.0-0.50)	72% of data has the value 0.
GGCS Fraction	0	Continuous	0.03 (0.0-0.50)	93% of data has the value 0.
GGFS Fraction	0	Continuous	0.02 (0.0-0.50)	95% of data has the value 0.
Limestone Fraction	0	Continuous	0.0 (0.0-0.13)	Effectively consists of 4 levels. 98% of data from 1 level (0).
Water / binder ratio	0	Continuous	0.55 (0.38-0.83)	
Aggregate / binder ratio	0	Continuous	5.6 (3.2-8.7)	
Stone / sand ratio	0	Continuous	1.4 (0.8-2.5)	

Table 6.5 Univariate characteristics of the covariates (continued)

Shrinkage testing variables

Variable	Percentage of missing data	Variable type	Mean (range in brackets) / Number of categories	Comment
Laboratory	0	Nominal	2 categories	Data fairly evenly divided between laboratories: 53% from WITS, 47% from UCT.
Age at First Drying	0	Continuous	25 (7-49) days	
2 * Volume To Surface Area	0	Continuous	47 (33-150) mm	
Temperature	0	Continuous	22.8 (21.0-25.0)ºC	
Humidity	0	Continuous	59 (43-72)%	

Concrete property variables

Variable	Percentage of missing data	Variable type	Mean (range in brackets) / Number of categories	Comment
Slump	6	Continuous	68 (10-115) mm	
28-day Compressive Strength	18	Continuous	46 (21-75) MPa	
28-day Elastic Modulus	33	Continuous	34 (15-52) GPa	

- o 28-day Elastic Modulus of the concrete: this was potentially a useful variable in describing the shrinkage behaviour of the concrete, but given the high level of missing data, it had to be excluded from the present study. However, the variable was correlated with another concrete property variable, the 28-day compressive strength (correlation coefficient = 0.70) which was available to be used in the regression analyses.

It is recommended that, in future studies, information on these variables (particularly the 28-day elastic modulus of the concrete) be gathered where at all possible, and that future modelling studies re-evaluate the possible importance of these covariates once more data has become available.

In order to ensure that the distributions of the *continuous* covariates were reasonably symmetrical and well spread out, these covariates were assessed for any transformations which might be necessary by inspection of their frequency distributions and assessment of their skewness and kurtosis statistics. Appropriate transformations were assessed for those covariates which exhibited significant skewness and / or kurtosis, and for which the coefficient of variation was greater than 0.25[55], since all the covariates represented positive physical measurements. The results are shown in Table 6.6. On the basis of these results it was decided to apply a natural logarithm transformation to the variables Stone Water Absorption, Cement Content, and (2 * Volume To Surface Area) to remedy positive skewness in the distributions of these variables. Transformations to remedy platykurtosis in several variables, as well as negative skewness in the Slump variable, were unsuccessful. Transformations of the cement content (FA Content, etc.) and fraction (FA Fraction, etc.) variables which contained a high proportion of data with a single value (usually equal to zero) were not attempted.

Table 6.6 Statistics used to determine the necessity for transformations of the continuous covariates

Covariate	z(skewness)	z(kurtosis)	|mean/standard deviation| (values < 4 are highlighted)	Outliers beyond 4 standard deviations from the mean
	Values significant at the 5% significance level are highlighted			
Stone Size	31.21	96.83	9.22	yes
Stone Relative Density	-0.16	-4.42	22.78	
Stone Loose Bulk Density	1.28	-3.50	15.74	
Stone Consolidated Bulk Density	-1.00	-3.84	16.85	
Stone Water Absorption	7.20	12.46	3.26	
ln (Stone Water Absorption)	-1.63	3.08	3.07	
Stone Coeff of Thermal Expansion	2.41	-3.99	5.10	
Stone Unconfined Compression	1.43	-2.78	2.65	
1 / Stone Unconfined Compression	7.79	4.75	2.15	yes
Stone 10% FACT	1.64	-3.86	2.94	
1 / Stone 10% FACT	2.37	-1.58	2.93	yes
Stone Elastic Modulus	1.77	-2.68	6.39	
Stone Flakiness Index	-10.16	8.35	7.87	
Sand Relative Density	0.00	-1.78	20.26	
Sand Fineness Modulus	-2.64	-1.69	8.80	
Cement Content	2.66	-1.47	2.69	
ln (Cement Content)	-1.04	-1.92	14.63	
CSF Content	41.72	161.77	0.13	yes
FA Content	15.68	20.26	0.35	yes
GGBS Content	8.08	2.13	0.57	
GGCS Content	22.35	46.50	0.26	yes
GGFS Content	30.18	92.85	0.21	yes
Limestone Content	44.30	184.71	0.13	yes
Water Content	1.64	-0.06	14.54	
Sand Content	2.49	0.65	6.30	
Stone Content	5.56	7.09	13.51	
Admix Content	11.81	17.37	0.53	yes
PVF	3.08	-2.08	7.78	
CSF Fraction	42.55	170.81	0.13	yes
FA Fraction	13.80	13.10	0.36	
GGBS Fraction	6.73	-0.75	0.59	
GGCS Fraction	19.94	33.60	0.27	yes
GGFS Fraction	25.84	61.51	0.22	yes
Limestone Fraction	40.29	147.04	0.13	yes
Water/Binder Ratio	1.69	-1.73	4.72	
Aggregate/Binder Ratio	0.94	-2.21	3.55	
1 / (Aggregate/Binder Ratio)	2.99	-2.92	3.32	
Stone/Sand Ratio	3.91	2.27	4.98	
Age at First Drying	-5.33	1.18	2.77	
2* Volume To Surface Area	27.61	78.44	2.82	yes
ln (2* Volume To Surface Area)	16.09	38.84	16.67	yes
Temperature	0.76	-2.02	19.24	
Humidity	-2.58	-3.71	8.21	
Slump	-3.44	-0.05	3.43	
√ (Slump)	-5.79	2.72	6.13	

Covariate	z(skewness)	z(kurtosis)	|mean/standard deviation| (values < 4 are highlighted)	Outliers beyond 4 standard deviations from the mean
	Values significant at the 5% significance level are highlighted			
1 / Slump	24.15	83.48	1.73	yes
28-day Compressive Strength	1.07	-2.11	3.45	
1 / 28-day Compressive Strength	4.40	1.01	3.16	

Next, the **bivariate** characteristics of the covariates were considered. A summary of the conclusions of bivariate plots of the covariates versus each other and versus the three dependent variables is given in Appendix F, while a correlation matrix for all the continuous variables in the data set is given in Appendix G. Although correlation coefficients as small as ±0.15 were found to be statistically significant at the 5% significance level, absolute values of correlation coefficients greater than 0.70 were considered to be practically significant.

It was interesting to note that there was not one continuous covariate which exhibited a strong correlation with any of the three dependent variables: the strongest correlation observed was between ln(β) and Sand Fineness Modulus ($r = -0.44$). The possibility of nonlinear relationships between the covariates and the dependent variables was also considered. However, inspection of the bivariate plots, as well as plots of α (ultimate shrinkage) versus transformations of covariates as used in other models for shrinkage (outlined in Chapters 2 and 7), did not suggest any useful nonlinear relationships.

As was to be expected, the data set contained a number of highly correlated covariates, so it was ultimately necessary to specify *sets of covariates*, for evaluation by means of multivariate multiple regression, which adequately described the composition and properties of the concrete, while avoiding collinearity between the variables. As a first step, the bivariate plots and the correlation matrix were used to identify **highly correlated relationships between the covariates**:

- The variables Stone Source and Stone Type were highly associated (most stone types were obtained from specific sources or locations). Since Stone Source is a much less important variable than Stone Type, the former variable was excluded as a candidate covariate. The same finding was made for the variables Sand Source and Sand Type.

- The Stone Size and the ln(2 * Volume To Surface Area) variables were correlated (correlation coefficient, r = 0.74) due to a combination of experimental conditions. Since these two variables were confounded to some extent, 97% of the values for Stone Size were identical and since it is known that the volume to surface area of the specimen affects the rate of shrinkage, it was decided to omit Stone Size from the analysis in favour of the ln(2 * Volume to Surface Area) variable.
- Some combinations of some the numeric variables describing the stone characteristics, viz. Relative Density, Loose Bulk Density, Consolidated Bulk Density, Coefficient of Thermal Expansion, 10% FACT and Elastic Modulus were highly correlated with each other (the absolute values of the correlation coefficient ranged from 0.71 to 0.92). Since none of these variables was characterised by a great deal of missing data, nor did any of them exhibit a trend with respect to any of the three dependent variables, it was not possible to eliminate or make a preferential selection of any variables from this group. They were subsequently evaluated, for their contribution to the multivariate regression model, in combinations which were not highly correlated, along with other (uncorrelated) stone descriptor variables.
- The recipe *content* variables (Cement Content, CSF Content, FA Content, GGBS Content, GGCS Content, GGFS Content, Limestone Content) were all, by definition, highly correlated with their corresponding recipe *fraction* variables (Cement Fraction, etc.) which meant that for the regression the concrete mix had to be specified in terms of either set of variables but not both at the same time.
- The Cement Content variable was highly correlated with the 28-day Compressive Strength of the concrete (r = 0.71). The 28-day Compressive Strength variable, in turn, was found to be highly correlated with the Water / Binder Ratio (r = -0.84) as well as the Aggregate / Binder Ratio (r = -0.81). These correlations were all as expected.
- By definition, the PVF, Water / Binder Ratio and the Aggregate / Binder Ratio were all highly correlated with each other. The absolute values of

the correlation coefficients ranged from 0.83 to 0.95. Likewise the Water Content and PVF variables were highly correlated (r = 0.81).

- The Sand Content and the Stone / Sand Ratio variables were, by definition, also highly correlated (r = -0.91).
- The Temperature and Humidity variables were found to be correlated (r = -0.71) due to the combinations of experimental conditions used.
- Finally, two unexpected correlations were found: Sand Fineness Modulus with Stone Coefficient of Thermal Expansion (-0.80) and Age at First Drying (0.79), respectively. These correlations do not have a practical explanation and arose from the particular combination of experimental conditions present in the overall data set.

Inspection of the 95% bivariate confidence ellipses for the plots of the covariates versus the three dependent variables revealed that, in most cases, there were some outliers. The outliers, when assessed across the three dependent variables, did not correspond to particular cases (experiments), so it was decided not to exclude any cases at this point but rather to assess the regression models and to remove outliers at that stage, if necessary.

The **covariates were then grouped into sets for evaluation by multivariate multiple regression**. The sets were constructed so that each set would adequately describe the composition and properties of the concrete, while avoiding the collinear relationships between the variables described above. Each set of covariates would then be subjected to variable selection as part of the multivariate multiple regression to find that subset of covariates within each set which best described the variability in the dependent variables. The models thus derived from each set of covariates would then be compared to find the overall best model. The construction of the covariate sets is shown in Figure 6.3.

A: Cement

Cement Early Strength Gain Class
plus one of the following sub-groups:

A1: Cement Type
OR
A2: CSF Content
FA Content
GGBS Content
GGCS Content
GGFS Content
Limestone Content
OR
A3: CSF Fraction
FA Fraction
GGBS Fraction
GGCS Fraction
GGFS Fraction
Limestone Fraction

AND

B: Stone

B1: Stone Type
OR
ln(Stone Water Absorption)
Stone Flakiness Index
plus one of the following combinations:
B2-1: Stone Relative Density, Stone Unconfined Compression
B2-2: Stone Loose Bulk Density, Stone 10% FACT
B2-3: Stone Consolidated Bulk Density, Stone Unconfined Compression
B2-4: Stone Coefficient of Thermal Expansion, Stone Unconfined Compression
B2-5: Stone Loose Bulk Density, Stone Elastic Modulus
B2-6: Stone Loose Bulk Density, Stone Unconfined Compression
B2-7: Stone Coefficient of Thermal Expansion, Stone Elastic Modulus

AND

C: Sand

C1: Sand Type
OR
C2: Sand Relative Density
Sand Fineness Modulus
(exclude Stone Coefficient of Thermal Expansion and Age at First Drying)

AND

D: Concrete mix specification / concrete properties

Admixture Type & Content
ln(Cement Content)
Water Content
Slump
plus one of the following:
D1-1: Sand Content, Stone Content, Water / Binder Ratio
D1-2: Sand Content, Stone Content, Aggregate / Binder Ratio
D1-3: Stone / Sand Ratio, Water / Binder Ratio
D1-4: Stone / Sand Ratio, Aggregate / Binder Ratio

OR

Admixture Type & Content
ln(Cement Content)
PVF
Slump
plus one of the following:
D2-1: Sand Content, Stone Content

D2-2: Stone / Sand Ratio

OR

Admixture Type & Content
28-day Compressive Strength
PVF
Slump
plus one of the following:
D3-1: Sand Content, Stone Content
D3-2: Stone / Sand Ratio

AND

E: Shrinkage test conditions

Laboratory
Age at First Drying
ln(2*Volume To Surface Area)
plus one of the following
E1: Temperature
E2: Humidity

Figure 6.3 Construction of the covariate sets used in the multivariate multiple regressions

Given that all combinations listed in Figure 6.3 would be tested, there were 3(A) x 8(B) x 2(C) x 8(D) x 2(E) = 768 sets of covariates to be used as a starting point for the multivariate multiple regressions.

The last step in the preparation of the covariates was the **coding of the categorical covariates**. These were coded as indicator variables, relative to a reference category (coded as 0) specified by Ballim and Alexander[(67)]:

- Cement type: CEM I
- Cement strength class: 32.5
- Cement Early Strength Gain Class: N
- Stone type: Andesite
- Sand type: Dolomite
- Admixture type: No admixture used
- Laboratory: WITS

For the Cement Early Strength Gain Class, where this classification was not yet in use (prior to 1998), the cases were classified as N (normal). The Cement Type variable coding for CEM II A-S, CEM II B-S and CEM IIIA was further split into three variables for each cement type to reflect the three types of slag (GGBS, GGCS and GGFS) used in these cements. Where mixtures of stone types were used in an experiment, this was represented by the use of fractions in the corresponding indicator variables. The same was done in cases where mixtures of sand types were used. Since the levels at which the various admixtures were used differed substantially in some cases, but always made a very small contribution to the overall concrete mix composition, it was decided to code the admixture content together with admixture type.

The resulting data file 'Data for multivariate multiple regression.xls' may be found on the included CD. A partial listing of the file is shown in Appendix H.

6.5 Multivariate multiple regression analysis

Variable selection in multivariate multiple regression, as in multiple regression with one dependent variable, may be done in the following ways(55):

- Forward addition: The procedure starts with a regression model containing the one covariate which is the most highly correlated with the dependent variables. The partial correlation coefficients are then examined to find the next covariate which explains the largest statistically significant portion of the error remaining from the first regression equation. This covariate is added to the regression model. The process is continued until no further covariates are eligible for addition to the model. One drawback of this method is that only one variable is considered for selection at a time; if two variables together explain a significant portion of the variance, but neither is significant by itself, neither variable would be considered for inclusion in the model. A further drawback is that there is no opportunity to remove variables from the model which no longer contribute significantly once a new variable has been added.
- Backward elimination: An initial regression equation is calculated with all the covariates and then all covariates which do not make a significant contribution are removed. There is no opportunity for eliminated variables to re-enter the model at a later stage.

- Forward stepwise regression: This procedure is similar to the forward addition procedure, but at each stage the variables already in the model are assessed to see if they still make a significant contribution to the model and are removed from the model if they do not. The approach suffers from the same drawback as forward addition, namely that if two variables together explain a significant portion of the variance, but neither is significant by itself, neither variable would be considered for inclusion in the model.
- Backward stepwise regression: This procedure is similar to backward elimination, but at each stage the variables not currently in the model are assessed to see if they make a significant contribution to the model and are included in the model if they do. Multicollinear variables can have a substantial impact on the final model description in both stepwise procedures: if one of two multicollinear variables enters the model, it is highly unlikely that the other variable will enter too, and it may be erroneously concluded that the second variable is not important.
- All-possible-subsets: All possible subsets of the covariate set are evaluated and the best-fitting set of covariates is identified. This is a computationally intensive procedure, and does not take multicollinear relationships between variables into account.

It was decided to use the **forward and backward stepwise variable selection** approaches for variable selection within each of the 768 groups of covariates identified in the previous section. Although the ratio of the number of observations to covariates (~ 3 to 7, given that the number of covariates in the covariate sets ranged from 26 to 59) was lower than the conservative figure of 50:1 for these techniques[55], these approaches offered the greatest flexibility in dealing with the large number of covariates in many of the covariate sets. Highly multicollinear covariates had already been separated by the definition of the covariate groups. All-possible-subsets regression was also attempted but could not be carried out due to the large number of variables in the covariate groups which led to computer memory allocation problems.

The multivariate multiple regression analyses were initially attempted in SAS[58]. Multivariate multiple regression in SAS requires the use of PROC GLM, which does not offer built-in forward or backward stepwise (or all-possible-subsets) regression. A SAS program for variable selection in multivariate multiple regression published by Al-Subaihi[68] was therefore adapted for this purpose. However, the design matrix (Z in Section 6.1) was singular in many cases, which meant that the parameter estimates could not be calculated (since $Z'Z$ could not be inverted). This typically occurred from the combination of missing data and the use of a large number of dummy variables which resulted in one or more columns of the design matrix having zero variance. Attempted Gram-Schmidt orthogonalisation[69] of Z'Z did not converge.

The forward and backward stepwise regressions were therefore carried out in STATISTICA[70] where interactive elimination of covariates from the starting set of covariates, to remove the problems with the design matrix described above, could be carried out more easily. Thus covariate sets where the design matrices were singular were reduced to combinations which could be subjected to regression analysis. The p-value for entry or removal from the model was 0.05 and the multivariate significance of the model terms was assessed by Wilk's lambda. Each experiment was given equal weighting in the regression. The only goodness-of-fit measure for multivariate multiple regression models available in STATISTICA was the adjusted R^2. The regression results are presented in Appendix I together with the *average* adjusted R^2 for the three dependent variables for each model.

In many cases, as expected from the initial work done in SAS, the covariate combinations gave rise to one or more columns in the design matrix with no variance, usually due to a combination of missing data and the use of indicator variables. It was thus necessary to exclude certain covariates from some of the covariate sets, as indicated in Appendix I. These covariates included the Admixture covariates, Slump, Cement Early Strength Gain Class (all of which were regarded as not being of pivotal importance in the prediction of shrinkage) as well as some of the continuous stone descriptor variables (which were potentially rather more important).

When the non-significant (pooled) covariates were removed completely from a particular model (the result of a forward or backward stepwise variable selection process), often the number of cases (experiments) available for the regression increased (due to missing data associated with the covariates no longer included in the regression). In many instances this resulted in further covariates becoming non-significant and having to be removed (i.e. the backward or forward variable selection process had to be repeated on a reduced set of covariates and a larger set of cases). Inevitably this led to a reduction in the average adjusted R^2 for that particular model. Since the aim of this work was to identify a group of good models from which the best model would ultimately be chosen, rather than to perform an exhaustive set of regressions, the further removal of non-significant variables after the initial forward or backward stepwise variable selection process was done only for those models with an *initial* average adjusted R^2 greater than or equal to 0.62, as indicated in Appendix I. This value was chosen to give a reasonable number of models for further evaluation, as will be shown shortly.

It was noted, after analysis of the first 34 covariate groups, that, for a given covariate set, the model derived by backward stepwise regression invariably had average adjusted R^2 values which were the same or higher than those for the model derived by forward stepwise regression. Therefore, forward stepwise variable selection was carried out only for those covariate sets for which backward stepwise variable selection had given a model with an (initial) average adjusted R^2 of 0.62 or greater. Likewise, it was also noted, after analysis of the first 68 covariate groups, that the covariate groups which included Temperature gave better models (as determined by the average adjusted R^2 values) than the corresponding covariate groups which included Humidity. As a result, it was decided to abandon the analysis of those covariate groups which included Humidity rather than Temperature (which effectively halved the number of covariate groups to be tested) except for one analysis in each Cement – Stone – Sand covariate grouping and for any covariate group for which the corresponding Temperature-containing covariate group had given an (initial) average adjusted R^2 of 0.62 or greater, all of which confirmed the earlier conclusion.

Eighteen distinct models with a final average adjusted R^2 of 0.62 or greater were identified as a result of the regression analyses described above. These models are highlighted in Appendix I and their significant covariates are listed in Table 6.7, which also introduces the simpler referencing of the models (Model 1*, Model 2*, etc.) which will be used henceforth. Models with similar covariates are grouped together. Table 6.7 also illustrates the consistency in the covariates which were found to be important in describing the variation in the three dependent variables. It is interesting to note the following:

- Cement Early Strength Gain Class did not feature in any of the models (but often had to be omitted due to large amounts of missing data).
- Cement composition description: All 18 models contained the Cement type descriptors (CEM V A, etc.) whereas none of the models contained the corresponding Content or Fraction covariates.
- Stone description: The models contained either the individual stone types or Stone Relative Density (in one case together with ln(Stone Water Absorption) and Stone Flakiness Index).
- Sand description: Seventeen models contained the individual sand types while only one model contained Sand Relative Density and Sand Fineness Modulus.
- Concrete mix specification: The admixture covariates had to be omitted in all cases. Seventeen models contained combinations of ln(Cement Content), Water Content, PVF, Water/Binder Ratio, Aggregate/Binder Ratio, Sand Content, Stone Content and Stone/Sand Ratio while one model contained the 28-day Compressive Strength. None of the models used the Slump covariate (which had to be omitted in some, but not all, cases).
- Shrinkage test conditions: All the models included ln(2 * Volume To Surface Area) and Temperature, while some included the Laboratory and Age at First Drying covariates as well. No models included Humidity in favour of Temperature (the use of these variables was mutually exclusive since they were highly correlated).

Table 6.7 Significant covariates in the nineteen best models for α, ln(β) and ln(γ)

f = forward stepwise variable selection; b = backward stepwise variable selection 2VTSA = 2 * Volume To Surface Area

Model 1*	**Model 2***	**Model 3***	**Model 4***	**Model 5***	**Model 6***
A1/B2-1/C2/D3-1/E1/b[†]	**A1/B2-1/C1/D1-3/E1/b**	**A1/B2-1/C1/D1-1/E1/f**	**A1/B2-1/C1/D1-2/E1/f**	**A1/B2-1/C1/D1-3/E1/f**	**A1/B2-1/C1/D1-4/E1/f**
CEM II A-L	CEM II B-S GGFS	CEM II B-S GGFS			
CEM II A V	CEM III A GGBS	CEM III A GGBS			
CEM II B-S GGBS		CEM V A	CEM V A	CEM V A	CEM V A
CEM II B-S GGFS	Stone Relative Density	Stone Relative Density	Stone Relative Density	Stone Relative Density	Stone Relative Density
CEM II B V	Sand Wits Quartzite	Sand Wits Quartzite	Sand Wits Quartzite	Sand Wits Quartzite	Sand Wits Quartzite
CEM III A GGBS	Shale	Shale	Shale	Shale	Shale
CEM III A GGFS	Granite	Granite	Granite	Granite	Granite
CEM V A	Dolerite	Dolerite	Dolerite	Dolerite	Dolerite
Stone Relative Density	Andesite	Andesite	Andesite	Andesite	Andesite
ln(Water	Klipheuwel Pit	Klipheuwel Pit	Klipheuwel Pit	Klipheuwel Pit	Klipheuwel Pit
Absorption)	Natural	Natural	Natural	Natural	Natural
Flakiness Index	Cape Flats	Cape Flats	Cape Flats	Cape Flats	Cape Flats
Sand Relative Density	River Vaal	River Vaal	River Vaal	River Vaal	River Vaal
Fineness Modulus	Tillite	Tillite	Tillite	Tillite	Tillite
	Quartzite	Quartzite	Quartzite	Quartzite	Quartzite
Laboratory	ln(Cement Content)	ln(Cement Content)	ln(Cement Content)	ln(Cement Content)	ln(Cement Content)
ln(2VTSA) [††]	Water Content		Water Content	Water Content	Water Content
Temperature	Stone/Sand Ratio	Stone Content	Stone Content	Stone/Sand Ratio	Stone/Sand Ratio
28-day Compressive	Water/Binder Ratio	Water/Binder Ratio	Aggregate/Binder Ratio	Water/Binder Ratio	Aggregate/Binder Ratio
Strength	ln(2VTSA)	ln(2VTSA)	ln(2VTSA)	ln(2VTSA)	ln(2VTSA)
	Temperature	Temperature	Temperature	Temperature	Temperature

Table 6.7 Significant covariates in the nineteen best models for α, ln(β) and ln(γ) (continued)

Model 7*	Model 8*	Model 9*	Model 10*	Model 11*	Model 12*
A1/B1/C1/D1-1/E1/b	**A1/B1/C1/D1-3/E1/b**	**A1/B1/C1/D1-2/E1/b**	**A1/B1/C1/D1-4/E1/b**	**A1/B1/C1/D1-3/E2/b**	**A1/B1/C1/D1-1/E1/f**
CEM III A GGBS	CEM III A GGBS	CEM V A	CEM V A	CEM V A	CEM V A
Stone Wits Quartzite	Stone Wits Quartzite			Stone Wits Quartzite	Stone Shale
Dolerite	Dolerite	Stone Wits Quartzite	Stone Wits Quartzite	Shale	Dolomite
Greywacke	Greywacke	Dolerite	Shale	Dolerite	Granite
Dolomite,	Dolomite, Granite	Greywacke	Dolerite	Greywacke	Tillite
Granite	Pretoria Quartzite	Dolomite,	Greywacke	Dolomite	Quartzite
Pretoria	Sand Wits Quartzite	Granite	Dolomite, Granite	Granite	Sand Wits Quartzite
Quartzite	Granite,	Pretoria	Sand Granite	Sand Granite	Shale
Sand Wits Quartzite	Dolerite	Quartzite	Dolerite, Andesite	Dolerite	Granite
Granite,	Andesite	Sand Wits Quartzite	Klipheuwel Pit	Andesite	Dolerite
Dolerite	Klipheuwel Pit	Granite,	Natural	Klipheuwel Pit	Klipheuwel Pit
Andesite	Natural	Dolerite	Cape Flats	Natural	Natural
Klipheuwel Pit	Cape Flats	Andesite	River Vaal	Cape Flats	River Vaal
Natural	River Vaal	Klipheuwel Pit	Pretoria Quartzite	River Vaal	ln(Cement Content)
Cape Flats	Tillite, Quartzite	Natural	Tillite, Quartzite	Pretoria Quartzite	Stone Content
River Vaal	River	Cape Flats	River	Tillite	Water/Binder Ratio
Tillite, Quartzite	ln(Cement Content)	River Vaal	ln(Cement Content)	Quartzite	ln(2VTSA)
River		Tillite, Quartzite	Water Content	River	Temperature
ln(Cement Content)	Stone/Sand Ratio	River	Stone/Sand Ratio	ln(Cement Content)	
	Water/Binder Ratio	ln(Cement Content)	Aggregate/Binder Ratio	Water Content	
Stone Content	Laboratory	Water Content	Laboratory	Water/Binder Ratio	
Water/Binder Ratio	Age at first drying	Stone Content		Stone/Sand Ratio	
Laboratory	ln(2VTSA)	Aggregate/Binder Ratio	ln(2VTSA)	Laboratory	
Age at first drying	Temperature	Laboratory	Temperature		
ln(2VTSA)					
Temperature		ln(2VTSA)			
		Temperature			

Table 6.7 Significant covariates in the nineteen best models for α, ln(β) and ln(γ) (continued)

Model 13*	**Model 14***	**Model 15***	**Model 16***	**Model 17***	**Model 18***
A1/B1/C1/D2-1/E1/b	**A1/B1/C1/D2-2/E1/b**	**A1/B1/C1/D3-1/E1/b**	**A1/B1/C1/D3-2/E1/b**	**A1/B1/C1/D1-2/E1/f**	**A1/B1/C1/D1-4/E1/f**
CEM II A-L, CEM II A V		CEM II A-S GGCS	CEM II A-S GGCS	CEM III A GGBS	CEM III A GGBS
CEM II A-S GGBS, GGCS,	CEM II A-S GGCS	CEM III A GGBS	CEM III A GGBS	CEM V A	
GGFS		CEM V A	CEM V A	Stone Granite	Stone Dolerite
CEM II B-S GGBS, GGCS,	CEM II B-S GGCS	Stone Wits Quartzite	Stone Wits Quartzite	Tillite	Granite
GGFS		Dolerite	Dolerite	Quartzite	Tillite
CEM II B V, CEM V A	CEM II B V, CEM V A	Greywacke	Greywacke	Sand Wits Quartzite	Quartzite
CEM III A GGBS, GGCS.	CEM III A GGBS, GGCS,	Dolomite, Granite	Dolomite, Granite	Shale	Sand Wits Quartzite
GGFS	GGFS	Pretoria Quartzite	Pretoria Quartzite	Granite	Shale
Stone Wits Quartzite	Stone	Sand Shale	Sand Shale	Dolerite	Granite
Dolerite,Greywacke	Dolerite,Greywacke	Granite, Dolerite	Granite, Dolerite	Andesite	Dolerite
Dolomite, Granite	Dolomite, Granite	Andesite	Andesite	Klipheuwel Pit	Andesite
Pretoria Quartzite	Pretoria Quartzite	Klipheuwel Pit	Klipheuwel Pit	Natural	Klipheuwel Pit
Sand Shale	Sand	Natural	Natural	River Vaal	Natural
Granite, Dolerite	Dolerite	Cape Flats	Cape Flats	ln(Cement Content)	Cape Flats
Andesite, Natural	Andesite, Natural	River Vaal	River Vaal	Water Content	River Vaal
Klipheuwel Pit	Klipheuwel Pit	Pretoria Quartzite	Pretoria Quartzite	Stone Content	Pretoria Quartzite
Cape Flats	Cape Flats	Tillite, River	Tillite, River	Aggregate/Binder Ratio	ln(Cement Content)
River Vaal	River Vaal	PVF	PVF	ln(2VTSA)	Water Content
Pretoria Quartzite	Pretoria Quartzite	Sand Content	Stone/Sand Ratio	Temperature	Stone/Sand Ratio
Tillite, River	Tillite, River	Laboratory	Laboratory		Aggregate/Binder Ratio
ln(Cement Content)	ln(Cement Content)	Age at first drying			ln(2VTSA)
PVF	Stone/Sand Ratio	ln(2VTSA)	ln(2VTSA)		Temperature
Laboratory, ln(2VTSA)	Laboratory, ln(2VTSA)	Temperature	Temperature		
Temperature	Temperature				

It was noted that Model 1* was based on only 118 of the 174 experiments in the regression data set (68% of the data set) whereas the other models were all based on 170 experiments (98%). This resulted from casewise deletion of missing data associated with the covariates used in Model 1*, most notably 28-day Compressive Strength and Sand Fineness Modulus. Regardless of the goodness-of-fit of Model 1*, it would be expected to be a less generalisable model (since it is based on less data) and its applicability to data not included in the 174 profiles would be more limited (again due to missing data in the covariates), compared to the other seventeen models. Therefore Model 1* was eliminated from further consideration.

Model diagnostic tests and plots were carried out for the remaining seventeen models. These diagnostics are similar to those used in univariate multiple regression and comprised[55,65]:

- Examination of the regression variate: Plots of studentised residuals versus predicted dependent variables, observed dependent variables and all independent variables should all be evenly dispersed about zero and show no obvious patterns or trends.
- Checks of the assumed linearity of the overall relationship: The plot of residuals versus the predicted dependent variable should not show obvious patterns. Partial regression plots for the residuals versus each independent variable should not exhibit any relationship or outliers.
- A check of the assumption of homoscedasticity: The plot of residuals versus the predicted dependent variable should not fan out.
- A check of the assumption of the independence of the error terms: The errors between experiments can be assumed to be independent since each experiment was based on a separately prepared concrete mix. The errors between parameters are not assumed to be independent (hence the use of multivariate multiple regression).
- Checks of the assumption of the normality of the error term distribution: The residuals should be approximately normally distributed and the normal probability plots of the residuals should be approximately linear.
- Examination for outliers and influential points**:** Outliers were identified by means of large standardised or studentised residuals and also from the partial regression plots. Influential and leverage points were identified by

means of the leverage, Mahalanobis D^2, studentised deleted residuals, Cook's distance and SDFFIT statistics[55].

The findings for all the models were very similar, and are thus discussed in general terms here. The detailed results of the diagnostic checks are tabulated in Appendix J for all the models. Over all models, plots of the studentised residuals versus the predicted dependent variables indicated only one outlier (defined here as the absolute value of a studentised residual near or greater than 4) for each of the three dependent variables: case 169 for α, case 170 for ln(β) and case 165 for ln(γ). (Case numbers refer to the data as listed in the data file 'Data for multivariate multiple regression.xls' on the included CD.) These outliers can also be observed in the plots of the studentised residuals versus the observed dependent variables as well as in the plots of the observed versus the predicted dependent variables (more so for α than for ln(β) and ln(γ)).

These three outlier cases all came from the same data set (set 22) which also contained nine other cases *not* found to be outliers or influential points. The value of α for case 169 (824) is the largest (by ~ 100 microstrain) within the 174 cases used in the regression analysis and so its identification as an outlier was not surprising. The originators of the data also commented on the particularly high shrinkage for this sample, relating it to the particular batch of cement used in the sample preparation[71]. The value of ln(β) for case 170 (-3.27) was itself not an outlier, but it should be noted that the same batch of cement as that used for case 169 was used in the manufacture of the sample for case 170. Finally, the value of ln(γ) for case 165 (0.41) was the largest value (by 0.14 units) within the 174 cases used in the regression analysis and so again its identification as an outlier in the regression model was perhaps not surprising.

The overall shapes of the residual plots indicated that there was no heteroscedasticity of errors, which was confirmed by the distributions of the studentised residuals as well as their normal probability plots. Plots of the residuals versus time (Test Year covariate) as well as versus covariates both included and not included in the models were all acceptable.

The Mahalanobis D^2 statistics showed that case 70 (for all models) as well as case 79 (for Model 11*) were very different to the other cases based on their combination of covariates. Case 70 is the lone representative of its data set (set 13) and also of the Quartz Alluvial stone type and the River Windhoek sand type. It is also characterised by a relatively low sand content, high stone content and the highest stone/sand ratio. Case 79 is the only case from its data set (set 15) to have been retained in the 174 profiles used in the regression analysis. It is the only representative of the Ecca Grit sand type (which is significant only in Model 11*).

The leverage statistics identified between 14 and 21 significant cases for each of the models, including some overlap of cases between the different models. The cases with significant leverage included cases 70 and 79. Since the cases exerting a high degree of leverage on the regressions were well spread across the data sets making up the overall data set (they came from 12 out of the 19 data sets represented in the 174 cases used in the regression analyses), they were not regarded as a cause for concern.

Examination of the studentised deleted residuals identified 3 to 8, 5 to 10 and 8 to 11 significant influential points for α, $\ln(\beta)$ and $\ln(\gamma)$ respectively in the different models. These included the three outliers identified above. The Cook's distance statistics identified case 70 (for all models except Model 10*) and case 79 for Model 11* as having a large influence on the regression equation. The SDFFIT statistics identified different cases for different models with significant influence on the regression equation (see Appendix J). These were cases 79, 136, 139, 165 and 169 (the latter two cases were also identified as outliers).

In the second step, **the three outliers (cases 165, 169 and 170) identified above were removed from the data set and the seventeen selected regression analyses were repeated**. The findings for all the models were again very similar, and are thus discussed in general terms here. Detailed results are tabulated in Appendix J.

The average adjusted R^2 for each model increased upon deletion of the outliers, as would be expected. All model terms remained significant, except for one term in Model 8* (see Appendix J). Some stone and sand terms were lost from the models as they were represented only by the deleted cases (see Appendix J). For all models, plots of the studentised residuals versus the predicted dependent variables indicated no outliers. The overall shapes of the residual plots indicated that there was no heteroscedasticity of errors, which was confirmed by the distribution of the residuals and their normal probability plots. Plots of the residuals versus time (Test Year covariate) as well as covariates both included and not included in the models were all acceptable.

The Mahalanobis D^2 statistics showed again that case 70 (for all models) as well as case 79 (for Model 11*) were very different to the other cases based on their combination of covariates. The leverage statistics identified between 12 and 19 significant cases for each of the models, including cases 70 and 79. The studentised deleted residuals identified 5-13, 6-11 and 7-13 significant influential points for α, $\ln(\beta)$ and $\ln(\gamma)$ respectively for the different models. The Cook's distance statistics again identified case 70 (for all models) and case 79 for Model 11* as having a large influence on the regression equation. The SDFFIT statistics identified different cases for different models with significant influence on the regression equation (see Appendix J). These were cases 70, 79, 136 and 139.

In the final step, **each model was re-evaluated after those cases (for a particular model) with very high Mahalanobis D^2 statistics (as indicated above) or significant Cook's distances or significant SDFFIT statistics were omitted from the model.** Again, the results are tabulated in Appendix J. The diagnostic plots for one model (Model 17*) are shown in Appendix K.

The average adjusted R^2 for each model increased in some cases and decreased slightly in other instances upon deletion of these cases. Model terms for which the deleted cases were the sole representative were obviously lost from the regression model. For some models, one of the cement terms was no longer significant and was removed from the model (see Appendix J).

For all models, plots of the studentised residuals versus the predicted dependent variables indicated no outliers. The overall shapes of the residual plots indicated that there was no heteroscedasticity of errors, which was confirmed by the distributions of the studentised residuals as well as their normal probability plots. Plots of the residuals versus time (Test Year covariate) did not show any trends over the testing period, indicating that the changes in variability of the profile estimates were not due to any changes in laboratory procedures or instrumentation over the years. This was also confirmed by a similar plot done separately for each of the two laboratories. Plots of the residuals versus covariates both included and not included in the models were all acceptable.

The Mahalanobis D^2 statistics now indicated no cases which were very different to the other cases based on their combination of covariates, as expected. The leverage statistics identified between 8 and 17 significant cases for each of the models. The studentised deleted residuals identified 5-14, 4-12 and 8-12 significant influential points for α, $\ln(\beta)$ and $\ln(\gamma)$ respectively. No influential cases were identified by the Cook's distance statistics. The SDFFIT statistics identified some new influential cases for a few of the models (see Appendix J). It was decided, based on the acceptable quality of the diagnostic plots, not to delete any further cases.

Whilst care had been taken to exclude multicollinear covariates from the models, **the final versions of each model**, i.e. with the cases as deleted in the final step above, **were examined for multicollinearity** by inspecting the variance inflation factors of the model terms. Variance inflation factors greater than 10 were considered significant[(55)] (corresponding to multiple correlations, between the covariate in question and the other covariates in the model, greater than 0.95). It was found that all the models containing the Laboratory covariate (viz. Models 7*, 8*, 9*, 10*, 11*, 13*, 14*, 15* and 16*) exhibited high (10) to extremely high (900) variance inflation factors for many covariates, including some Cement Type, Stone Type and Sand Type covariates as well as the ln (2 * Volume To Surface Area), Age at First Drying and Temperature covariates. (Removal of the Laboratory covariate from these models led to a number other covariates in the model becoming non-significant, as a result of the multicollinearity.) In contrast, for remaining eight models which did not contain the Laboratory covariate, only the (2 * Volume To Surface Area) covariate exhibited a significant variance inflation factor ranging between 10 and 12.

High multicollinearity results in difficulties in the correct estimation of regression coefficients (coefficients may have an incorrect sign and magnitude at high collinearity) as well as in separating the contribution of each covariate. Since the aim of this work was to produce a model which has a practical interpretation, and there was a range of candidate models, it was decided to retain only those models with a relatively low degree of multicollinearity (viz. Models 2*, 3*, 4*, 5*, 6*, 12*, 17* and 18*) for further consideration. Moreover, use of a model which does not contain the Laboratory covariate would make the predictive use of the model more generally acceptable. This is not to say that there may be differences in shrinkage measurements as performed in the two laboratories included in the data set, but these differences appear to be confounded with the Stone and Sand types, as was mentioned in Chapter 3 and which has now been confirmed by the multicollinearity of the models containing the Laboratory term. Further investigation into these possible inter-laboratory differences was recommended in Chapter 3.

6.6 Exploration of alternative weighting schemes for the multivariate multiple regression analysis

The results presented in the previous section were based on equal weighting of the cases. Alternative weighting schemes were considered which reflected differences in the precision with which the dependent variables α, $\ln(\beta)$ and $\ln(\gamma)$ had been estimated for each shrinkage profile:

- Weighting each case by the inverse of the generalised variance and

- weighting each case in proportion to the number of data points contained in its shrinkage profile.

These weighting schemes were applied to the covariate sets from which Models 2* - 18* had been obtained.

6.6.1 Weighting by the inverse of the generalised variance

The generalised variance for a particular case (experiment) was calculated by taking the square root of the determinant of the variance-covariance matrix for the fit of the growth curve model to the shrinkage-time measurements for that particular case. The SAS program listing is given in Appendix L and a partial listing of the weights is shown at the end of the table in Appendix H (with a full listing in 'Data for multivariate multiple regression.xls'). The weights were found to range from 32 to 1688745, reflecting differences in both the number of data points as well as the goodness-of-fit of the growth curve model across the 174 cases comprising the regression data set. The distribution of these weights, shown in Figure 6.4 (on a $\log_{10}$ scale) , indicated that there was too much variation in the magnitude of the weights for them to be used in a sensible weighting scheme. This weighting scheme was thus abandoned.

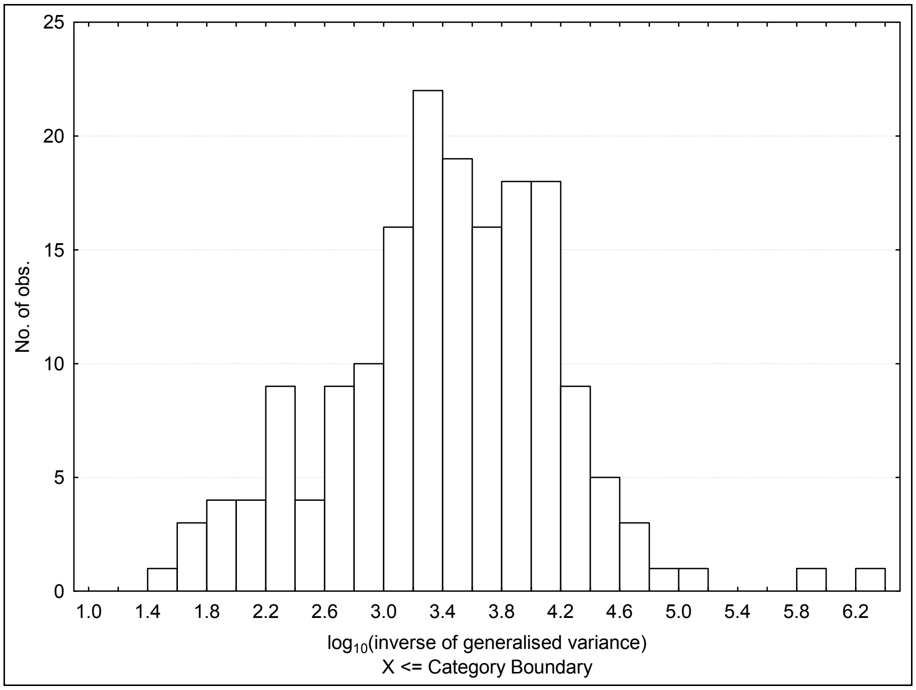


Figure 6.4 Frequency distribution of the case weights from the inverse of the generalised variance, on a $\log_{10}$ scale

6.6.2 Weighting each case in proportion to the number of data points contained in its shrinkage profile

The number of data points in each shrinkage profile contained in the 174 cases used in the regression analysis ranged from 5 to 55. The distribution of weights is shown in Figure 6.5. This level of variation in the weights was deemed to be acceptable for the construction of a weighting scheme.

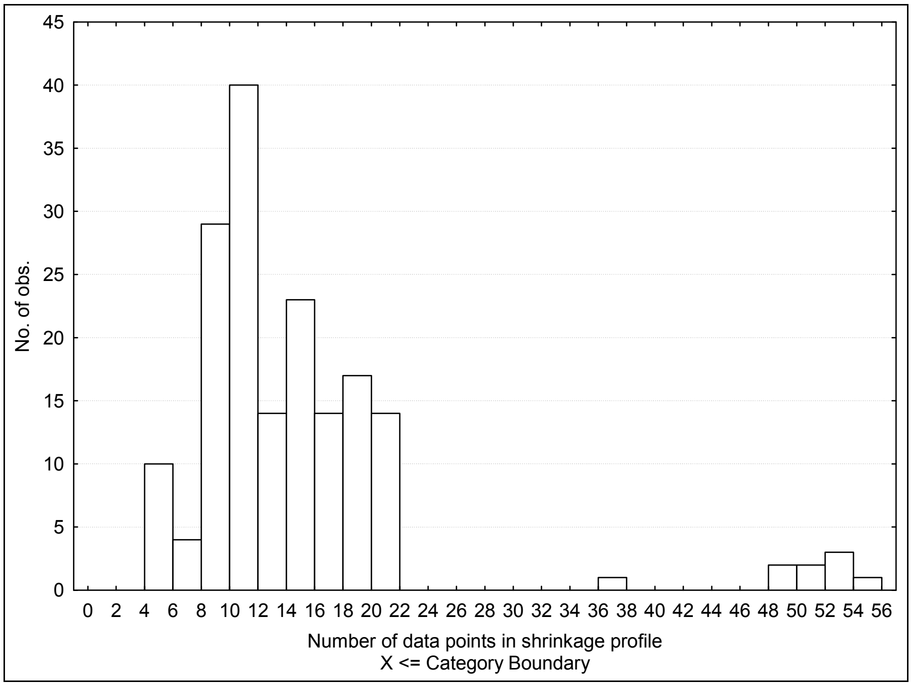


Figure 6.5 Frequency distribution of the number of data points in each shrinkage profile

These values were scaled, for use in STATISTICA, so that the lowest weight would have a value of 1, which gave a range of weights between 1 and 11. (STATISTICA carries out weighted regression by entering cases with a weight of *n* into the regression *n* times. Non-integer weights are rounded to the nearest integer.) Since it was known that the unweighted models all identified cases 165, 169 and 170 as outliers (see discussion in previous section), the initial weighted regression analyses were carried out on a reduced data set of 171 cases which did not contain the three outlier cases. For each of the seventeen covariate sets identified in the previous section, the weighted regressions were carried out by performing either forward or backward stepwise regression (whichever had been used for the corresponding unweighted regression model).

As a result of the manner in which weighted regression is carried out in STATISTICA, the number of cases in each regression analysis was now much higher (~ 520 rather than 171) which gave rise to artificially increased power and higher degrees of freedom. In order to remove model terms which were not actually significant, the weighted multivariate multiple regression with the significant terms identified by the regression in STATISTICA was repeated in SAS (but without variable selection) and non-significant terms were manually removed one by one (starting with the least significant term) until no further non-significant terms remained. The SAS program used in this step is listed in Appendix M.

A further twelve distinct models (designated Models 19*W – 30*W), all with average adjusted R^2 values of 0.62 or greater, were thus derived. (Some covariate sets gave rise to identical weighted models, or models which were subsets of other models.) The models' significant covariates are listed in Table 6.8 (models with similar covariates are grouped together), while the results of the diagnostic checks are tabulated in Appendix N. The findings for all the models were very similar, and are thus discussed in general terms here.

In all cases the weighted regression models had somewhat different significant covariates compared to their unweighted counterparts derived from the same covariate sets, indicating that the weighting did make a difference. The average adjusted R^2 values were, in each case, higher for the weighted regression than for the corresponding unweighted regression derived from the same covariate set.

Table 6.8 Significant covariates in the twelve best weighted models for α, ln(β) and ln(γ).
f = forward stepwise variable selection; b = backward stepwise variable selection. 2VTSA = 2 * Volume To Surface Area

Model 19*W	**Model 20*W**	**Model 21*W**	**Model 22*W**	**Model 23*W**	**Model 24*W**
A1/B2-1/C1/D1-3/E1/b	**A1/B2-1/C1/D1-1/E1/f**	**A1/B2-1/C1/D1-2/E1/f**	**A1/B1/C1/D1-1/E1 or E2/b**	**A1/B1/C1/D1-2/E1/f**	**A1/B1/C1/D1-1/E1/f**
CEM II A-S GGCS	**A1/B2-1/C1/D1-3/E1/f**	**A1/B2-1/C1/D1-4/E1/f**	**A1/B1/C1/D1-3/E1/b**		CEM II A-S GGCS
CEM II B-S GGCS	CEM II A-S GGCS	CEM II A-S GGCS	CEM II A-S GGCS	CEM II A-S GGCS	CEM II B-S GGFS
CEM III A GGCS	CEM II B-S GGBS	CEM II B-S GGBS	CEM II B-S GGCS	CEM II B-S GGCS	CEM III A GGBS
CEM III A GGFS	CEM II B-S GGFS	CEM II B-S GGFS	CEM III A GGCS	CEM III A GGCS	CEM III A GGFS
CEM V A	CEM III A GGBS	CEM III A GGBS	CEM III A GGFS, CEM V A	CEM III A GGFS, CEM V A	Stone Wits Quartzite
Stone Relative	CEM III A GGFS	CEM III A GGFS	Stone Wits Quartzite	Stone Wits Quartzite	Shale, Dolerite
Density	Stone Relative	Stone Relative	Shale, Dolerite	Shale, Dolerite	Greywacke
Sand Wits Quartzite	Density	Density	Greywacke	Greywacke	Dolomite, Granite
Shale	Sand Wits Quartzite	Sand Wits Quartzite	Dolomite, Granite	Dolomite, Granite	Quartz Alluvial
Granite	Shale	Shale	Quartz Alluvial	Quartz Alluvial	Pretoria Quartzite
Dolerite	Granite	Granite	Pretoria Quartzite	Pretoria Quartzite	Tillite
Andesite	Dolerite	Dolerite	Tillite	Tillite	Sand Wits Quartzite
Klipheuwel Pit	Andesite	Andesite	Sand Wits Quartzite	Sand Wits Quartzite	Granite, Dolerite
Natural	Klipheuwel Pit	Klipheuwel Pit	Granite, Dolerite	Granite, Dolerite	Andesite, Klipheuwel
River Windhoek	Natural	Natural	Andesite,	Andesite,	Pit
Cape Flats	River Windhoek	River Windhoek	Klipheuwel Pit	Klipheuwel Pit	Natural, Cape Flats
River Vaal	Cape Flats	Cape Flats	Natural, Cape	Natural, Cape	River Vaal
Tillite	River Vaal	River Vaal	Flats	Flats	Pretoria Quartzite
Quartzite	Tillite	Tillite	River Vaal	River Vaal	Quartzite, River
ln(Cement Content)	Quartzite	Quartzite	Pretoria Quartzite	Pretoria Quartzite	ln(Cement Content)
Water Content	River	River	Quartzite, River	Quartzite, River	Water Content
Stone/Sand Ratio	ln(Cement Content)	ln(Cement Content)	Ecca Grit	Ecca Grit	Stone Content
Water/Binder Ratio	Water Content	Water Content	ln(Cement Content)	ln(Cement Content)	Water/Binder Ratio
ln(2VTSA)	Water/Binder Ratio	Aggregate/Binder Ratio	Water Content	Water Content	Laboratory
Temperature	ln(2VTSA)	ln(2VTSA)	Water/Binder Ratio	Aggregate/Binder Ratio	Age at first drying
	Temperature	Temperature	Laboratory, ln(2VTSA)	Laboratory, ln(2VTSA)	ln(2VTSA)
			Temperature	Temperature	Temperature

Table 6.8 Significant covariates in the twelve best weighted models for α, ln(β) and ln(γ) (continued)

Model 25*W	**Model 26*W**	**Model 27*W**	**Model 28*W**	**Model 29*W**	**Model 30*W**
A1/B1/C1/D2-1/E1/b	**A1/B1/C1/D2-2/E1/b**	**A1/B1/C1/D3-1/E1/b**	**A1/B1/C1/D3-2/E1/b**	**A1/B1/C1/D1-2/E1/f**	**A1/B1/C1/D1-4/E1/f**
CEM II A-S GGCS CEM II B-S GGFS CEM III A GGBS CEM III A GGFS Stone Wits Quartzite Shale, Dolerite Greywacke Dolomite, Granite Quartz Alluvial Pretoria Quartzite Tillite Sand Wits Quartzite Granite, Dolerite Andesite, Klipheuwel Pit Natural, Cape Flats River Vaal Pretoria Quartzite Quartzite, River ln(Cement Content) PVF, Laboratory ln(2VTSA) Temperature	CEM II A-S GGBS, GGCS, GGFS CEM II B-S GGBS, GGCS, GGFS CEM III A GGBS, GGCS GGFS and CEM V A Stone Wits Quartzite Shale, Dolerite Greywacke Dolomite, Granite Quartz Alluvial Pretoria Quartzite Tillite Sand Wits Quartzite Granite, Dolerite Andesite, Klipheuwel Pit Natural, Cape Flats River Vaal Pretoria Quartzite Quartzite, River ln(Cement Content) PVF, Laboratory ln(2VTSA) Temperature	CEM II A-S GGCS CEM II B-S GGFS CEM III A GGBS CEM III A GGFS CEM V A Stone Wits Quartzite Shale, Dolerite Greywacke Dolomite, Granite Quartz Alluvial Pretoria Quartzite Tillite Sand Wits Quartzite Granite, Dolerite Andesite, Klipheuwel Pit Natural, Cape Flats River Vaal Pretoria Quartzite Quartzite, River PVF, Sand Content Laboratory Age at first drying ln(2VTSA) Temperature	CEM II A-S GGCS CEM II B-S GGFS CEM III A GGBS CEM III A GGFS CEM V A Stone Wits Quartzite Shale, Dolerite Greywacke Dolomite, Granite Quartz Alluvial Pretoria Quartzite Tillite Sand Wits Quartzite Granite, Dolerite Andesite, Klipheuwel Pit Natural, Cape Flats River Vaal Pretoria Quartzite Quartzite, River PVF, Stone/Sand Ratio Laboratory Age at first drying ln(2VTSA) Temperature	CEM II A-S GGCS CEM II B-S GGBS CEM II B-S GGFS CEM III A GGBS CEM III A GGFS Stone Wits Quartzite Shale, Greywacke Granite Quartz Alluvial Pretoria Quartzite Tillite Sand Wits Quartzite Granite, Dolerite Andesite Klipheuwel Pit Natural, Cape Flats River Vaal Pretoria Quartzite Quartzite, Ecca Grit ln(Cement Content) Water Content Stone Content Aggregate/Binder Ratio Laboratory, ln(2VTSA) Temperature	CEM II A-S GGCS CEM II B-S GGFS CEM III A GGBS CEM III A GGFS CEM V A Stone Wits Quartzite Shale, Dolerite Greywacke Granite Quartz Alluvial Pretoria Quartzite Tillite Sand Wits Quartzite Granite, Dolerite Andesite Klipheuwel Pit Natural, Cape Flats River Vaal Pretoria Quartzite Quartzite, Ecca Grit Water Content Aggregate/Binder Ratio Laboratory ln(2VTSA) Temperature

For all models except Model 22*W, plots of the studentised residuals versus the predicted dependent variables indicated no outliers. The two outliers (cases 20 and 25) for Model 22*W were removed. The overall shapes of the residual plots for all the models indicated that there was no heteroscedasticity of errors, which was confirmed by the distributions and normal probability plots of the studentised residuals. Plots of the residuals versus time (Test Year covariate) did not show any trends over the testing period, indicating that the changes in variability of the profile estimates were not due to any changes in laboratory procedures or instrumentation over the years. This was also confirmed by a similar plot done separately for each of the two laboratories. Plots of the residuals versus covariates both included and not included in the models were all acceptable.

The Mahalanobis D^2 statistics indicated no cases which were very different to the other cases based on their combination of covariates, except for case 79 for Model 22*W which was also removed. Case 70, which had been problematic in the unweighted regressions, did not feature in the weighted models due to its relatively high weighting. The leverage statistics identified between 14 and 35 significant cases for each of the models which was slightly more than in the case of the unweighted regressions. The studentised deleted residuals identified 10-17, 10-15 and 11-14 significant influential points for α, $\ln(\beta)$ and $\ln(\gamma)$ respectively for the different models which was also generally slightly more than in the case of the unweighted regressions. No influential cases were identified by the Cook's distance or SDFFIT statistics.

Inspection of the variance inflation factors of the model terms once again indicated that the models containing the Laboratory covariate (viz. Models 22*W - 30*W) exhibited high (10) to extremely high (14700) variance inflation factors for many covariates, including some Cement Type, Stone Type and Sand Type covariates as well as the ln (2 * Volume To Surface Area), Age at First Drying, Laboratory and Temperature covariates. In the remaining three weighted models which did not contain the Laboratory covariate, only the (2 * Volume To Surface Area) and Sand Type Dolerite covariates exhibited significant variance inflation factors ranging between

10 and 17. Thus, it was again decided to retain only those models with a relatively low degree of multicollinearity (viz. Models 19*W, 20*W and 21*W) for further consideration.

In the next section, the goodness-of-fit of the remaining eleven models (eight from unweighted regression and three from weighted regression) are assessed by a number of goodness-of-fit measures and the best model is selected using MCDA.

6.7 Selection of the best model

Five goodness-of-fit criteria were selected for evaluation of the eleven candidate models. These are defined below in the context of multivariate multiple regression. Where possible, definitions based on an unbiased (rather than a Maximum Likelihood) estimate of the common covariance matrix of the multivariate multiple regression were used. In all cases, the following definitions apply:

n = number of cases (experiments)

k = number of parameters (terms) in the model (including the intercept)

q = number of dependent variables (= 3 in this study)

MSE (mean square error matrix) = $\dfrac{SSE}{n-k} = \dfrac{1}{n-k}(\mathbf{Y}-\mathbf{Z}\hat{\varphi})'(\mathbf{Y}-\mathbf{Z}\hat{\varphi})$

where **Y** is the matrix of dependent variables, **Z** is the design matrix and $\hat{\varphi} = (\mathbf{Z'Z})^{-1}\mathbf{Z'Y}$ is the matrix containing the estimates of the parameters of the regression (see Section 6.1)

u = unbiased

ML = Maximum Likelihood

The **multivariate final prediction error (MFPE)** is defined as follows[72,73]:

$$MFPE_u = |MSE|\left(\frac{n+k}{n}\right)^q .$$

The best models have the lowest MFPE values.

Aikaike's Information Criterion (AIC) is defined for multivariate multiple regression as[72,73]

$$AIC_u = log\left|MSE\right| + \frac{2kq + q(q+1)}{n}.$$

According to Kim[72], this statistic tends to overfit when there are too many parameters relative to the number of cases, defined as $\frac{n}{3k+6} < 40$. In this study, n ranges between 165 and 167 and k between 19 and 25, which makes the overfitting criterion applicable. Therefore the **Corrected Aikaike's Information Criterion** (CAIC)[72] was used instead:

$$CAIC_u = log\left|MSE\right| + \frac{2kq + q(q+1)}{n-k-q-1}.$$

The best model is indicated by the lowest CAIC value.

The **Bayesian Information Criterion** (BIC) is defined for multivariate multiple regression as[72]

$$BIC_{ML} = n\,log\left|\frac{SSE}{n}\right| + \left\{kq + \frac{1}{2}q(q+1)\right\}log(n).$$

Again, the best model is indicated by the lowest BIC value.

The Hannan and Quinn (HQ) Information Criterion, is defined as[72]

$$HQ_{ML} = n\,log\left|\frac{SSE}{n}\right| + \{2kq + q(q+1)\}log(log(n)).$$

This statistic can overfit when applied to relatively small samples, according to the same definitions as used for the CAIC, so in this study the **Corrected Hannan and Quinn** (CHQ) statistic was used instead:

$$CHQ_u = log\left|MSE\right| + \frac{2kq\,log(log(n))}{n-k-q-1}.$$

The best models have the lowest CHQ values.

Finally, the **average adjusted R^2** (for the three dependent variables) was also used as a measure of goodness-of-fit. The adjusted R^2 is derived from the R^2 value for the regression by adjusting for the number of cases and the number of parameters in the model:

$$Adjusted\ R^2 = 1 - \frac{(n-1)(1-R^2)}{n-k}.$$

The best model has the highest average adjusted R^2 value.

The values of the first four criteria for the eleven models were calculated by means of a SAS program, adapted from a program published by Al-Subaihi[68]. The program listing is given in Appendix O. The adjusted R^2 values were taken from the results of the multivariate multiple regressions carried out in STATISTICA. These results are presented in Table 6.9 and are illustrated by means of a radar chart in Figure 6.6.

Table 6.9 Goodness-of-fit scores for the eleven multivariate multiple regression model

Model	Number of cases	Number of parameters	Goodness-of-fit score				
			MFPE	CAIC	BIC	CHQ	Average adjusted R^2
Model 2*	166	21	8.74	2.789	585.87	3.269	0.660
Model 3*	166	21	9.17	2.837	593.89	3.317	0.664
Model 4*	166	20	8.96	2.781	580.81	3.230	0.667
Model 5*	166	20	9.12	2.800	583.82	3.248	0.661
Model 6*	166	20	9.03	2.789	582.09	3.238	0.662
Model 12*	166	19	9.51	2.809	581.38	3.228	0.658
Model 17*	165	20	8.43	2.725	568.15	3.176	0.666
Model 18*	165	22	8.10	2.753	580.04	3.265	0.664
Model 19*W	167	25	8.79	2.929	626.39	3.530	0.691
Model 20*W	167	25	8.34	2.877	617.76	3.478	0.697
Model 21*W	167	25	8.31	2.873	617.05	3.474	0.698

For each goodness-of-fit statistic, the models with the first, second and third best scores are highlighted in dark orange, light orange and yellow, respectively

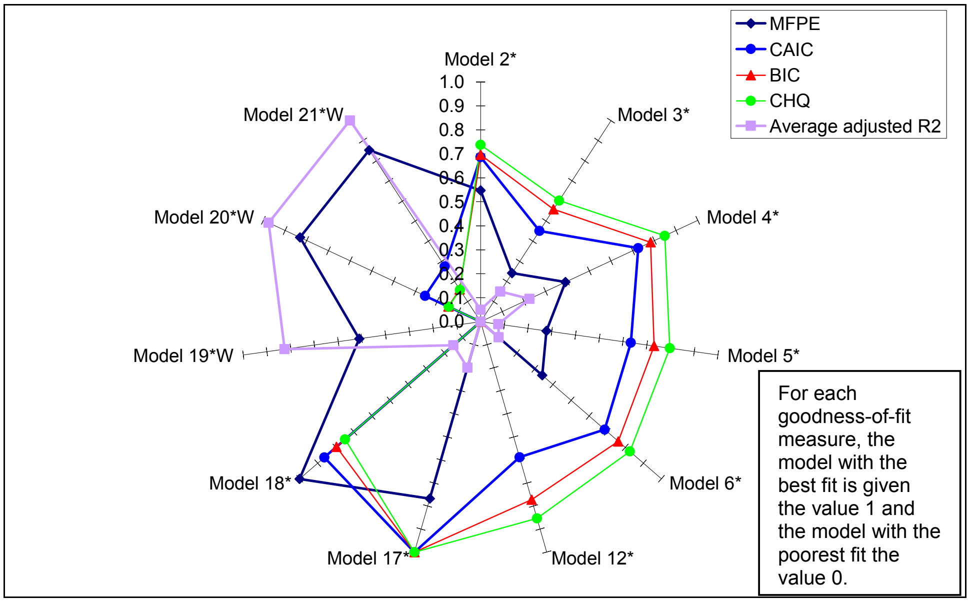


Figure 6.6 Comparison of the goodness-of-fit scores for the eleven multivariate multiple regression models

The unweighted models had higher CAIC, BIC and CHQ scores than the weighted models, while the weighted models had higher average adjusted R^2 values than the unweighted models. There was no one model which had the best scores across all five goodness-of-fit measures, although Model 17* had the highest scores for CAIC, BIC and CHQ. Therefore, selection of the best model was carried out by MCDA, based on the five goodness-of-fit measures. The Value Function and Outranking methods, described in Chapter 5, were used once again.

For the **Value Function** method, only the average adjusted R^2 criterion had a defined range (0 to 1) but since the actual values of this statistic were clustered in a narrow range (0.66-0.70), it was decided to use local scaling for all five criteria. Equal weights were assigned to all five criteria. The partial and overall value functions for the eleven models are given in Table 6.10 and the overall value functions are compared in Figure 6.7.

Table 6.10 Partial and overall value functions for the eleven multivariate multiple regression models

Model	Standardised partial value functions					Value function
	MFPE	CAIC	BIC	CHQ	Average adjusted R^2	
Model 2*	55	69	70	74	5	55
Model 3*	24	45	56	60	15	41
Model 4*	39	73	78	85	23	61
Model 5*	28	63	73	80	8	51
Model 6*	34	69	76	82	10	55
Model 12*	0	59	77	85	0	44
Model 17*	77	100	100	100	20	81
Model 18*	100	86	80	75	15	72
Model 19*W	51	0	0	0	83	34
Model 20*W	83	25	15	15	98	56
Model 21*W	85	27	16	16	100	58

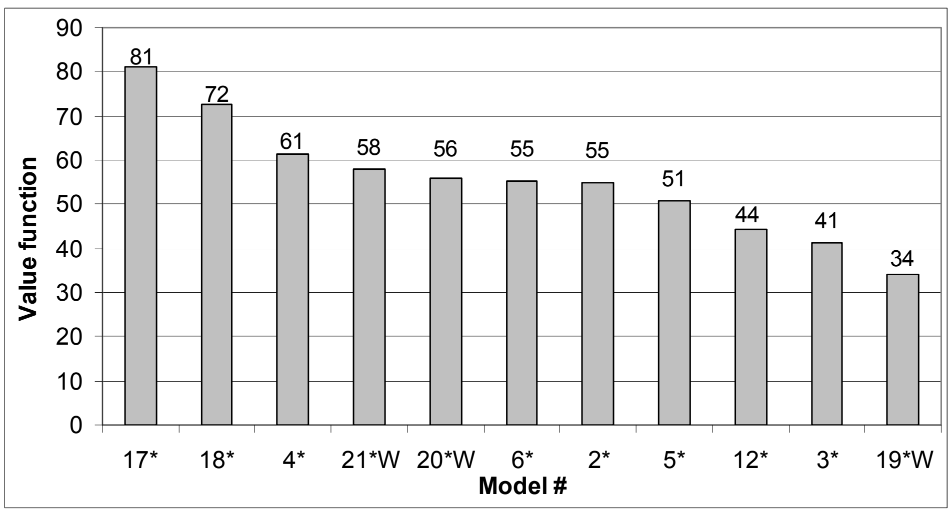


Figure 6.7 Comparison of the value function scores for the eleven multivariate multiple regression models

Model 17* obtained the highest value function score, followed by Model 18*, after which there was a considerable decrease in value function scores for the rest of the models.

The **Outranking method** was also applied. The parameters for this method were set as follows in order to obtain an outranking relationship between the models which did not contain too many incomparable relationships:

- For the concordance index, the five goodness-of-fit criteria were weighted equally, as for the value function method.
- For the discordance index, twice the standard deviation for each criterion (across the eleven models) was used as the veto threshold.
- The concordance threshold was set at C* = 0.5 (the values of the concordance index range from 0 to 1).
- The discordance threshold was set at D* = 0.1 (the discordance index can take on values of only 0 or 1).

The tabulated concordance and discordance indices and the outcome of the pairwise outranking comparisons are presented in Tables 6.11 and 6.12 respectively. The outranking relationship between the eleven multivariate multiple regression models is depicted in Figure 6.8.

Table 6.11 Concordance and discordance indices for the ELECTRE I outranking method for the eleven multivariate multiple regression models

Concordance Index (a=column, b=row)											
Model	2*	3*	4*	5*	6*	12*	17*	18*	19*W	20*W	21*W
2*	1.0	0.2	0.8	0.6	0.8	0.4	1.0	1.0	0.4	0.4	0.4
3*	0.8	1.0	1.0	0.8	0.8	0.6	1.0	1.0	0.4	0.4	0.4
4*	0.2	0.0	1.0	0.0	0.0	0.2	0.8	0.6	0.4	0.4	0.4
5*	0.4	0.2	1.0	1.0	1.0	0.4	1.0	0.8	0.4	0.4	0.4
6*	0.4	0.2	1.0	0.0	1.0	0.4	1.0	0.8	0.4	0.4	0.4
12*	0.6	0.4	0.8	0.6	0.6	1.0	1.0	0.8	0.4	0.4	0.4
17*	0.0	0.0	0.2	0.0	0.0	0.0	1.0	0.2	0.2	0.4	0.4
18*	0.0	0.2	0.4	0.2	0.2	0.2	0.8	1.0	0.2	0.2	0.2
19*W	0.8	0.6	0.6	0.6	0.6	0.6	0.8	0.8	1.0	1.0	1.0
20*W	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.8	0.0	1.0	1.0
21*W	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.8	0.0	0.0	1.0

Discordance Index (a=column, b=row)											
Model	2*	3*	4*	5*	6*	12*	17*	18*	19*W	20*W	21*W
2*	0	0	0	0	0	0	0	0	3	0	0
3*	0	0	0	0	0	0	0	0	0	0	0
4*	0	0	0	0	0	0	0	0	3	1	1
5*	0	0	0	0	0	0	0	0	3	0	0
6*	0	0	0	0	0	0	0	0	3	0	0
12*	0	0	0	0	0	0	0	0	3	1	1
17*	0	0	0	0	0	1	0	0	3	3	3
18*	0	1	0	1	1	1	0	0	3	1	1
19*W	0	0	0	0	0	1	0	0	0	0	0
20*W	1	1	0	1	1	2	0	1	0	0	0
21*W	1	1	0	1	1	2	0	1	0	0	0

Table 6.12 Outranking relation for the pairs of alternatives (multivariate multiple regression models) according to the ELECTRE I method

Outranking relation (a=column, b=row)										
Model	2*	3*	4*	5*	6*	12*	17*	18*	19*W	20*W
3*	2* > 3*									
4*	4* > 2*	4* > 3*								
5*	5* > 2*	5* > 3*	4* > 5*							
6*	6* > 2*	6* > 3*	4* > 6*	6* > 5*						
12*	2* > 12*	12* > 3*	4* > 12*	5* > 12*	6* > 12*					
17*	17* > 2*	17* > 3*	17* > 4*	17* > 5*	17* > 6*	17* > 12*				
18*	18* > 2*	18* > 3*	18* > 4*	18* > 5*	18* > 6*	18* > 12*	17* > 18*			
19*W	2* > 19*W	3* > 19*W	4* > 19*W	5* > 19*W	6* > 19*W	inc	17* > 19*W	18* > 19*W		
20*W	inc	inc	4* > 20*W	inc	inc	inc	17* > 20*W	inc	20*W > 19*W	
21*W	inc	inc	4* > 21*W	inc	inc	inc	17* > 21*W	inc	21*W > 19*W	21*W > 20*W

The symbol ">" indicates outranking and "inc" indicates an incomparable relationship.

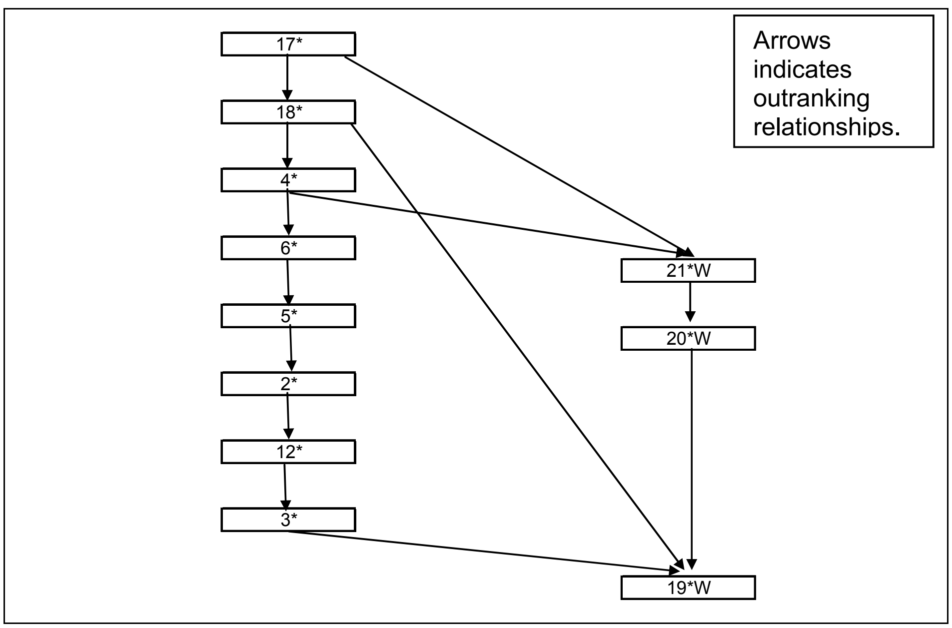


Figure 6.8 The outranking relationship for the eleven multivariate multiple regression models

The results showed that Model 17* was ranked as the best performing model, as was the case in the application of the Value Function method. The other models followed roughly the same ranking as given by the Value Function method, but as a result of a number of incomparable relationships (see Table 6.12) the ranking given by the Outranking method was not as clear. This was however not of concern since the aim of the MCDA was only to identify the best model.

6.8 Validation of the model

Due to the unbalanced nature of the data, it was not feasible to validate the model by means of a split sample or retained sample approach. However, the PRESS statistic(55,69) was used to validate the model, whereby each case is omitted in turn and predicted from the regression model derived from the remaining cases. The PRESS statistics for the top three models, as identified by MCDA in the previous section, are shown in Table 6.13.

Table 6.13 Sum of squares of PRESS residuals for the three best multivariate multiple regression models

Dependent variable:	**α**	**ln(β)**	**ln(γ)**
Model 17*	654 759	**17.18**	6.53
Model 18*	625 981	18.06	6.93
Model 4*	**614 014**	19.16	**6.48**
% difference between Model 17* and model with lowest sum of squares of PRESS residuals	6.6%	0	0.8%

The lowest value for each dependent variable is marked in bold.

The results showed that the PRESS statistics for the best model, chosen on the basis of the five goodness-of-fit criteria discussed in the previous section, are similar to those for the two models which were ranked second and third by MCDA.

6.9 Description of the best model

Firstly, the **range of applicability** of the model must be specified. The cement, stone and sand types covered by the model are limited to the cement, stone and sand types and levels included in the 165 profiles used in the final multivariate multiple regression analysis giving rise to Model 17* *and also their combinations* since interaction effects were not specified in the model. These ranges and combinations are summarised in Table 6.14.

The ranges of the continuous variables covered by the model are given in Table 6.15.

Table 6.14 Types, levels and combinations of cement, stone and sand covered by Model 17*

Stone Type (all range 0 to 1)	**Cement Type**										
	CEM I	CEM II A-D	CEM II A-L	CEM II A-M(L)	CEM II A-S	CEM II A-V	CEM II B-M (V/L)	CEM II B-S	CEM II B-V	CEM III A	CEM V A
Andesite	√		√	√	√		√	√	√	√	√
Dolerite	√	√	√		√	√		√	√	√	
Dolomite	√							√		√	
Granite	√							√	√	√	
Greywacke	√					√			√	√	
Pretoria Quartzite	√										
Quartzite	√										
Shale										√	
Tillite	√							√	√	√	
Wits Quartzite	√									√	

Sand Type (range 0 to 1 unless otherwise indicated)	**Cement Type**										
	CEM I	CEM II A-D	CEM II A-L	CEM II A-M(L)	CEM II A-S	CEM II A-V	CEM II B-M (V/L)	CEM II B-S	CEM II B-V	CEM III A	CEM V A
Andesite	√										
Cape Flats	√					√			√		
Dolerite	√	√							√	√	
Dolomite	√		√		√	√		√	√	√	
Ecca grit									√		
Granite	√		√	√			√	√	√	√	√
Klipheuwel pit	√									√	
Natural	√				√			√		√	
Pretoria Quartzite	√										
Quartzite (range 0 to 0.8)	√										
River (range 0 to 0.25)	√									√	
River Vaal (range 0 to 0.2)	√							√	√	√	
Shale										√	
Tillite (range 0 to 0.8)	√							√	√	√	
Wits Quartzite	√									√	

Table 6.14 Types, levels and combinations of cement, stone and sand covered by Model 17* (continued)

Sand Type (range 0 to 1 unless otherwise indicated)	**Stone Type** (all range 0 to 1)									
	Ande-site	Dole-rite	Dolo-mite	Gra-nite	Grey-wacke	Pretoria Quartzite	Quartz-ite	Shale	Tillite	Wits Quartzite
Andesite	√									
Cape Flats					√					
Dolerite		√								
Dolomite		√	√							
Ecca grit		√								
Granite	√			√						
Klipheuwel pit					√					
Natural	√	√								
Pretoria Quartzite						√				
Quartzite (range 0 to 0.8)							√			
River (range 0 to 0.25)										√
River Vaal (range 0 to 0.2)	√		√	√	√	√	√		√	√
Shale										√
Tillite (range 0 to 0.8)									√	
Wits Quartzite								√		√

Table 6.15 Ranges of the continuous variables covered by Model 17*

Covariate	**Range**
Cement Content	112 – 536 kg/m^3
Water Content	160 – 225 kg/m^3
Stone Content	900 – 1400 kg/m^3
Aggregate / Binder Ratio	3.18 – 8.74
2*Volume To Surface Area	33 - 150
Temperature	21 – 25 ºC

The **coefficients** of the model terms for the three dependent variables are given in Table 6.16. It is interesting to note that the ln(Cement Content) term was not significant for any of the three dependent variables, despite its multivariate significance. It can also be seen that different covariates were significant for different dependent variables.

Table 6.16 Coefficients for Model 17*

Dependent variable	α			ln(β)		
Model term	coefficient	p-value	standardised coefficient	coefficient	p-value	standardised coefficient
Intercept	**-2245.19**	**<0.001**		**9.76**	**<0.001**	
CEM III A GGBS	8.63	0.678	0.03	**-0.29**	**0.007**	-0.21
CEM V A	-3.85	0.937	-0.00	**0.79**	**0.002**	0.16
Stone Granite	-43.02	0.117	-0.08	**0.33**	**0.017**	0.15
Stone Tillite	**211.75**	**<0.001**	0.29	-0.32	0.050	-0.10
Stone Quartzite	**302.21**	**<0.001**	0.27	-0.00	0.997	-0.00
Sand Wits Quartzite	**260.15**	**<0.001**	0.53	**-0.64**	**<0.001**	-0.29
Sand Shale	**321.77**	**<0.001**	0.30	**-0.49**	**0.034**	-0.10
Sand Granite	**201.75**	**<0.001**	0.55	**-0.36**	**0.003**	-0.22
Sand Dolerite	134.27	0.089	0.17	**1.22**	**0.003**	0.34
Sand Andesite	**99.54**	**0.002**	0.13	0.04	0.786	0.01
Sand Klipheuwel pit	**139.92**	**<0.001**	0.43	**0.43**	**<0.001**	0.30
Sand natural	**170.57**	**<0.001**	0.54	**-0.42**	**<0.001**	-0.30
Sand River Vaal	**431.08**	**<0.001**	0.25	-0.35	0.494	-0.05
ln(Cement Content)	55.71	0.052	0.17	0.01	0.919	0.01
Water Content	**2.54**	**<0.001**	0.27	0.00	0.662	0.03
Stone Content	-0.05	0.439	-0.03	**-0.00**	**0.001**	-0.16
Aggregate/Binder Ratio	**25.35**	**<0.001**	0.33	-0.05	0.092	-0.14
ln(2*Volume To Surface Area)	**173.17**	**0.010**	0.33	**-1.76**	**<0.001**	-0.75
Temperature	**44.34**	**<0.001**	0.44	**-0.26**	**<0.001**	-0.57

Table 6.16 Coefficients for Model 17* (continued)

Dependent variable	ln(γ)		
Model term	coefficient	p-value	standardised coefficient
Intercept	3.04	0.057	
CEM III A GGBS	-0.01	0.909	-0.04
CEM V A	0.19	0.223	0.07
Stone Granite	**0.34**	**<0.001**	**0.27**
Stone Tillite	0.05	0.607	0.04
Stone Quartzite	0.01	0.934	0.01
Sand Wits Quartzite	**-0.44**	**<0.001**	**-0.31**
Sand Shale	**-0.40**	**0.006**	**-0.14**
Sand Granite	**-0.36**	**<0.001**	**-0.39**
Sand Dolerite	**0.50**	**0.046**	**0.34**
Sand Andesite	-0.03	0.734	-0.02
Sand Klipheuwel pit	0.02	0.694	0.00
Sand natural	-0.12	0.055	**-0.20**
Sand River Vaal	**-2.58**	**<0.001**	**-0.69**
ln(Cement Content)	0.16	0.074	0.24
Water Content	0.00	0.275	0.06
Stone Content	**-0.00**	**<0.001**	**-0.22**
Aggregate/Binder Ratio	**0.08**	**<0.001**	**0.44**
ln(2*Volume To Surface Area)	**-0.62**	**0.004**	**-0.62**
Temperature	**-0.08**	**0.004**	**-0.39**

Significant terms (5% confidence level) are marked in bold.

The final step in the description of the model was to define the 95% confidence interval for the predicted shrinkage-time profile, derived from the two-stage model. The multivariate multiple regression analysis provided standard errors for the predicted values of α, ln(β) and ln(γ). (Alternatively these can be calculated as shown in Section 6.1. The values of $(Z'Z)^{-1}$ and

$\hat{\Sigma}$ required for these calculations are given in Appendix P.) The standard errors for the predicted values of α, ln(β) and ln(γ) had to be used to derive a standard error for each prediction of *shrinkage* when the growth curve model

$$Shrinkage(t-t_0)=\alpha\left[1-e^{-\beta(t-t_0)}\right]^{\gamma}, \quad \alpha,\beta,\gamma>0$$

was applied at different values of drying time, t-t_0. This problem can be formulated as follows:

$$Shrinkage(t-t_0)=f(\hat{\alpha},\hat{\beta},\hat{\gamma},t-t_0)+e$$
$$Var(Shrinkage(t-t_0))=Var\left[f(\hat{\alpha},\hat{\beta},\hat{\gamma},t-t_0)\right]+Var[e]$$

Approximate 95% confidence limits for a predicted shrinkage value at drying time, t-t_0 are then obtained via the Normal distribution:

$$Shrinkage(t-t_0)\pm 1.96*\sqrt{Var(Shrinkage(t-t_0)}\,.$$

Two methods were used to derive the $Var(Shrinkage(t-t_0)$ term, viz.

- the Delta Method and
- simulation.

The **Delta Method** is described by Weisberg[(74)]. Let $\underset{(1\times m)}{X}$ be a vector for which we have the variance-covariance matrix, $\underset{(m\times m)}{Var(X)}$. Let $\underset{(1\times n)}{G(X)}$ be a vector function for which we want to calculate $\underset{(n\times n)}{Var(G(X))}$. Then, the approximate relationship holds:

$$Var(G(X))\approx G'(X)Var(X)G'(X)^T$$

where $\underset{(n\times m)}{G'(X)}$ is the matrix of first derivatives of $G(X)$.

The method was first applied to derive the variances and covariances of the estimates $\hat{\beta}$ and $\hat{\gamma}$ from the known variances and covariances of the estimates $ln(\hat{\beta})$ and $ln(\hat{\gamma})$. Setting $X=\begin{bmatrix}\hat{\alpha}\\ ln(\hat{\beta})\\ ln(\hat{\gamma})\end{bmatrix}$ and $G(X)=\begin{bmatrix}\hat{\alpha}\\ e^{ln(\hat{\beta})}=\hat{\beta}\\ e^{ln(\hat{\gamma})}=\hat{\gamma}\end{bmatrix}$,

we obtain

$$Var\begin{bmatrix}\hat{\alpha}\\ \hat{\beta}\\ \hat{\gamma}\end{bmatrix}=\begin{bmatrix}Var(\hat{\alpha}) & \hat{\beta}\cdot Cov(\hat{\alpha},ln(\hat{\beta})) & \hat{\gamma}\cdot Cov(\hat{\alpha},ln(\hat{\gamma}))\\ \hat{\beta}\cdot Cov(\hat{\alpha},ln(\hat{\beta})) & \hat{\beta}^2\cdot Var(ln(\hat{\beta})) & \hat{\beta}\hat{\gamma}\cdot Cov(ln(\hat{\beta}),ln(\hat{\gamma}))\\ \hat{\gamma}\cdot Cov(\hat{\alpha},ln(\hat{\gamma})) & \hat{\beta}\hat{\gamma}\cdot Cov(ln(\hat{\beta}),ln(\hat{\gamma})) & \hat{\gamma}^2\cdot Var(ln(\hat{\gamma}))\end{bmatrix}$$

The Delta Method was then applied again to derive the variance for the estimate of the shrinkage at time t-t_0, i.e. $Var\left[f(\hat{\alpha},\hat{\beta},\hat{\gamma},t-t_0)\right]$. Here,

$$X=\begin{bmatrix}\hat{\alpha}\\ \hat{\beta}\\ \hat{\gamma}\end{bmatrix},$$

$$Var(X)=\begin{bmatrix}Var(\hat{\alpha}) & Cov(\hat{\alpha},\hat{\beta}) & Cov(\hat{\alpha},\hat{\gamma})\\ Cov(\hat{\alpha},\hat{\beta}) & Var(\hat{\beta}) & Cov(\hat{\beta},\hat{\gamma})\\ Cov(\hat{\alpha},\hat{\gamma}) & Cov(\hat{\beta},\hat{\gamma}) & Var(\hat{\gamma})\end{bmatrix}$$ and

$$G(X)=\hat{\alpha}\left[1-e^{-\hat{\beta}(t-t_0)}\right]^{\hat{\gamma}}.$$

The approximate variance for the estimate of the shrinkage was then derived as

$$\begin{aligned}Var\left[G(X)\right]\approx\ & Var(\hat{\alpha})\left[\frac{\partial G(X)}{\partial\hat{\alpha}}\right]^2+Var(\hat{\beta})\left[\frac{\partial G(X)}{\partial\hat{\beta}}\right]^2+Var(\hat{\gamma})\left[\frac{\partial G(X)}{\partial\hat{\gamma}}\right]^2\\ & +2Cov(\hat{\alpha},\hat{\beta})\frac{\partial G(X)}{\partial\hat{\alpha}}\frac{\partial G(X)}{\partial\hat{\beta}}+2Cov(\hat{\alpha},\hat{\gamma})\frac{\partial G(X)}{\partial\hat{\alpha}}\frac{\partial G(X)}{\partial\hat{\gamma}}\\ & +2Cov(\hat{\beta},\hat{\gamma})\frac{\partial G(X)}{\partial\hat{\beta}}\frac{\partial G(X)}{\partial\hat{\gamma}}\end{aligned}$$

$$\begin{aligned}=\ & Var(\hat{\alpha})\left[1-e^{-\hat{\beta}(t-t_0)}\right]^{2\hat{\gamma}}\\ & +\hat{\beta}^2\cdot Var(ln(\hat{\beta}))\cdot\left[\hat{\alpha}\hat{\gamma}(t-t_0)\left[e^{-\hat{\beta}(t-t_0)}\right]\left[1-e^{-\hat{\beta}(t-t_0)}\right]^{\hat{\gamma}-1}\right]^2\\ & +\hat{\gamma}^2\cdot Var(ln(\hat{\gamma}))\cdot\left[\hat{\alpha}\left[1-e^{-\hat{\beta}(t-t_0)}\right]^{\hat{\gamma}}ln\left[1-e^{-\hat{\beta}(t-t_0)}\right]\right]^2\\ & +2\hat{\beta}\cdot Cov(\hat{\alpha},ln(\hat{\beta}))\cdot\hat{\gamma}(t-t_0)\left[e^{-\hat{\beta}(t-t_0)}\right]\left[1-e^{-\hat{\beta}(t-t_0)}\right]^{\hat{\gamma}-1}\\ & +2\hat{\gamma}\cdot Cov(\hat{\alpha},ln(\hat{\gamma}))\cdot\left[1-e^{-\hat{\beta}(t-t_0)}\right]^{\hat{\gamma}}ln\left[1-e^{-\hat{\beta}(t-t_0)}\right]\\ & +2\hat{\beta}\hat{\gamma}\cdot Cov(ln(\hat{\beta}),ln(\hat{\gamma}))\cdot\hat{\alpha}(t-t_0)\left[e^{-\hat{\beta}(t-t_0)}\right]\left[1-e^{-\hat{\beta}(t-t_0)}\right]^{\hat{\gamma}-1}\left\{1+g\cdot ln\left[1-e^{-\hat{\beta}(t-t_0)}\right]\right\}\end{aligned}$$

$Var[e]$ was calculated from the pooled residual variance of the 206 fitted growth curve models (see section 5.7):

$$Var[e] = \frac{\sum_{i=1}^{206} SSE}{\sum_{i=1}^{206} df} = 208.9$$

where *df* refers to the degrees of freedom.

The **simulation method** is described by Chernick[75]. For a particular shrinkage prediction at a given drying time t-t_0, values of α, ln(β) and ln(γ) were sampled from a correlated multivariate Normal distribution, after which the values of β and γ were calculated. Finally, a value of e was sampled from a Normal $(0, Var(e) = 208.9)$ distribution. These values were used to calculate an estimate of the shrinkage at the given drying time:

$$Shrinkage(t - t_0) = f(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, t - t_0) + e\,.$$

This process was repeated 10,000 times. The mean of these values was set equal to the predicted shrinkage at the given drying time. The 2.5th and 97.5th percentiles of the distribution of these values formed the 95% confidence limits of the predicted shrinkage values at the given drying time. The method was repeated for each drying time t-t_0 at which a shrinkage prediction was required.

The simulation of α, ln(β) and ln(γ) from a correlated multivariate Normal distribution was carried out as follows[76]: For each of the 10,000 samples (see above), the vector $X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$ was simulated from $X_i \sim N(0,1),\ \ i = 1,2,3$. Simulated values for the three dependent variables were then obtained by

$$Y = \begin{bmatrix} \alpha \\ ln(\beta) \\ ln(\gamma) \end{bmatrix} = AX + E(Y)$$

where A is the square root matrix of the model's variance-covariance matrix Σ (see Appendix P for values) and is obtained by a Cholesky decomposition of Σ and where E(Y) is the vector of the predicted values of α, ln(β) and ln(γ) for the shrinkage profile in question.

The predicted shrinkage profile and its 95% confidence interval for one experiment, calculated by both the Delta Method and by simulation, is illustrated in Figure 6.9. The confidence intervals yielded by the two methods were very similar, indicating that both methods were suitable for the calculation of the confidence intervals. The Delta Method was used henceforward for the calculation of the confidence intervals for all the shrinkage profiles in the data set.

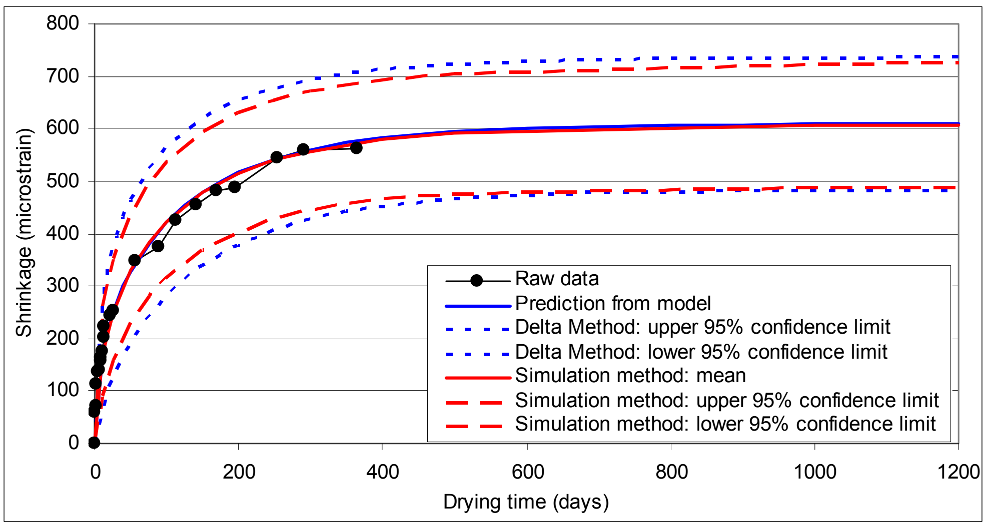


Figure 6.9 Mean and 95% confidence interval predictions for a typical shrinkage profile (# 003) by the Delta method and by simulation

The predicted shrinkage profiles, together with their 95% confidence intervals, for

- the 165 cases (experiments) in the data set used to derive the multivariate multiple regression model (Model 17*) and
- a further 47 experiments in the overall data set of 245 profiles which were not used to derive Model 17*, but which qualified to be predicted by the model

are shown in the file "Predicted shrinkage profiles.xls" contained in the enclosed CD. To give an overview of the fit of the model to the raw data, the two shrinkage profiles with predicted values of α closest to the observed asymptote as well as the two shrinkage profiles with predicted values of α furthest from the observed asymptote, from the 165 cases in the multivariate multiple regression data set, are shown in Figure 6.10. **Model 17* was henceforward known as the WITS Model.**

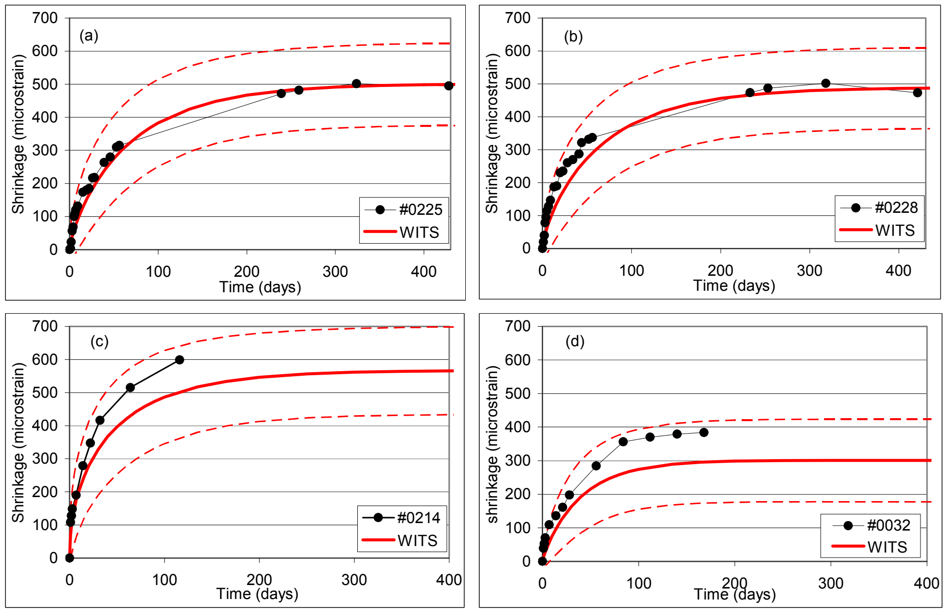


Figure 6.10 Mean and 95% confidence interval predictions for the two shrinkage profiles with predicted values of α ((a) and (b)) closest to the observed asymptote and ((c) and (d)) furthest from the observed asymptote

The American Concrete Institute notes that the variability of shrinkage test measurements prevents models from matching measured data closely and it is thus unrealistic to expect results from shrinkage prediction models to be within less than 20% of the test data[3]. In this case, the two predicted profiles exhibiting the *worst* deviations from the raw data (see Figure 6.10 (c) and (d)) had deviations of 16.2 and 22.9%, respectively, from the last raw data point, which confirmed that the model was indeed acceptable. Comparisons between the WITS model and previously published models are presented in Chapter 7.

6.10 Interpretation of the model

The standardised coefficients (see Table 6.16) were used to establish which covariates had the most significant effect on each dependent variable. They are listed in descending order of absolute magnitude of the β-coefficient in Table 6.17.

Table 6.17 The ranking of the significant terms for each of the three dependent variables for the multivariate multiple regression model according to the absolute magnitude of the standardised coefficient

α	ln(β)	ln(γ)
Sand Granite (+)	ln(2*Volume to Surface Area) (-)	Sand River Vaal (-)
Sand natural (+)	Temperature (-)	ln(2*Volume to Surface Area) (-)
Sand Wits Quartzite) (+)	Sand Dolerite (+)	Aggregate/Binder Ratio (+)
Temperature (+)	Sand Natural (-)	Sand Granite (-)
Sand Klipheuwel Pit (+)	Sand Klipheuwel pit (+)	Sand Wits Quartzite (-)
Aggregate/Binder Ratio (+)	Sand Wits Quartzite (-)	Temperature (-)
ln(2*Volume to Surface Area)(+)	Sand Granite (-)	Stone Granite (+)
Sand Shale (+)	Cement CEM III A GGBS (-)	Sand Dolerite (+)
Stone Tillite (+)	Cement CEM V A (+)	Stone Content (-)
Stone Quartzite (+)	Stone Content (-)	Sand Shale (-)
Water Content (+)	Stone Granite (+)	
Sand River Vaal (+)	Sand Shale (-)	
Sand Andesite (+)		

The sign of the coefficient is indicated. Colour coding is used for ease of interpretation.

The interpretation of the model terms should be read in conjunction with the definitions, background and present state of knowledge regarding the factors known to influence the shrinkage of concrete given in Chapter 2. We first consider the dependent variable **α**, which represents the **ultimate shrinkage**.

- The different sand types feature very strongly in the regression equation for this dependent variable, with three sand types (granite, natural and Wits Quartzite) making the largest contribution to high values of α. All seven sand types which were found to be significant had positive coefficients relative to the reference sand type, dolomite, which is generally accepted to be the least shrinking aggregate of those covered by this study[67].
- The next most important term was temperature. As expected, a higher temperature leads to a higher ultimate shrinkage.

- The effect of the aggregate / binder ratio may be understood in terms of the diluting effect of the aggregate.
- The specimen size effect (as represented by the ratio of the volume to the surface area of the specimen) is unexpected: the positive sign of the coefficient suggests that specimens with a lower surface area available for moisture loss, relative to their volume, are expected to reach higher levels of ultimate shrinkage. It must be noted that this variable was subject to some multicollinearity (see Section 6.5) (mostly with sand type dolerite and temperature) and thus the meaning of the magnitude and sign of the coefficient should not be over-interpreted.
- The two stone types which were found to be significant had positive coefficients relative to the reference stone type, andesite.
- Finally, the water content of the concrete also plays a role in influencing ultimate shrinkage. As expected, increasing the water content of the concrete mix increases the ultimate shrinkage.

As mentioned in Chapter 2, much less is known about the factors which influence the rate of shrinkage. However, the significant terms in the regression equation for **ln(β)**, which represents the **rate of shrinkage development with time**, may be interpreted as follows:

- As expected, the specimen size effect is the strongest contributor to the rate of shrinkage development with time: specimens with a higher surface area available for moisture loss, relative to their volume, lose moisture more rapidly.
- The rate of shrinkage decreases with increasing temperature. This is unexpected, but is likely to be linked to the very narrow range of temperatures covered by this study (21-25ºC) (see further discussion in Chapter 8).
- Sand type plays an important role in determining the rate of shrinkage. In contrast to the findings for the ultimate shrinkage, here the signs of coefficients for the different sand types are both positive and negative.
- A few cement types have an effect of the rate of shrinkage. Cements with high levels (36-65%) of GGBS (i.e. CEM III A) appeared to slow the rate of shrinkage. The effect was not statistically significant for comparable concretes containing Corex or Ferro-manganese slag.

Cements containing both GGBS (18-30%) and fly ash (18-30%) (i.e. CEM V A) appeared to increase the rate of shrinkage.

- Increasing stone content decreases the rate of shrinkage, presumably due to the diluting effect of the aggregate. One stone type, granite, was found to increase the rate of shrinkage relative to the reference stone type, andesite.

The **shape parameter, ln(γ)** is itself difficult to interpret in the context of the growth curve equation, and thus its interpretation in terms of the regression terms is even more difficult and was not attempted.

The effects of the significant regression terms discussed above are illustrated graphically in Figures 6.11 to 6.18 below. For each numerical covariate, the extremes and the midpoint of the range covered by the model are illustrated (keeping the other covariates at fixed levels).

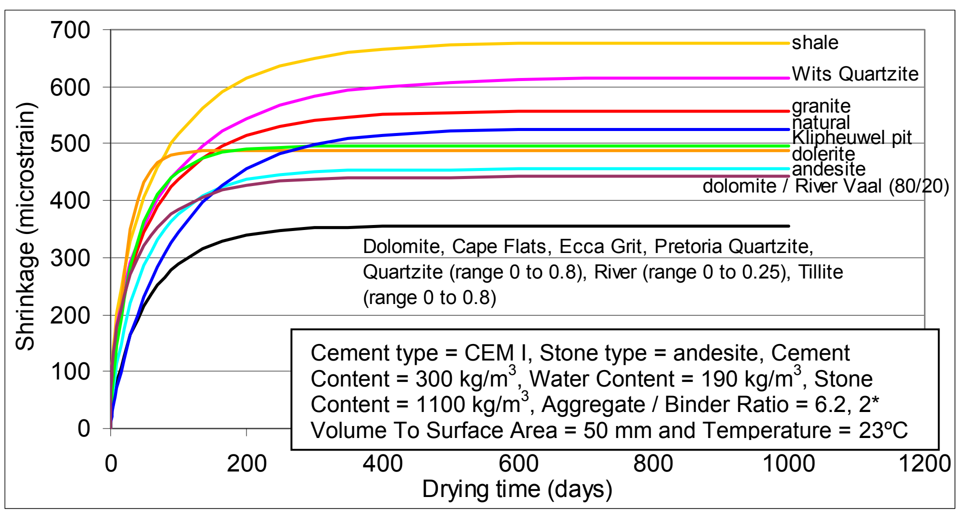


Figure 6.11 The effect of sand type on the predicted shrinkage profile

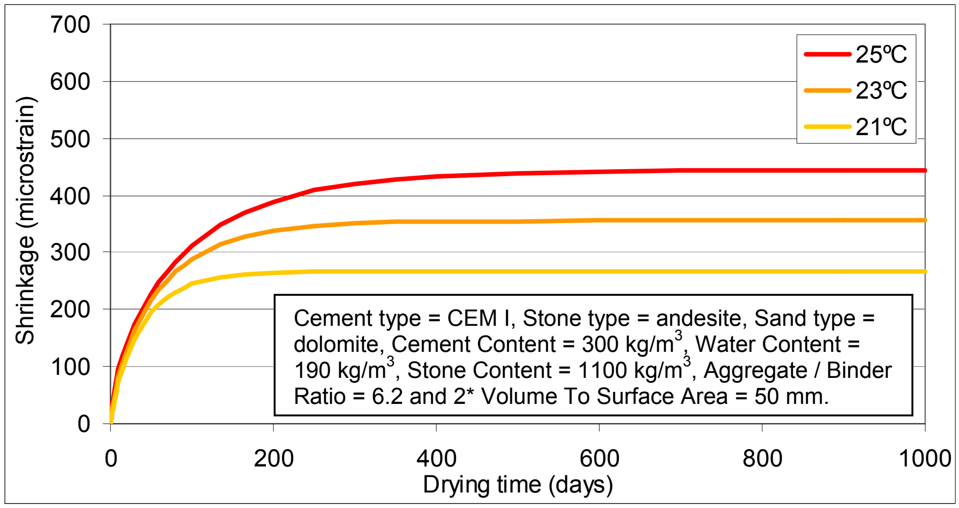


Figure 6.12 The effect of temperature on the predicted shrinkage profile

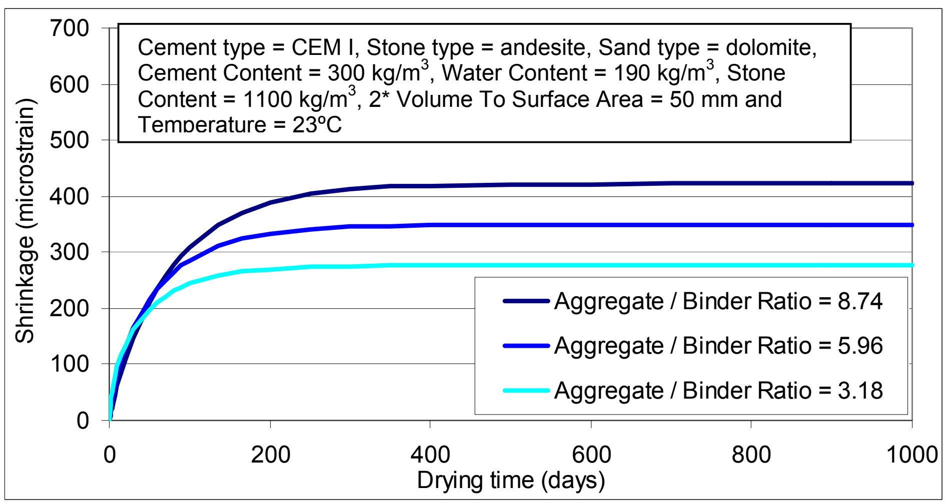


Figure 6.13 The effect of the aggregate / binder ratio on the predicted shrinkage profile

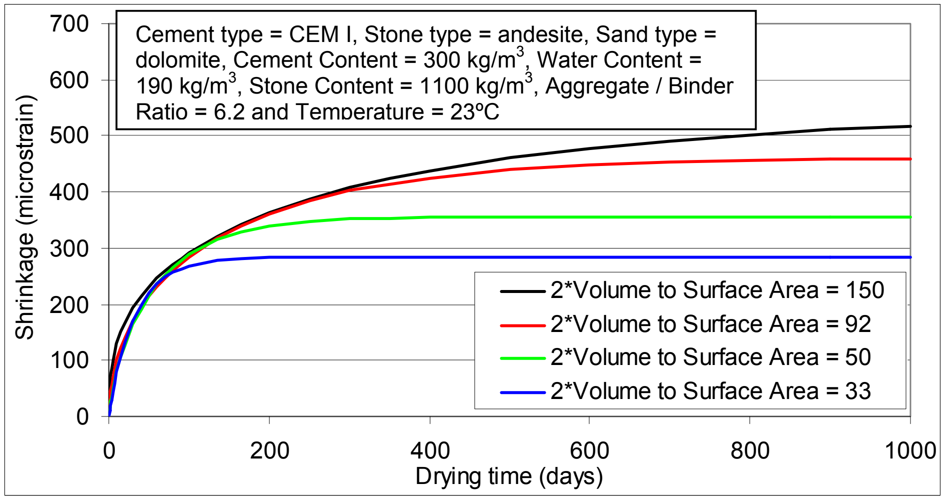


Figure 6.14 The effect of the specimen volume to surface area ratio on the predicted shrinkage profile

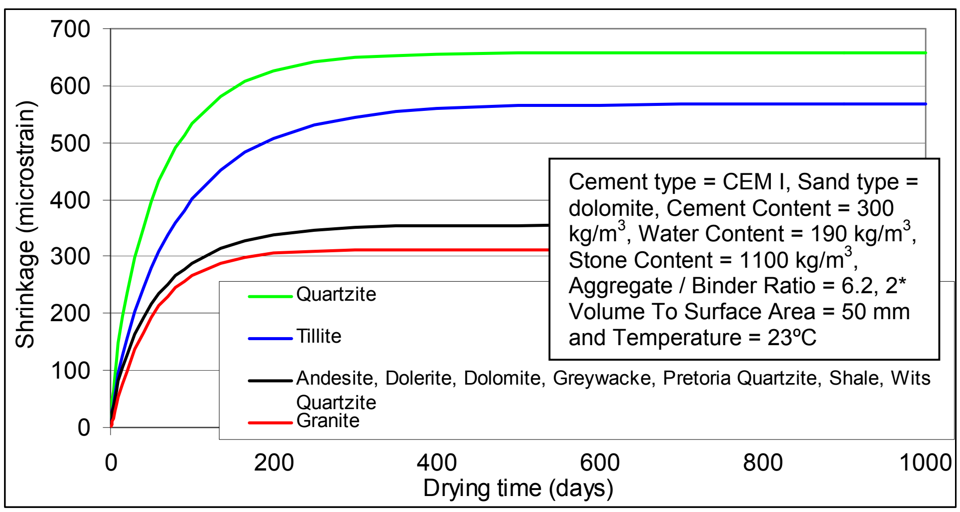


Figure 6.15 The effect of the stone type on the predicted shrinkage profile

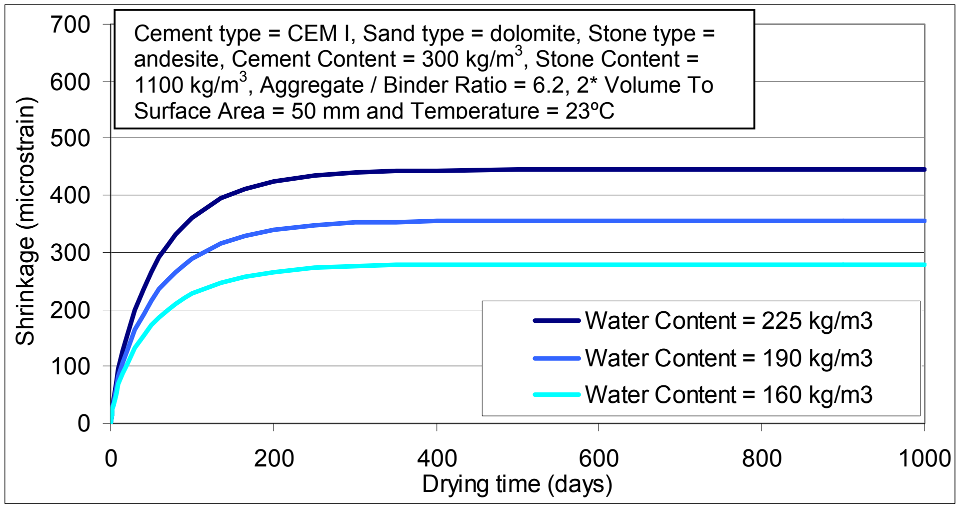


Figure 6.16 The effect of water content on the predicted shrinkage profile

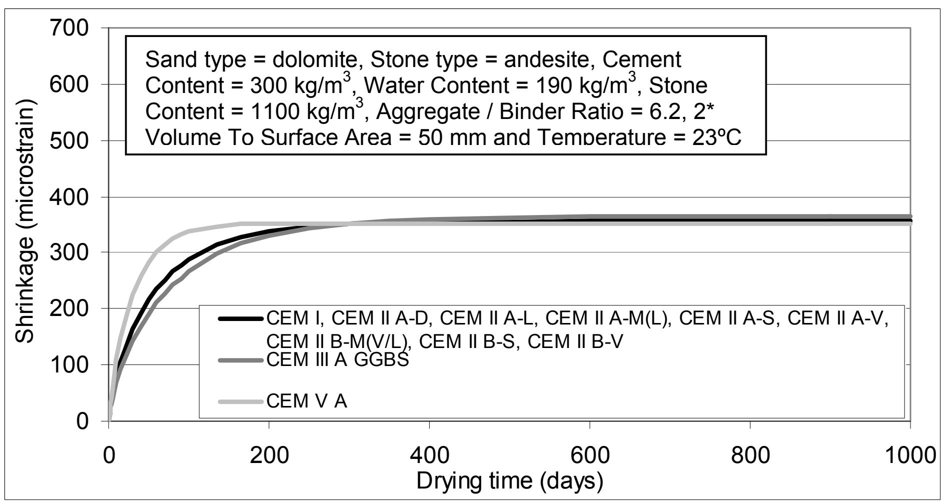


Figure 6.17 The effect of cement type on the predicted shrinkage profile

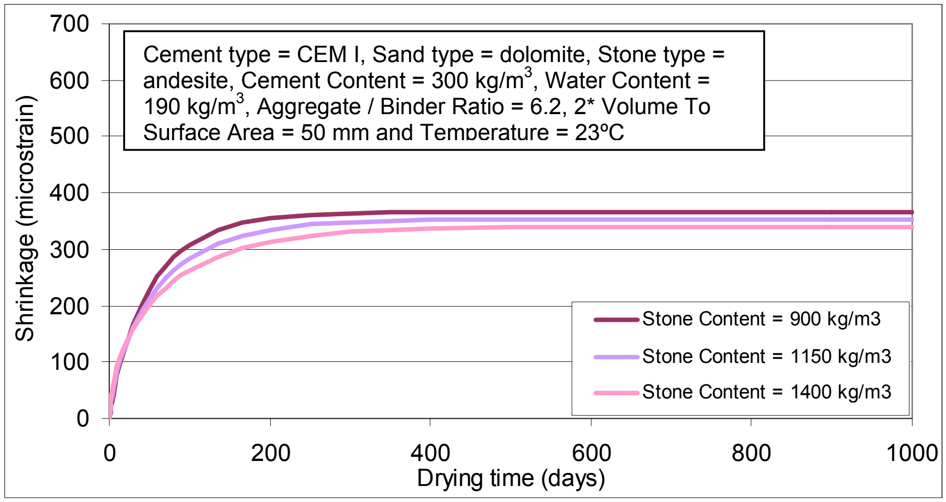


Figure 6.18 The effect of stone content on the predicted shrinkage profile

In a further attempt to interpret the role of the sand and stone types in the model, their regression model coefficients (for the three dependent variables α, ln(β) and ln(γ)), were plotted against all the numerical descriptors available for the stone and sand types (such as relative density, elastic modulus, etc. - see Table 3.2) as well as the water demand of several sand types which became available later on in the project[77]. No meaningful correlations were found, for several reasons:

- There were only one or two stone types with non-zero coefficients in the regression equation (depending on the dependent variable under consideration) while there were eight or nine stone types with a coefficient of zero. Further, the significance of the Stone Content variable (at least for ln(β) and ln(γ)) reduced the impetus for finding a numeric descriptor to interpret the role of the stone type in the regression equation.
- While eight of the fifteen sand types in the model had non-zero coefficients for at least one dependent variable, there were still too many sand types which all had a coefficient of zero to achieve meaningful relationships when plotted against the numeric descriptor variables associated with the sand characteristics.
- Many of the sand descriptor variables suffered from missing data, as has been discussed previously.

Imamoto[78], in a study on Japanese aggregates, found that the shrinkage of concrete was proportional to the surface area of the aggregates. However, in this case, no surface area data were available to check whether this was true for the South African aggregates used in this study.

6.11 Sensitivity of the model predictions to small changes in input parameters

As mentioned in Chapter 4, one of the requirements for a good shrinkage model is that it should not be highly sensitive to small changes in input parameters. To test this requirement, the six continuous variables in the model were, in turn, varied by 5% from their baseline values (taken as their mean value in the data set) and the resulting percentage changes from baseline in the predicted values of α, $\ln(\beta)$ and $\ln(\gamma)$ were recorded. The results are shown in Table 6.18.

The results show that the model was most sensitive to a 5% change in temperature. A 5% change in temperature from the average temperature in the regression data set (22.8ºC) spanned the range 21.7 – 23.9 ºC which covered a large part of the temperature range covered by the data set (21-25 ºC). Temperature is known to have a strong influence on shrinkage (see Figure 6.12). Thus, this was not a cause for concern. Changes in the stone content level had a large influence on the values of $\ln(\gamma)$, but as can be seen from Figure 6.18, which covers a much wider range of stone content, this had very little effect on the overall shrinkage profile. Thus it was concluded that the model is not overly sensitive to small changes in the input parameters.

Table 6.18 The percentage change in the predicted values of α, ln(β) and ln(γ) resulting from 5% changes in the values of the continuous input parameters to the model

Percentage change in input parameters						Percentage change in predicted values		
Cement Content (kg/m^3)	**Water Content (kg/m^3)**	**Stone Content (kg/m^3)**	**Aggregate / Binder Ratio**	**2* Vol. To Surface Area (mm)**	**Tempe-rature (°C)**	**α**	**ln(β)**	**ln(γ)**
0	0	0	0	0	0	0	0	0
+5%	0	0	0	0	0	+0.9%	0%	-2.4%
-5%	0	0	0	0	0	-0.9%	0%	+2.5%
0	+5%	0	0	0	0	+7.5%	-0.3*	-5.7%
0	-5%	0	0	0	0	-7.5%	+0.3*	+5.7%
0	0	+5%	0	0	0	-0.9%	+1.5%	**14.4%**
0	0	-5%	0	0	0	+0.9%	-1.5%	**-14.4%**
0	0	0	+5%	0	0	+2.3%	+0.4%	-7.0%
0	0	0	-5%	0	0	-2.3%	-0.4%	+7.0%
0	0	0	0	+5%	0	+2.6%	+2.1%	+9.1%
0	0	0	0	-5%	0	-2.7%	-2.2%	-9.6%
0	0	0	0	0	+5%	**+15.6%**	+7.2%	**27.3%**
0	0	0	0	0	-5%	**-15.6%**	-7.2%	**-27.3%**

Changes in excess of 10% are marked in bold.

In this chapter, the second step in the hierarchical non-linear modelling process was completed: a multivariate multiple regression model was identified for modelling the relationship between the parameters of the growth curve model and the covariates describing the concrete raw materials, the concrete mix and the shrinkage testing conditions. Preparation of both the dependent variables and the covariates for regression analysis was discussed, as well as the grouping of covariates into meaningful groups for variable selection in the multiple regression analysis while avoiding multicollinearity. The results of the multivariate multiple regressions were presented and different weighting schemes for the cases (experiments) were explored. MCDA was used to identify the best regression model from a number of candidates, using five goodness-of-fit measures. The chosen model was then described in terms of its range of applicability and significant terms and a 95% confidence interval for the predicted shrinkage profiles was developed. The model was interpreted in terms of what is already known about the factors affecting concrete shrinkage. Finally, a sensitivity analysis was performed to investigate the effect of changes in the input parameters on the predicted growth curve parameters.

In the next chapter, the model developed in this study will be compared to the previously published models introduced in Chapter 2. The ability of all the models to predict the shrinkage profiles in the data set will be compared by a variety of summary statistics. These summary statistics will also be critically assessed.

7 Comparison of the WITS model to other published models for concrete shrinkage

In this chapter, the other published models for concrete shrinkage, first highlighted in Chapter 2, are presented in detail. These models are then compared to the WITS model, regarding their predictive ability with respect to the South African data set used in this study. A wide variety of graphical and numerical goodness-of-fit measures are used to assess the relative performance of the models.

7.1 Overview of the published models for concrete shrinkage used as comparisons to the WITS model

As outlined in Chapter 2, five published models were used as comparisons to the WITS model developed in this study:

- The ACI 209R-92 model developed by the American Concrete Institute[9]
- The RILEM B3 model developed by Bažant and co-workers[10,11]
- The CEB MC90-99 model developed by the Comité European du Beton[12]
- The GL2000 model developed by Gardner and co-workers[13]
- The SANS 10100-1 model adopted by the South African Bureau of Standards[14]

The background to these models was given in Chapter 2. Here, the models are described more fully. The covariates (other than drying time) used in each of these models are given in Table 7.1 and the ranges of applicability of each model are summarised in Table 7.2. The corresponding information for the WITS model is included for comparison. The ranges of applicability of the WITS model given in Table 7.2 are equivalent to the ranges of the data used in fitting the model; the actual ranges of applicability could well be wider.

Table 7.1 Covariates included in the published models as well as the WITS model

Covariates	Model					
	ACI 209R-92	**RILEM B3**	**CEB MC90-99**	**GL2000**	**SABS 0100**	**WITS**
Concrete raw materials and composition:						
Cement type	*	√	√	√		√
Cement content	√	*				√
Water content		√			√	√
Water / cement mass ratio		*				
Air content	√					
Sand type						√
Stone type						√
Stone content						√
Sand / total aggregate mass ratio	√					
Aggregate / cement mass ratio		*				
Aggregate / binder mass ratio						√
Testing conditions:						
Curing method	*	√	*			
Age at first drying	√	√	*			
Specimen shape		√				
Specimen volume to surface area ratio	√	√	√	√		√
Specimen ratio of cross-sectional area to exposed perimeter					√	
Temperature	*		*			√
Humidity	√	√	√	√	√	
Concrete properties:						
28-day compressive strength		√	√	√		
28-day elastic modulus		√				
Slump	√					

The symbol √ denotes that the covariate is required for model prediction calculations, while the symbol * denotes that the covariate is only required to assess the applicability of the model.

Table 7.2 Ranges of applicability for the published models as well as the WITS model

Constraints	Model					
	ACI 209R-92	**RILEM B3**	**CEB MC90-99**	**GL2000**	**SABS 0100**	**WITS**
Concrete raw materials and composition:						
Cement type	Type I and III	Type I, II and III				see Table 6.14
Cement content	279-446 kg/m^3	160-720 kg/m^3				112- 536 kg/m^3
Water content					150- 230 kg/m^3	160- 225 kg/m^3
Water / cement mass ratio		0.35-0.85				
Aggregate / cement mass ratio		2.5-13.5				
Aggregate / binder mass ratio						3.18- 8.74
Sand type						see Table 6.14
Stone type						
Stone content						900-1400 kg/m^3
Testing conditions:						
Curing method and time	moist: ≥ 1 day or steam: 1-3 days	moist: ≥ 1 day or steam	moist ≤ 14 days	moist: ≥ 1 day or steam		
Specimen volume to surface area ratio	1.2*exp (-0.00472*V/S) ≥ 0.2					16.5-75
Temperature	21.2-25.2 ºC		10-30 ºC			21-25 ºC
Humidity	40-100%	40-100%	40-100%	20-100%	20-100%	(43-72%)
Concrete properties:						
28-day (cylinder) compressive strength		17-70 MPa	15-120 MPa	16-82 MPa		

The specifications of the five models are shown in Figures 7.1-7.5. The specification of the WITS model is given in Figure 7.6 for comparison. Brief notes on some of the models follow.

According to Bažant[18,45] the hyperbolic tangent curve was chosen for the **RILEM B3** model to fit the development of shrinkage with time because it satisfies the requirements of diffusion theory at both short and very long drying times:

- at short drying times ($t-t_0 << \tau_{sh}$), $\tanh(x) \approx x$, i.e. shrinkage $\propto \sqrt{\dfrac{t-t_0}{\tau_{sh}}}$ as required by diffusion theory.
- at long drying times ($t-t_0 >> \tau_{sh}$), $1-\tanh(x) \approx 2e^{-2x}$, i.e. the difference between the shrinkage and the asymptotic value is approximately $2\exp\left[-2\sqrt{\dfrac{t-t_0}{\tau_{sh}}}\right]$, which agrees with diffusion theory.

(t = age of the concrete in days; t_0 = age at first drying in days; τ_{sh} = half-time of shrinkage in days, i.e. the time required for the shrinkage to reach half of its ultimate value.)

The **CEB MC90-99** model separates the specification of drying and autogenous shrinkage. Autogenous shrinkage results from the chemical reaction between the cement and water when the concrete is mixed(32), and is considered to make a small contribution to overall concrete shrinkage(44).

The covariates used in the published models included some which were not well represented in the data set used to derive the WITS model (see Table 7.1). This resulted from the fact that the shrinkage measurements in the data set used in this study were mostly carried out on standard specimen shapes and dimensions, under standard laboratory conditions. Given their use in the published models, it is clear that these covariates may be of potential importance in any shrinkage model, including future versions of the WITS model. The following covariates were used in one or more of the published models, but could not be included in the WITS model, as a result of high levels of missing data or the limited range covered by the covariate:

- Curing method: 95% of data set comprised one curing method.
- Age at first drying: The range was adequate at 7-49 days but the covariate was found to be non-significant in the final model.
- Specimen shape: This was not considered as a separate variable since it was confounded with the Laboratory covariate.
- Humidity: The extremes of the range, 43-72%, were possibly inadequately represented for this covariate to be significant.

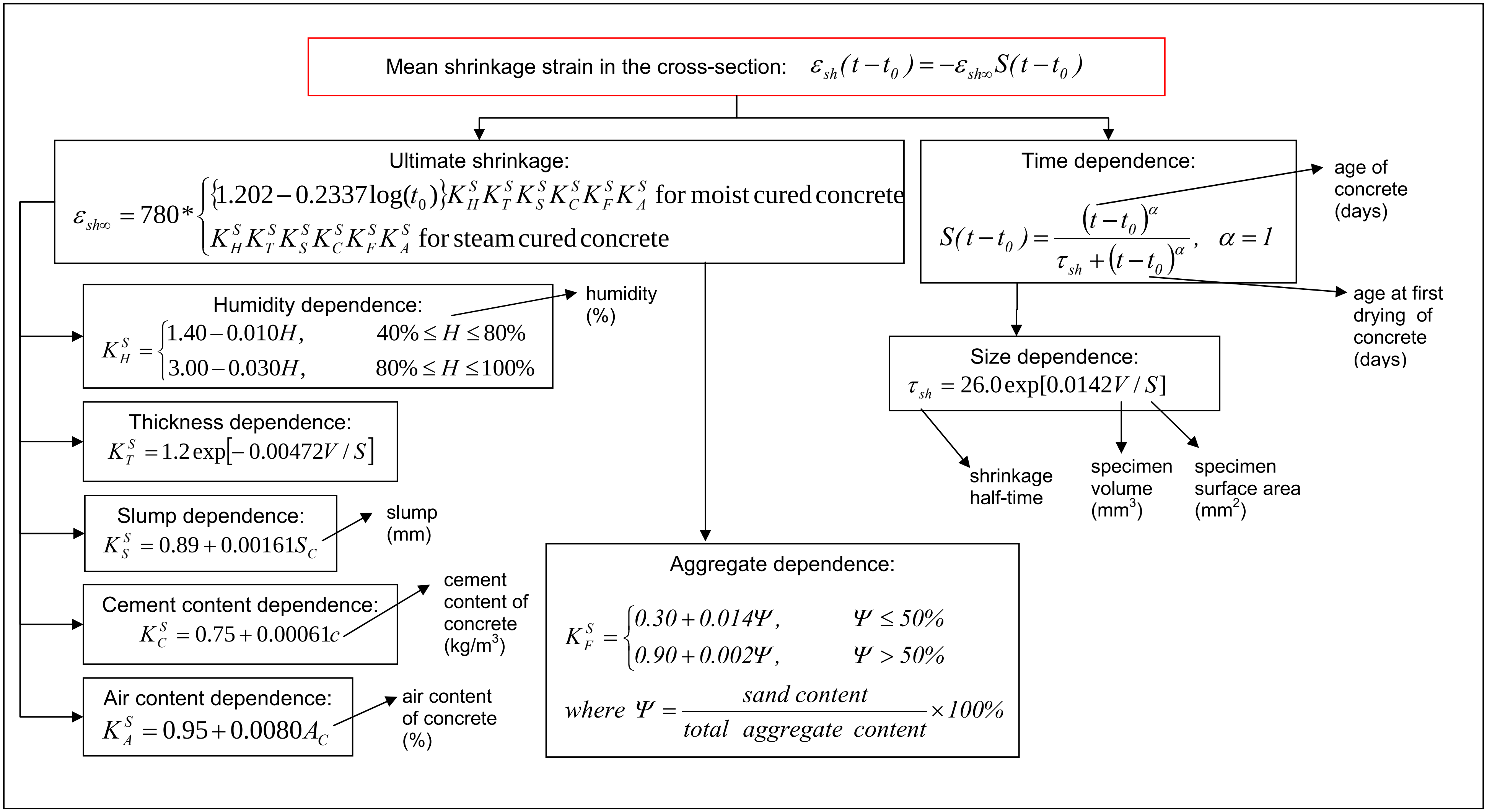


Figure 7.1 Summary of the ACI 209R-92 model for concrete shrinkage

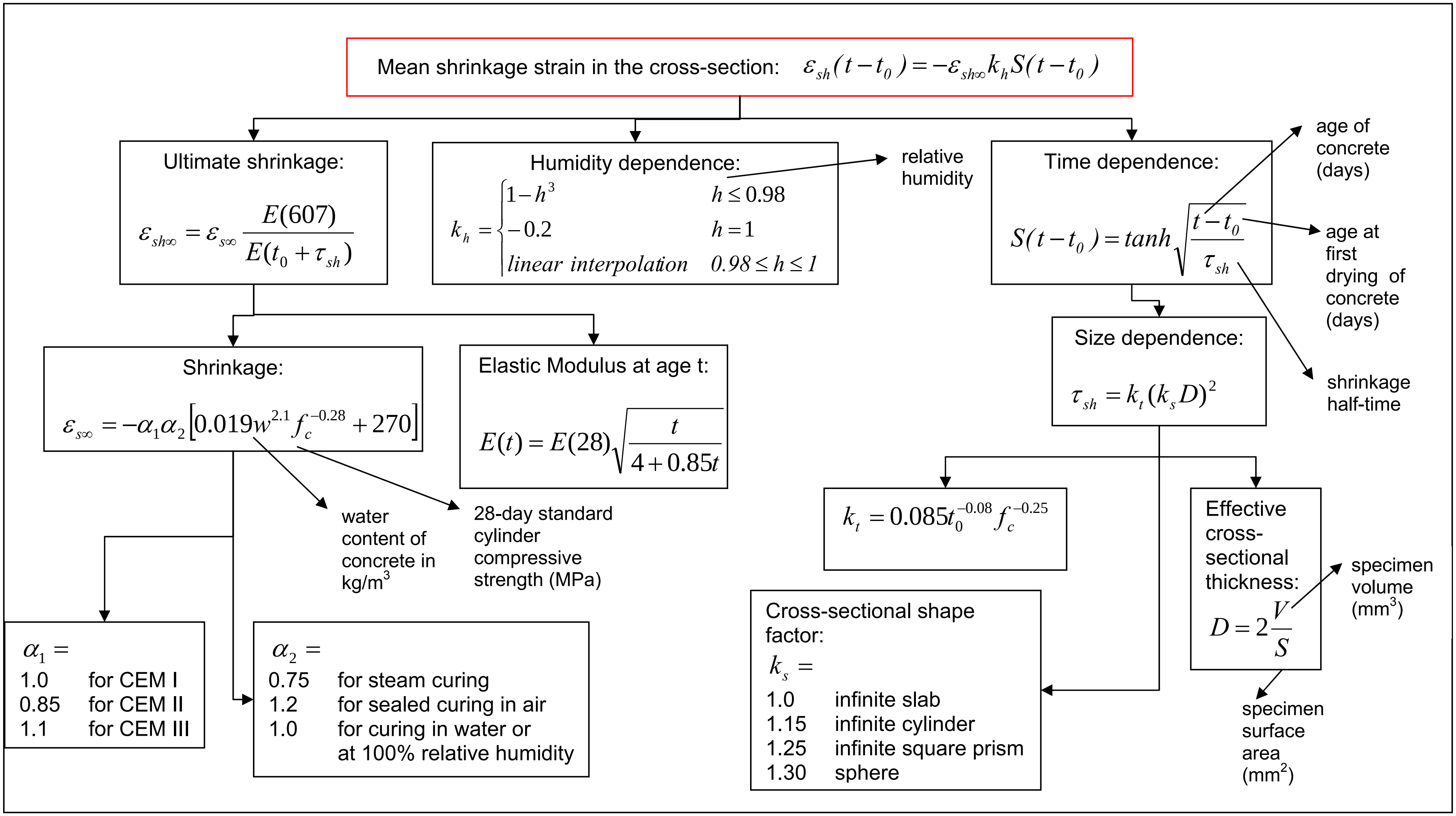


Figure 7.2 Summary of the RILEM B3 model for concrete shrinkage

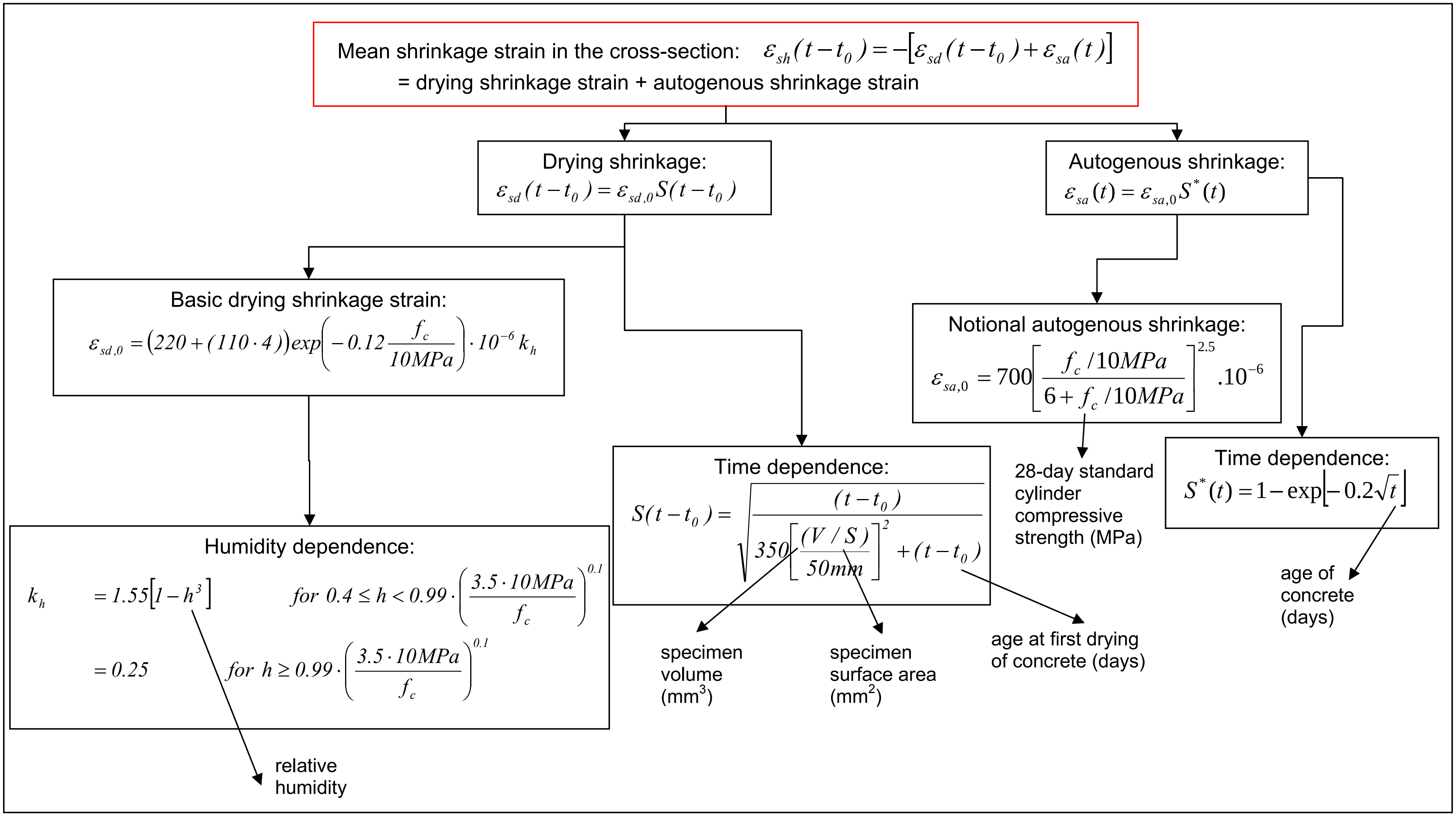


Figure 7.3 Summary of the CEB MC90-99 model for concrete shrinkage

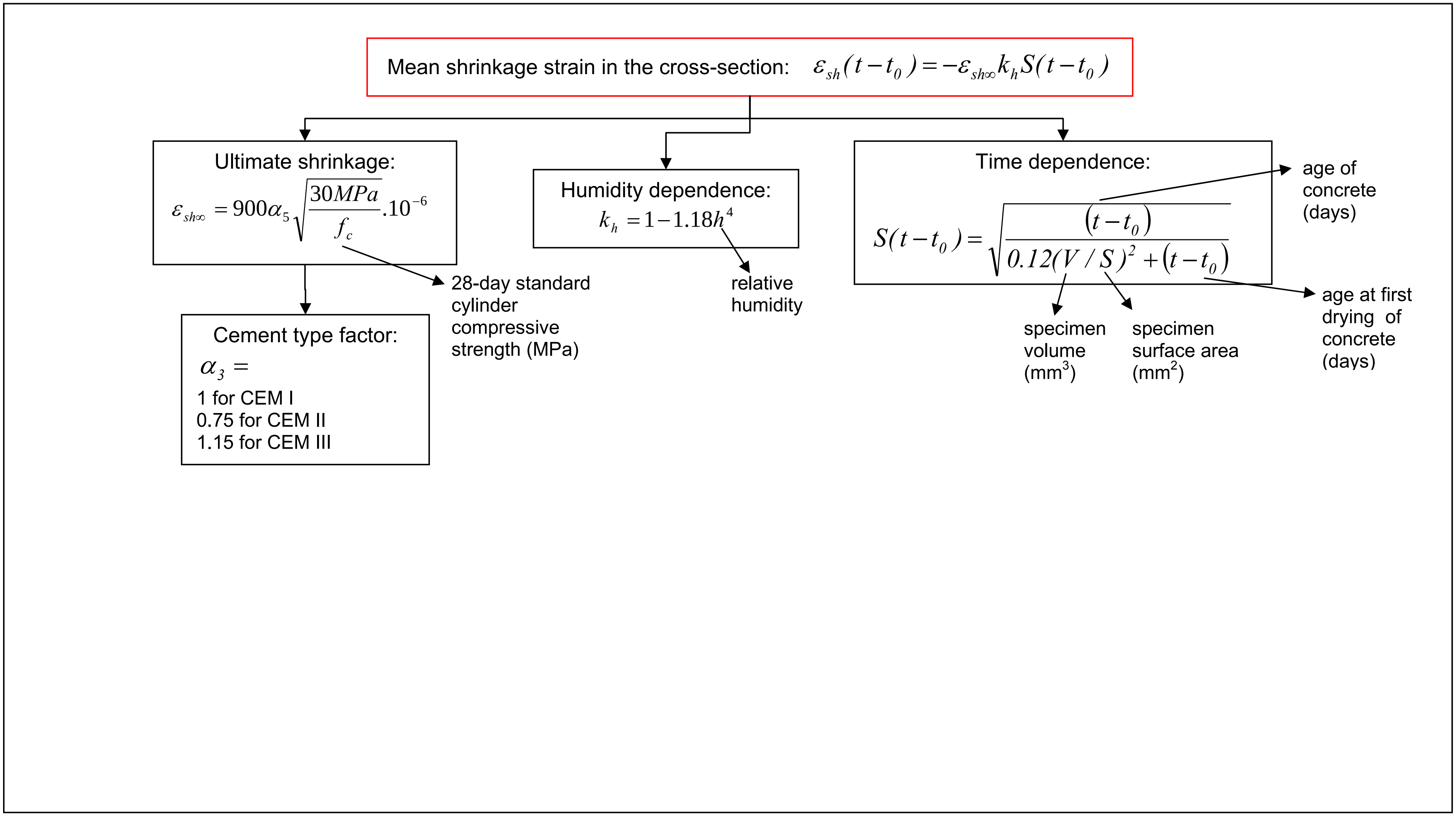


Figure 7.4 Summary of the GL2000 model for concrete shrinkage

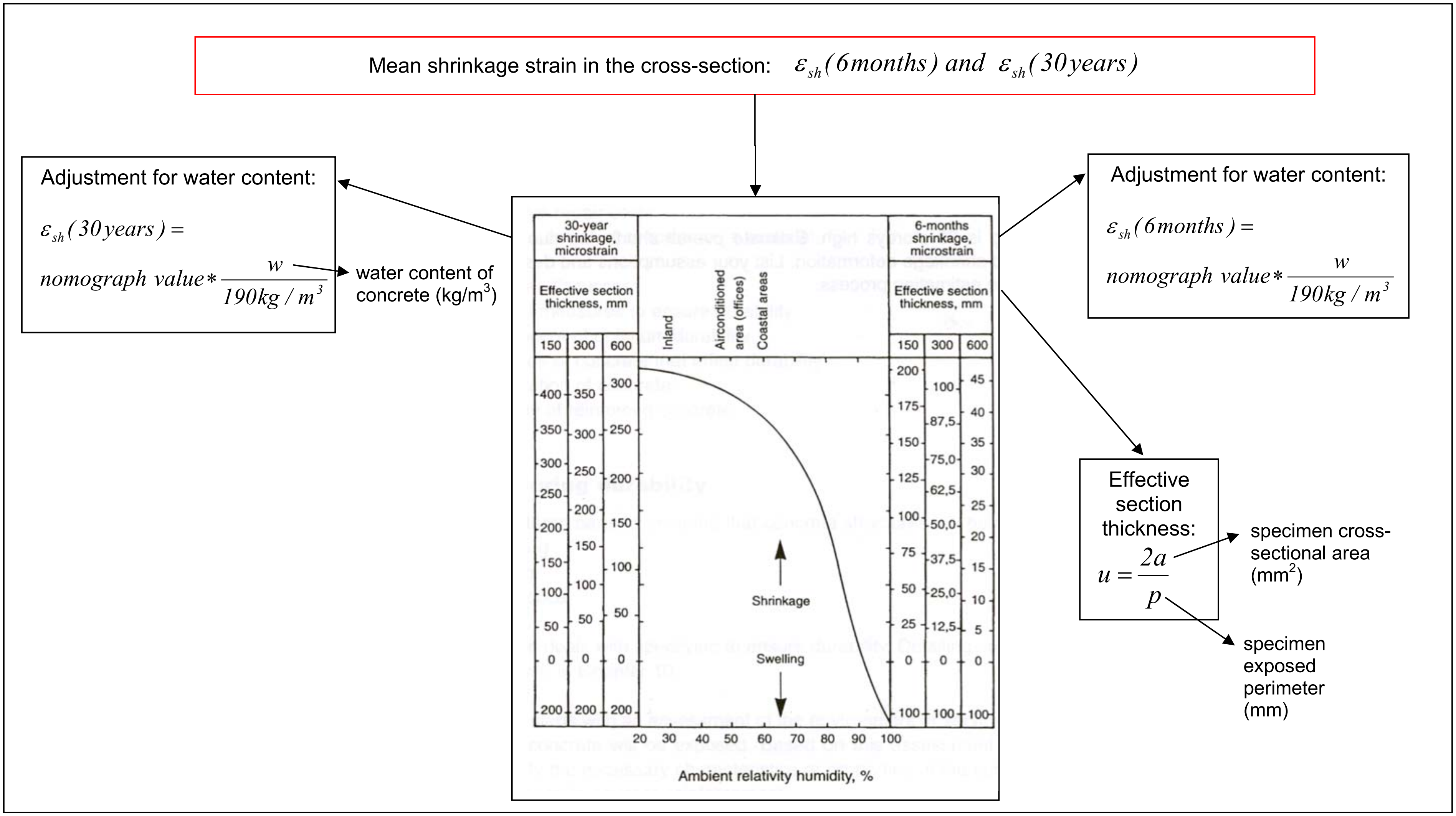


Figure 7.5 Summary of the SANS 10100-1 model for concrete shrinkage

Mean shrinkage strain in the cross-section: $\varepsilon_{sh}(t-t_0)=\alpha(1-e^{-\beta(t-t_0)})^{\gamma}$

α = -2245.19
+ Cement type factor:
- -3.85 CEM V A
- 0 CEM I, CEM II A-D, CEM II A-L, CEM II A-M(L), CEM II A-S, CEM II A-V, CEM II B-M(V/L), CEM II B-S, CEM B-V, CEM III A GGCS and CEM III A GGFS
- 8.63 CEM A GGBS

+ Stone type factor: (all range 0 to 1)
- -43.02 Granite
- 0 Andesite, Dolerite, Dolomite, Greywacke, Pretoria Quartzite, Shale, Wits Quartzite
- 211.75 Tillite
- 302.21 Quartzite

+ Sand type factor: (range 0 to 1 unless otherwise indicated)
- 0 Cape Flats, Dolomite, Ecca Grit, Pretoria Quartzite, Quartzite (0 to 0.8), River (0 to 0.25), Tillite (0 to 0.8)
- 99.54 Andesite
- 134.27 Dolerite
- 139.92 Klipheuwel Pit
- 170.57 Natural
- 201.75 Granite
- 260.15 Wits Quartzite
- 321.77 Shale
- 431.08 River Vaal (0 to 0.2)

+ 55.71 * ln(Cement Content)
+ 2.54 * Water Content
- 0.05 * Stone Content
+ 25.35 * Aggregate / Binder Ratio
+ 173.17 * ln(2*Volume To Surface Area Ratio)
+ 44.34 * Temperature

ln(β) = 9.76
+ Cement type factor:
- -0.29 CEM A GGBS
- 0 CEM I, CEM II A-D, CEM II A-L, CEM II A-M(L), CEM II A-S, CEM II A-V, CEM II B-M(V/L), CEM II B-S, CEM B-V, CEM III A GGCS and CEM III A GGFS
- 0.79 CEM V A

+ Stone type factor: (all range 0 to 1)
- -0.32 Tillite
- 0 Andesite, Dolerite, Dolomite, Greywacke, Pretoria Quartzite, Shale, Wits Quartzite, Quartzite
- 0.33 Granite

+ Sand type factor: (range 0 to 1 unless otherwise indicated)
- -0.64 Wits Quartzite
- -0.49 Shale
- -0.42 Natural
- -0.36 Granite
- -0.35 River Vaal (0 to 0.2)
- 0 Cape Flats, Dolomite, Ecca Grit, Pretoria Quartzite, Quartzite (0 to 0.8), River (0 to 0.25), Tillite (0 to 0.8)
- 0.04 Andesite
- 0.43 Klipheuwel Pit
- 1.22 Dolerite

+ 0.01 * ln(Cement Content)
- 0.05 * Aggregate / Binder Ratio
- 1.76 * ln(2*Volume To Surface Area Ratio)
- 0.26 * Temperature

ln(γ) = 3.04
+ Cement type factor:
- -0.01 CEM A GGBS
- 0 CEM I, CEM II A-D, CEM II A-L, CEM II A-M(L), CEM II A-S, CEM II A-V, CEM II B-M(V/L), CEM II B-S, CEM B-V, CEM III A GGCS and CEM III A GGFS
- 0.19 CEM V A

+ Stone type factor: (all range 0 to 1)
- 0 Andesite, Dolerite, Dolomite, Greywacke, Pretoria Quartzite, Shale, Wits Quartzite
- 0.01 Quartzite
- 0.05 Tillite
- 0.34 Granite

+ Sand type factor: (range 0 to 1 unless otherwise indicated)
- -2.58 River Vaal (0 to 0.2)
- -0.44 Wits Quartzite
- -0.40 Shale
- -0.36 Granite
- -0.12 Natural
- -0.03 Andesite
- 0 Cape Flats, Dolomite, Ecca Grit, Pretoria Quartzite, Quartzite (0 to 0.8), River (0 to 0.25), Tillite (0 to 0.8)
- 0.02 Klipheuwel Pit
- 0.50 Dolerite

+ 0.16 * ln(Cement Content)
+ 0.08 * Aggregate / Binder Ratio
- 0.62 * ln(2*Volume To Surface Area Ratio)
- 0.08 * Temperature

For allowable combinations of cement, stone and sand covered by model: see Table 6.14

Figure 7.6 Summary of the WITS model for concrete shrinkage

- Concrete properties (28-day Compressive Strength, 28-day Elastic Modulus and Slump): These were omitted from the model due to large amounts of missing data. The contribution of these variables may, however, have been expressed in the WITS model through the use of the sand and stone type covariates.

It is recommended that exploration of these variables be considered in future experiments designed to develop the WITS model further.

Conversely, the WITS model contained some covariates which do not appear in the published models:

- Stone and sand types.
- Stone content (although the ACI 209R-92 model uses the ratio of sand to total aggregate content).
- Aggregate / binder ratio.
- Temperature (significant despite its small range of 21-25 ºC).

7.2 Comparison of the predictive ability of the published and WITS models with respect to the South African data set

In the published literature, the most thorough model comparisons have been based on the RILEM data bank(24), a collection of 490 concrete shrinkage profiles mainly from North American and European research groups. The RILEM B3 model was developed on an older version of the current RILEM data bank(10,11). In this study, model comparisons were based on the local data set, which may be considered as a smaller, South African, version of the RILEM data bank.

A predicted shrinkage profile was calculated for each model, including the WITS model, for each qualifying experiment in the data set (in terms of the range of applicability of each model – see Table 7.2) and the goodness-of-fit of these predictions to the actual data was assessed. Since the WITS model was also derived from this data set, its assessments were divided into two groups, viz. the experiments used to derive the model and the experiments which qualified to be predicted by the model but which were not used to derive the model. Combined goodness-of-fit statistics for the two groups are also presented.

The ACI 209R-92, RILEM B3, CEBMC90-99 and GL2000 models were implemented according to the model specifications given in the American Concrete Institute Committee 209's guide for modelling and calculating shrinkage and creep in hardened concrete(3). Model implementation was checked against the example provided in Appendix C of this reference. The SANS 10100-1 model was implemented according to the SANS 10100-1 standard(14) for the prediction of shrinkage in concrete. The following points regarding the implementation of the models should be noted:

- In the ACI 209R-92 model, the air content was set at the standard value of 6%, making the air content factor equal to one, since data on this variable were not available.
- For the RILEM B3, CEB MC90-99 and GL2000 models, 28-day cube compressive strengths were converted to the corresponding cylinder strengths using the conversion table given in the British Standard *Common Rules for Buildings and Civil Engineering Structures*(53).
- The effective section thicknesses of the specimens used in this study were smaller than the minimum value of 150 mm presented in the SANS 10100-1 model. The values given in the standard were extrapolated to the required effective section thickness (and interpolated to the required relative humidity) by applying separate quadratic models (for the 6-month and 30-year shrinkage) fitted to the shrinkage data read from the nomograph at relative humidities of 40, 50, 60, 70 and 80% and effective section thicknesses of 150, 300 and 600 mm:

  6-month shrinkage (microstrain) = $314.0 + 1.035H - 1.025u + 0.003494Hu - 0.02962H^2 + 0.0007302u^2$ (R^2=0.994)

  30-year shrinkage (microstrain) = $395.6 + 6.231H - 0.6173u + 0.003239Hu - 0.09595H^2 + 0.0002922u^2$ (R^2=0.994)

  where H is the humidity in % and u is the effective section thickness in mm.

  After correction for the water content of the concrete (see Figure 7.5), the predicted 6-month and 30-year shrinkage values were used to determine the time-shift factor, α, in the hyperbolic growth curve[29,79] used to determine the shrinkage at other drying times:

$$\varepsilon_{sh}(t-t_0)=\frac{t}{\alpha+t}\cdot\varepsilon_{sh}(30\,years)$$

$$\text{where } \alpha=183\,days\cdot\frac{\varepsilon_{sh}(30\,years)-\varepsilon_{sh}(6\,months)}{\varepsilon_{sh}(6\,months)}$$

Firstly, the **proportion of the data set which could be predicted by a particular model** was determined. The results are presented in Figure 7.7.

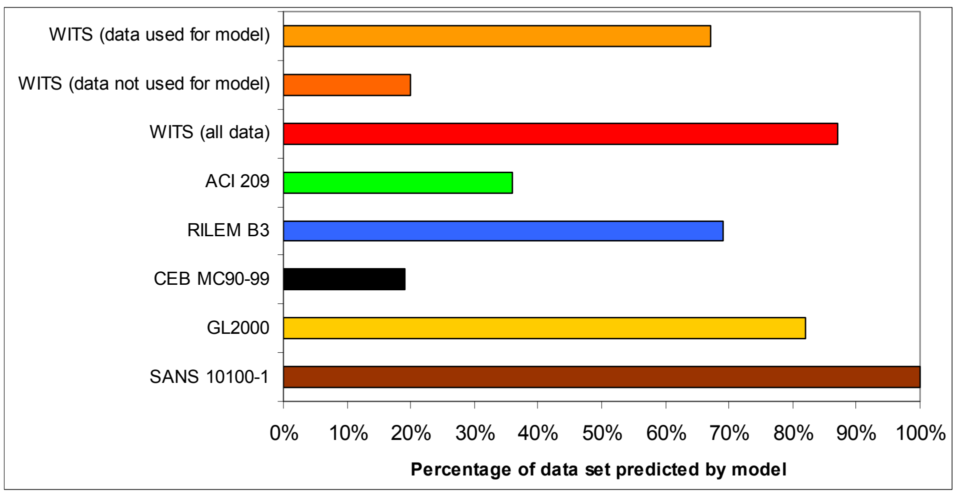


Figure 7.7 The percentage of the data set which could be predicted by each of the models

As a result of its simple requirements, the SANS 10100-1 model could be used to predict all the experiments in the data set. The WITS model had the next highest proportion of experiments which could be predicted (87%). Those experiments which did not qualify, failed to do so mostly because they used aggregate types which were not included in the derivation of the model. Sixty-seven percent of the data set was used to develop the model, leaving 20% (47 experiments) of the data set which qualified for prediction by the model but which was not used in the development of the model. The latter set was not regarded as a true validation data set since it comprised those shrinkage profiles which were excluded from the multivariate multiple regression analysis for data quality reasons (see Section 6.2). In the discussions which follow, reference to the WITS model includes consideration of all three subsets of data unless specifically indicated otherwise.

The GL2000 model was able to predict 82% of the experiments in the data set. In fact, all the experiments qualified for prediction, as a result of simple qualifying criteria, but missing data reduced the number of experiments for which predictions could be made. The RILEM B3 model could be used to predict 80% of the experiments in the data set based on the qualifying criteria (most failures to qualify resulted from missing 28-day compressive strength data) but missing data (mostly the 28-day elastic modulus) reduced this figure to 69%. As a result of the qualifying criteria relating to cement type and cement content, only 36% of the data set qualified for prediction by the ACI 209R-92 model. The CEB MC90-99 model fared even worse with only 19% of the data qualifying for prediction. This stemmed mainly from the qualifying criterion on curing time (age at first drying) which was exceeded in most cases by the experiments in the data set.

The **fit of the various models to each experiment in the data set** is shown in the spreadsheet “Model predictions.xls” on the accompanying CD. Six of these shrinkage profiles are illustrated in Figure 7.8. The illustrated profiles were randomly selected from the 57 experiments to which five or all six of the models could be fitted. It is clear that the WITS model performed well, which was to be expected since the model was derived from the data. The SANS 10100-1 and GL2000 models tended to under- and over-predict, respectively. No trend regarding the performance of the other models was immediately obvious from an inspection of the model fits to the raw data.

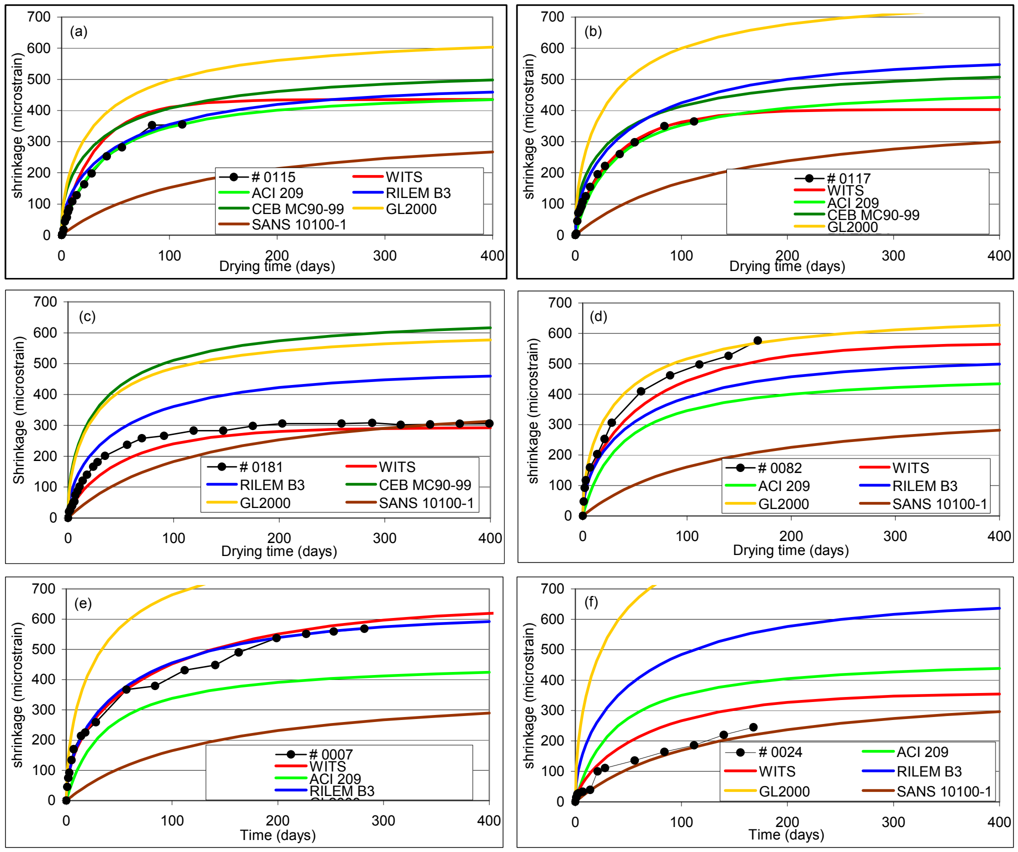


Figure 7.8 Illustration of the fit of the models to six experiments in the data set. The experiments illustrated in (a) – (d) were part of the data set used to derive the WITS model, while the experiments illustrated in (e) and (f) were not

Many **goodness-of-fit measures** have been used by different researchers who have developed models for concrete shrinkage. The American Concrete Institute is of the opinion that “the statistical indicators available are not adequate to uniquely distinguish between models”(3). In this study, a broad range of these measures, both graphical and numerical, was used, and a critical assessment of these measures from a statistical point of view is also given where possible.

Plots of the **actual versus predicted shrinkage values** are shown in Figure 7.9. The points from an ideal model would lie entirely on the unity line. Bažant(11,80) recommends focusing on the longer drying times, where predictions are most important, since such plots are typically dominated by data gathered at short drying times. The plots showed that the WITS model, particularly at longer drying times, remained closest to the unity line, thus indicating the best fit to the data. The adjusted R^2, as well as the slope and intercept, of the fitted lines to the plots are given in Table 7.3. A good model would have a slope close to 1 and an intercept close to 0, as well as a high adjusted R^2 value. The WITS model exhibited the least scatter (highest R^2) and the slope closest to 1, while its intercept was the second closest to zero of all the models. On balance of the three regression criteria, the CEB MC90-99 and ACI 209R-92 models performed the next best. The over-prediction by the RILEM B3, ACI 209R-92, CEB MC90-99 and GL2000 models as well as the under-prediction by the SANS 10100-1 model was evident from both the plots and the regression statistics.

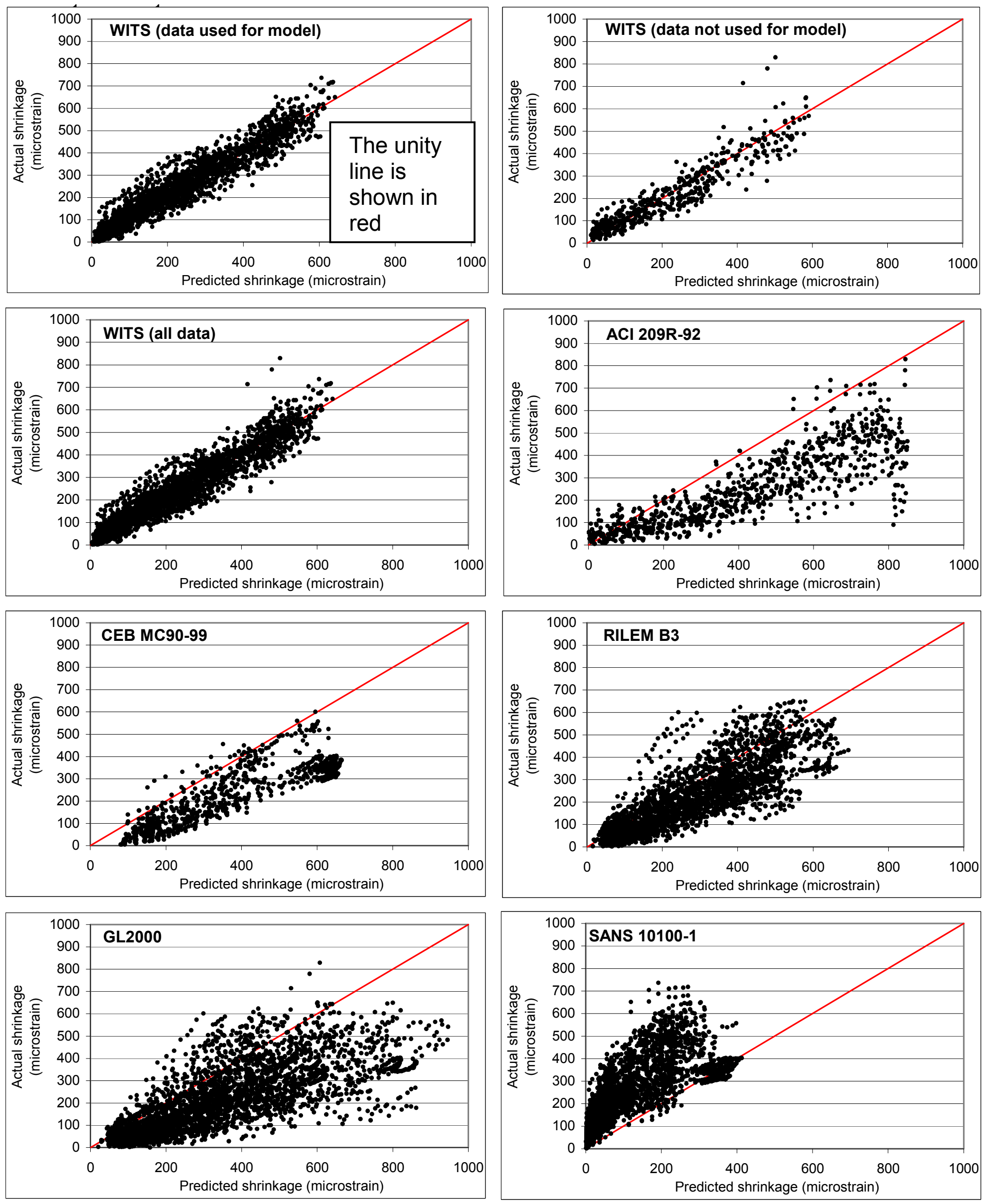


Figure 7.9 Plots of the actual vs. predicted shrinkage values for the models

Table 7.3 Parameters of the linear relationship between the actual and predicted shrinkage values for the models

Model	n	adjusted R^2	slope (95% confidence interval)	intercept (95% confidence interval)
WITS (data used for model development)	2603	0.92	0.97 (0.96 – 0.99)	18 (15 – 21)
WITS (data NOT used for model development)	483	0.87	0.89 (0.86 – 0.92)	19 (11 – 27)
WITS (all data)	3086	0.91	0.96 (0.95 – 0.97)	18 (15 – 21)
ACI 209R-92	850	0.71	0.96 (0.92 – 1.0)	53 (42 – 64)
RILEM B3	2581	0.67	0.73 (0.71 – 0.75)	17 (10 – 23)
CEB MC90-99	868	0.76	0.62 (0.59 – 0.64)	-6 (-16 – 4)
GL2000	3005	0.53	0.52 (0.50 – 0.54)	39 (31 – 46)
SANS 10100-1	3376	0.59	1.08 (1.05 – 1.11)	118 (113 – 123)

The models with the first, second and third best scores are highlighted in dark orange, light orange and yellow, respectively.

Plots of the **actual versus predicted shrinkage values on a log scale** are also useful since they illustrate the relative errors, which should decrease as shrinkage strain increases as a consequence of the homoscedasticity of errors[11,80]. These plots, and the parameters of their fitted lines, are shown in Figure 7.10 and Table 7.4 respectively. The WITS model was again closest to the unity line and exhibited the least scatter, indicating it to be the best model when viewed according to these criteria. The poorer performance of the WITS model on the data not included in the data set from which the model was derived, was more apparent in these plots than in the previous set. The RILEM B3 model exhibited more scatter of the data around the unity line (i.e. greater positive and negative residuals) at longer drying times than the WITS model, while the performance of the other models was worse.

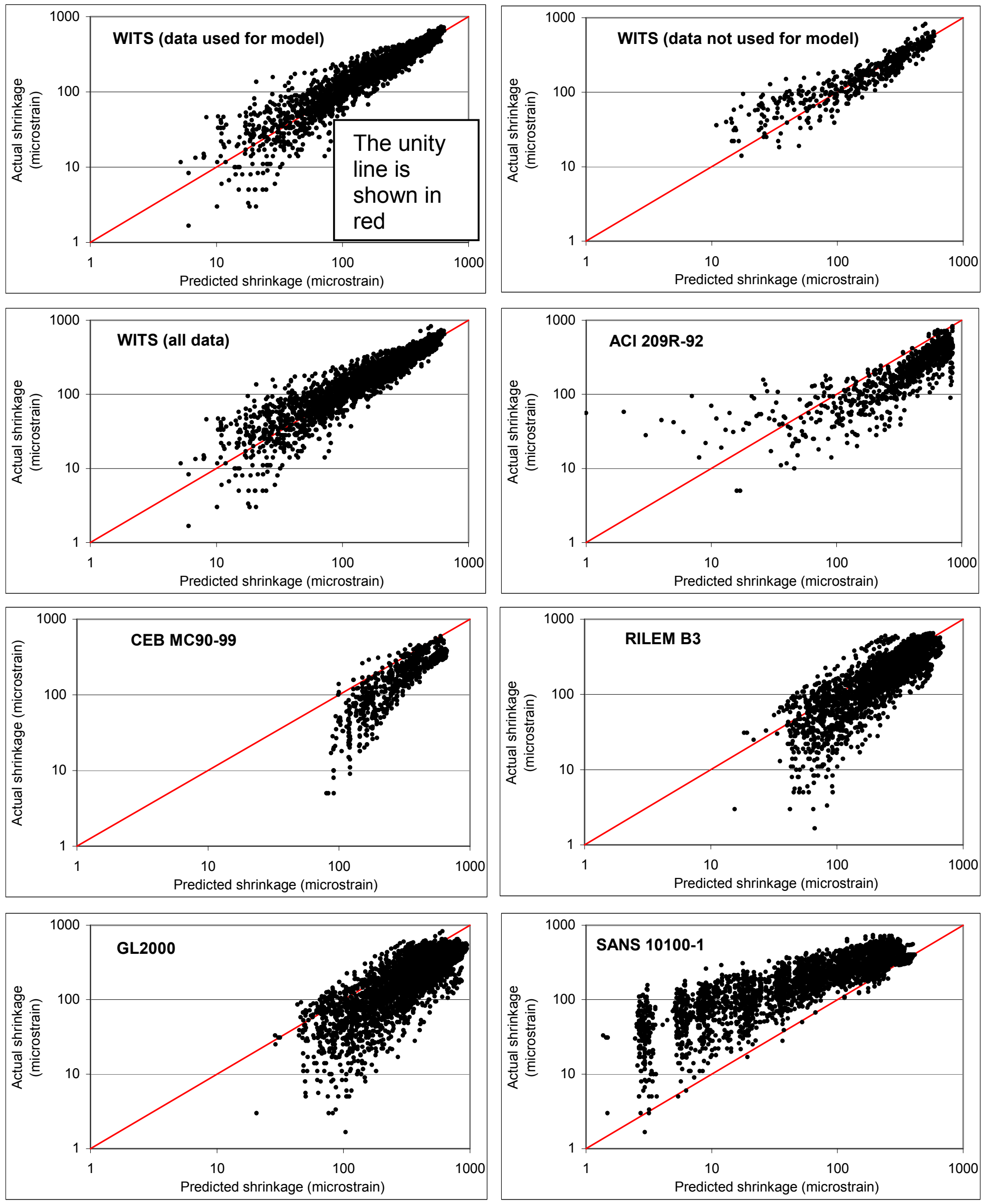


Figure 7.10 Plots of the actual vs. predicted shrinkage values, on a log scale, for the models

Table 7.4 Parameters of the linear relationship between the actual and predicted shrinkage values, on a log scale, for the models

Model	n	adjusted R^2	slope (95% confidence interval)	intercept (95% confidence interval)
WITS (data used for model development)	2603	0.86	0.96 (0.94 – 0.97)	0.11 (0.08 – 0.14)
WITS (data NOT used for model development)	483	0.84	0.76 (0.73 – 0.79)	0.55 (0.48 – 0.62)
WITS (all data)	3086	0.85	0.93 (0.91 – 0.94)	0.18 (0.15 – 0.22)
ACI 209R-92	850	0.67	0.70 (0.67 – 0.74)	0.76 (0.67 – 0.84)
RILEM B3	2579	0.68	1.06 (1.03 – 1.09)	-0.29 (-0.36 – -0.22)
CEB MC90-99	868	0.67	1.33 (1.27 – 1.39)	-1.1 (-1.3 – -0.96)
GL2000	3005	0.59	1.04 (1.01 – 1.07)	-0.35 (-0.43 – -0.28)
SANS 10100-1	3374	0.71	0.54 (0.53 – 0.55)	1.31 (1.28 – 1.33)

The models with the first, second and third best scores are highlighted in dark orange, light orange and yellow, respectively.

In the above plots, older, lower-strength concretes, which exhibit higher levels of shrinkage strain at a given drying time, were weighted more heavily. To correct for this, as suggested by civil engineering researchers, plots of **actual versus predicted shrinkage values, both multiplied by (28-day compressive strength for the experiment / average 28-day compressive strength for the data set)$^{1/2}$** were examined[11,80]. This effectively normalised shrinkage according to compressive strength. The results are shown in Figure 7.11 and Table 7.5. The conclusions were similar to those made from the previous plots: the WITS model performed the best, while the RILEM B3 and ACI 209R-92 models exhibited the next best performance.

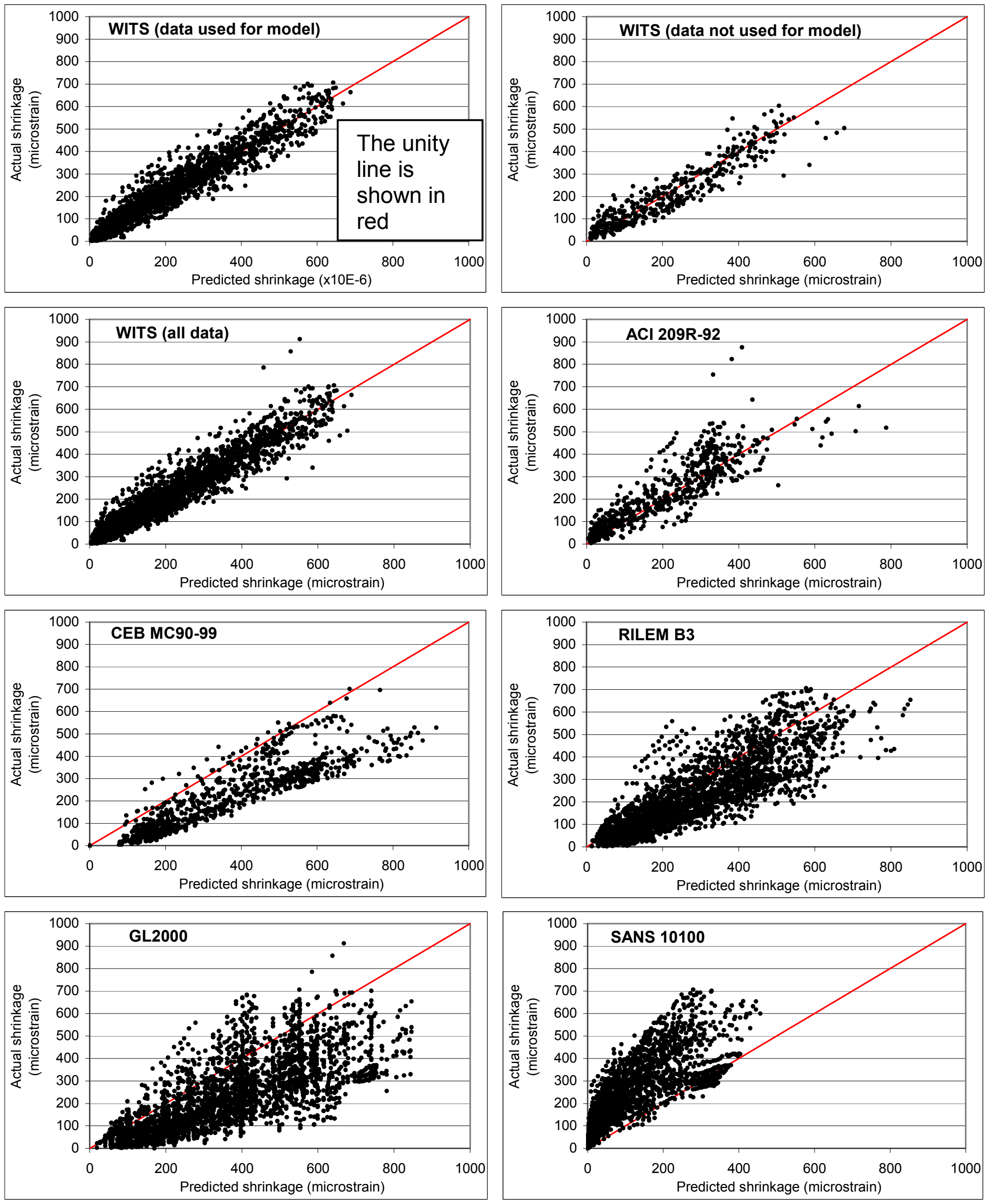


Figure 7.11 Plots of the actual vs. predicted shrinkage values, both multiplied by (28-day compressive strength for the experiment / average 28-day compressive strength for the data set)$^{1/2}$, for the models

Table 7.5 Parameters of the linear relationship between the actual and predicted shrinkage values, both multiplied by (28-day compressive strength for the experiment / average 28-day compressive strength for the data set)$^{1/2}$, for the models

Model	n	adjusted R^2	slope (95% confidence interval)	intercept (95% confidence interval)
WITS (data used for model development)	2132	0.92	0.98 (0.96 – 0.99)	18 (15 – 22)
WITS (data NOT used for model development)	472	0.87	0.91 (0.88 – 0.95)	16 (7.0 – 25)
WITS (all data)	2602	0.90	0.96 (0.95 – 0.98)	18 (15 – 22)
ACI 209R-92	616	0.80	0.89 (0.86 – 0.93)	40 (31 – 49)
RILEM B3	2579	0.70	0.76 (0.74 – 0.78)	8 (2 – 15)
CEB MC90-99	868	0.75	0.68 (0.66 – 0.71)	-28 (-39 – -17)
GL2000	3005	0.56	0.61 (0.59 – 0.63)	13 (5 – 21)
SANS 10100-1	2885	0.62	1.11 (1.07 – 1.14)	108 (103 – 113)

The models with the first, second and third best scores are highlighted in dark orange, light orange and yellow, respectively.

Plots of the **residuals vs. log$_{10}$(time)** are shown in Figure 7.12. These should not fan out or show any other obvious pattern or trend(81,82). The residuals for the WITS model conform to these criteria. However, the residuals for the other models all fan out, indicating increased deviation of the models from the raw data at longer drying times (where prediction is more important). The over- and under-prediction of the different models, as discussed previously, can clearly be seen in these plots.

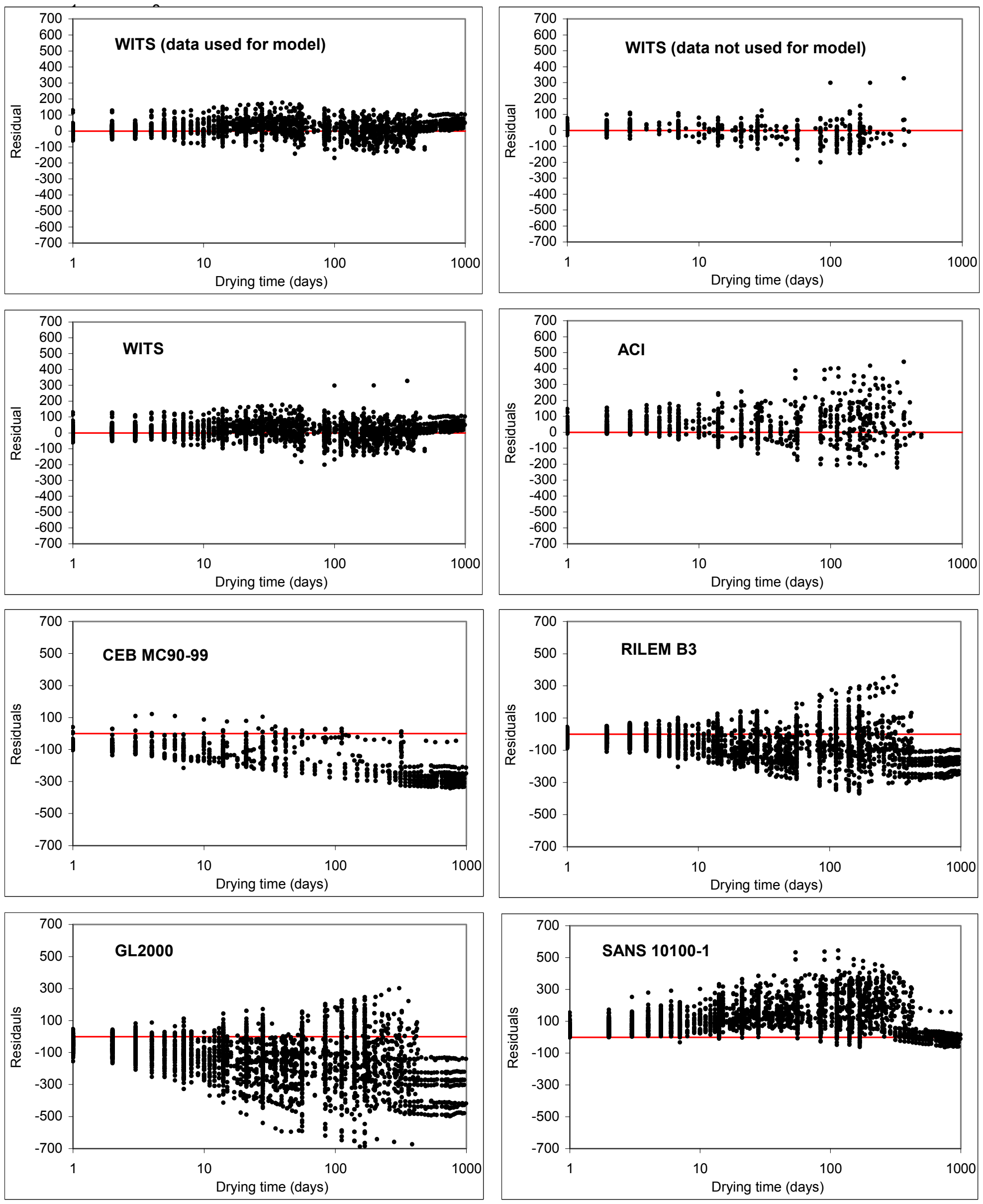


Figure 7.12 Plots of the residuals vs. $\log_{10}$(time) for the models

The **distribution of residuals for each decade of time (on a $\log_{10}$ scale)** spanned by the shrinkage profile was also examined[15]. Subdivision into decades of drying time was carried out since shrinkage is a decaying process, so a time increment of, say, ten days makes a much larger difference at a drying time of 10 days than at 1000 days. The shape of the histograms, their mean and standard deviations were examined for normality. The results are summarised in Figures 7.13 – 7.15.

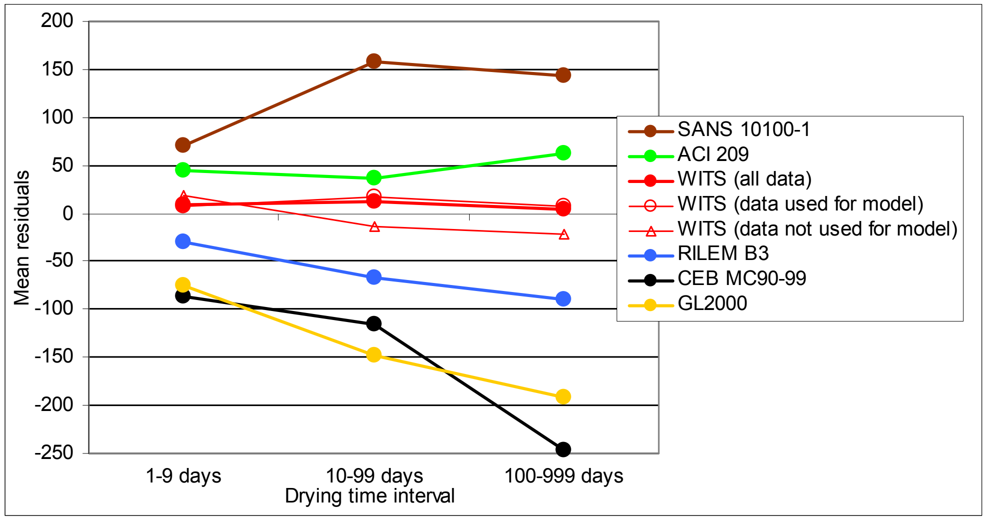


Figure 7.13 Plot of the mean residuals, for each decade of time, for the models

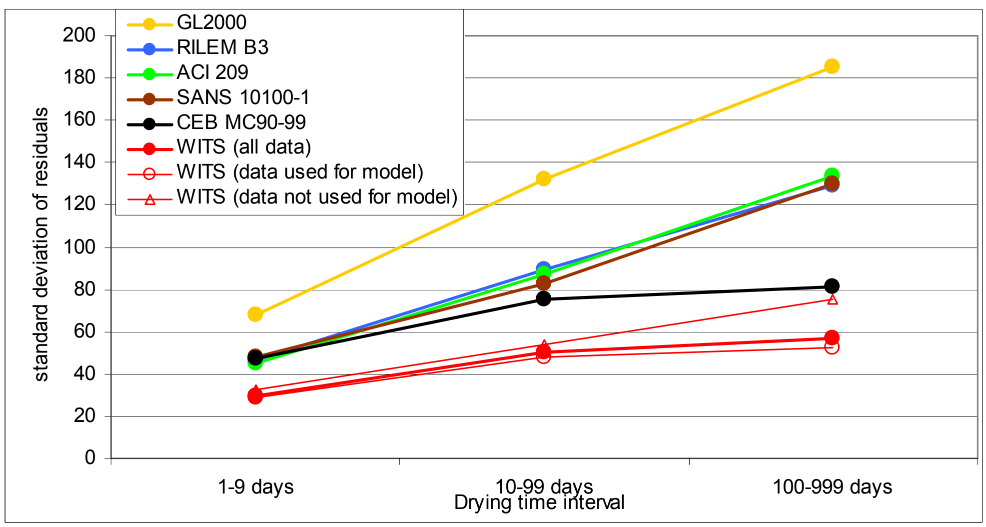


Figure 7.14 Plot of the standard deviation of the residuals, for each decade of time, for the models

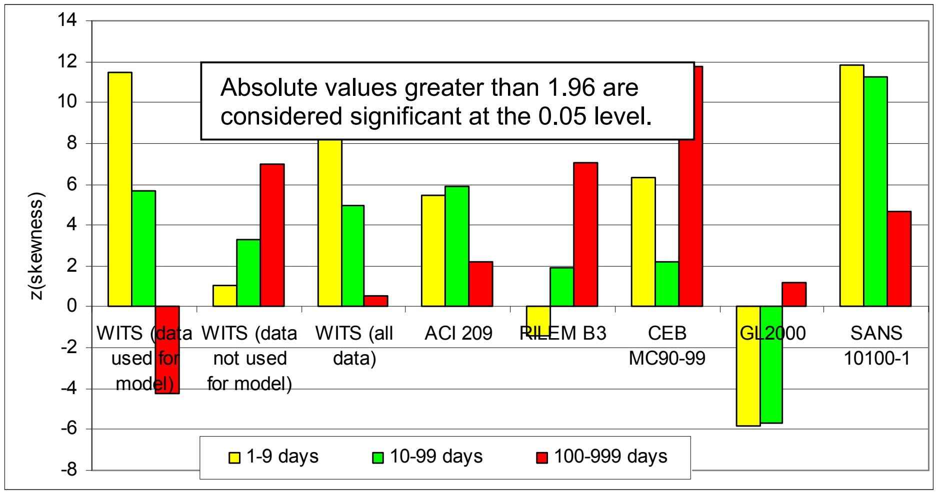


Figure 7.15 Plot of z(skewness) = coefficient of skewness/√(6/n) of the residuals, for each decade of time, for the models

The mean residuals for the WITS model were the closest to zero and the most consistent across the different intervals of drying time, whereas the absolute values of the mean residuals of the other models tended to increase with drying time. The ACI 209R-92 and RILEM B3 models exhibited the next best performance on this criterion. The standard deviations of the residuals increased with drying time, as expected. Again, the WITS model exhibited the smallest standard deviations. The residuals from all the models exhibited significant skewness for at least one interval of drying time, and it was difficult to draw any meaningful conclusions from these data.

Next, various **numerical goodness-of-fit summary statistics** were calculated. The results are given in Table 7.6. The **mean absolute deviation** (MAD) is defined as

$$MAD = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{l}\left|\Delta_{ij}\right|$$

- where Δ_{ij} is the difference between the *i*th actual and predicted shrinkage value for the *j*th experiment and *N* is the grand total of the number of data points. While this measure gives an indication of the spread of the residuals, it suffers from two disadvantages in the current context, namely that it weights all data points equally (whereas data points at longer drying times are regarded as being more important) and it does not explicitly indicate positive or negative bias of the predictions if any is present.

Table 7.6 Summary of the numerical goodness-of-fit statistics calculated for the models

Model	WITS (data used for model development)	WITS (data NOT used for model development)	WITS (all data)	ACI 209R-92	RILEM B3	CEB MC90-99	GL2000	SANS 10100-1
Mean absolute deviation (MAD) (microstrain)	35	34	41	74	91	158	156	129
Mean squared prediction error (MSPE) (microstrain^2)	2 138	3 222	2 308	10 497	13 430	33 925	39 976	25 837
Bažant's coefficient of variation,								
ω_{BP} (overall) (%)	25	35	27	49	84	84	130	76
ω_{BP} (1-9 days) (%)	37	45	39	70	76	133	136	105
ω_{BP} (10-99 days) (%)	21	25	22	36	49	59	87	75
ω_{BP} (100-999 days) (%)	13	20	13	33	40	71	67	48
CEB coefficient of variation								
V_{CEB} (overall) (%)	22	32	23	43	54	83	95	63
V_{CEB} (0-10 days) (%)	37	45	39	68	75	126	135	104
V_{CEB} (11-100 days) (%)	21	25	21	36	49	58	86	74
V_{CEB} (101-365 days) (%)	13	20	14	33	36	57	58	53
V_{CEB} (366-730 days) (%)	13		13	18	50	79	87	22
V_{CEB} (731-1095 days) (%)	15		15		54	77	93	5
CEB mean square error								
F_{CEB} (overall) (%)	35	34	34	38	141	174	212	53
F_{CEB} (0-10 days) (%)	70	45	66	55	293	356	426	84
F_{CEB} (11-100 days) (%)	24	30	25	38	76	89	135	66
F_{CEB} (101-365 days) (%)	15	24	17	33	45	65	85	45
F_{CEB} (366-730 days) (%)	13		13	16	55	83	91	17
F_{CEB} (731-1095 days) (%)	14		14		55	80	94	10

Table 7.6 Summary of the numerical goodness-of-fit statistics calculated for the models (continued)

Model	WITS (data used for model development)	WITS (data NOT used for model development)	WITS (all data)	ACI 209R-92	RILEM B3	CEB MC90-99	GL2000	SANS 10100-1
CEB mean relative deviation								
M_{CEB} (overall)	0.99	1.01	0.99	0.96	1.25	1.40	1.41	0.76
M_{CEB} (0-10 days)	1.04	0.91	1.02	0.89	1.47	1.76	1.75	0.42
M_{CEB} (11-100 days)	0.98	1.05	0.99	0.98	1.20	1.30	1.37	0.60
M_{CEB} (101-365 days)	1.01	1.05	1.02	0.97	1.14	1.25	1.22	0.77
M_{CEB} (366-730 days)	0.97		0.97	1.01	1.23	1.34	1.35	0.97
M_{CEB} (731-1095 days)	0.94		0.94		1.23	1.33	1.37	1.02
Gardner mean residuals								
3-9.9 days	10	16	11	49	-36	-94	-89	84
10-31.5 days	16	-3	12	35	-56	-120	-137	140
31.6-99 days	18	-33	10	38	-79	-107	-163	181
100-315 days	-3	-23	-6	72	-60	-196	-144	188
316-999 days	26	53	27	7	-159	-263	-282	50
Gardner root mean square of residuals								
3-9.9 days	32	39	33	70	64	108	116	97
10-31.5 days	48	44	47	81	97	137	180	156
31.6-99 days	54	61	55	114	127	140	222	202
100-315 days	54	79	59	149	145	219	236	222
316-999 days	50	144	54	143	184	272	317	116
Gardner coefficient of variation								
ω_G (overall) (%)	17	27	18	36	49	71	83	60

The models with the first, second and third best scores are highlighted in dark orange, light orange and yellow, respectively.

The WITS model had the lowest MAD, followed by the ACI 209R-92, RILEM B3 and then the SANS 10100-1 models (see Table 7.6).

The **mean squared prediction error** (MSPE) is defined as

$$MSPE = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{l}\Delta_{ij}^{2} \ .$$

This measure suffers from the same drawbacks as the Mean Absolute Deviation. The relative performance of the models with respect to the MSPE was the same as that for the MAD.

The numerical goodness-of-fit measures which follow were derived by researchers who developed the other models for concrete shrinkage. They are not necessarily good from a statistical perspective, but have been used here for comparative purposes since they are widely used in the concrete industry[3].

Bažant's coefficient of variation, $\overline{\omega}_{BP}$ [11,81,81] summarises the deviations of the model from the observed data. The coefficient of variation for a given experiment (shrinkage profile) *j* is defined as

$$\omega_j = \frac{s_j}{\bar{\varepsilon}_j} = \frac{1}{\bar{\varepsilon}_j}\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(w_{ij}\Delta_{ij}\right)^2}$$

where

- s_j is the standard deviation of the differences Δ_{ij} between the actual and predicted (model) shrinkage values for experiment *j*,
- $\bar{\varepsilon}_j = \frac{1}{n}\sum_{i=1}^{n} w_{ij}\varepsilon_{ij}$ is the average value of shrinkage strain in experiment *j*, based on the individual shrinkage strain measurements ε_{ij}, their weights w_{ij} and the number of data points *n* in the experiment and

the weights $w_{ij} = \frac{1}{n_d n_{d(ij)}}$ are based on n_d, the number of decades of time (on a $\log_{10}$ scale) spanned by the shrinkage profile and $n_{d(i)}$, the number of points in the decade to which point *ij* belongs. The weights ensure that the same weighting is given to each decade of drying time spanned by the experiment. The coefficients for the individual experiments are then combined into the overall coefficient of variation for the data set, as follows:

$$\overline{\omega}_{BP} = \sqrt{\frac{1}{N_j}\sum_{j=1}^{N_j}\omega_j^2} \cdot 100$$

where N_j is the number of shrinkage profiles in the overall data set. In this way, the same weighting is given to each shrinkage profile rather than to each data point, which represents a significant advantage over the goodness-of-fit measures discussed earlier.

Bažant's coefficient of variation as described above, may also be **calculated separately for each decade of time (on a $\log_{10}$ scale) spanned by the shrinkage profile**: i.e. for 1-9 days, 10-99 days, 100-999 days and 1000 days or more[11,80]:

$$\omega_t = \frac{\sqrt{\frac{1}{n-1}\sum_{k=1}^{n}\Delta_k^2}}{\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k} \cdot 100$$

where the subscript t refers to a particular time period, the Δ_k and ε_k are the residuals and shrinkage strain measurements in that time period and n is the total number of data points in that time period. Pooled data from the entire data set is used for each time period, but since the coefficient of variation is now calculated separately for each time period, it does not matter that all data points within that time period are now weighted equally.

The results of these calculations for all the models are shown in Table 7.6 and Figure 7.16. The WITS model exhibited the lowest coefficient of variation overall, as well as across all three intervals of drying time, followed by the ACI 209R-92 model. It should also be noted that the coefficient of variation of the WITS model of 27% (calculated on the South African data set from which the model was derived) compares favourably to the coefficient of variation for the RILEM B3 model calculated on the RILEM database, from which it was derived, of 34%.

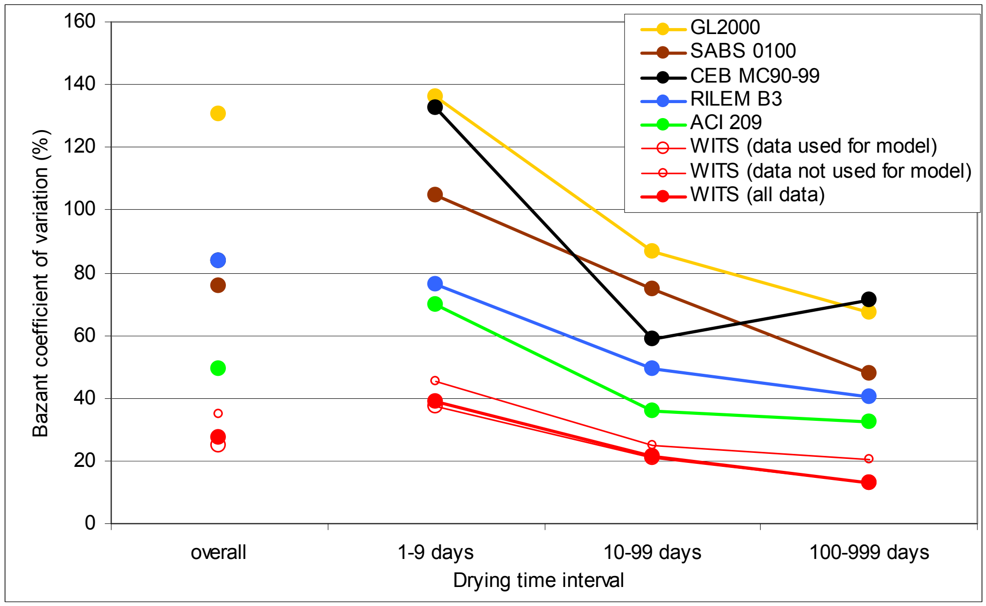


Figure 7.16 Bažant's coefficient of variation: overall and for each decade of time (on a $\log_{10}$ scale) spanned by the shrinkage profile

In the **CEB-FIP coefficient of variation** approach(81), all the data are pooled and then grouped into six intervals of drying time: 0-10 days, 11-100 days, 101-365 days, 366-730 days, 731-1095 days and above 1095 days. Firstly, the **CEB coefficient of variation** is calculated (definition of variables as above):

Coefficient of variation for time period *t*: $V_t = \frac{s_t}{\bar{\varepsilon}_t} \cdot 100$

where $\bar{\varepsilon}_t = \frac{1}{n}\sum_{k=1}^{n} \varepsilon_k$ and $s_t = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n} \Delta_k^2}$

CEB coefficient of variation: $V_{CEB} = \sqrt{\frac{1}{K}\sum_{t=1}^{K} V_t^2}$

where *K* is the number of time periods.

Next, the **CEB mean square error**, an overall indicator of the error of the predicted values, is calculated, for each of the six time intervals:

Mean square of relative errors for time period *t*: $F_t = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n} f_k^2}$

where $f_k = \frac{\varepsilon_k(predicted) - \varepsilon_k(actual)}{\varepsilon_k(actual)} \cdot 100$

CEB mean square error: $F_{CEB} = \sqrt{\frac{1}{K}\sum_{t=1}^{K} F_t^2}$

Finally, the **CEB mean relative deviation**, an indicator of systematic over- or under-estimation by the model, is calculated for each of the six time intervals:

Deviation for time period *t*: $M_t = \frac{1}{n}\sum_{k=1}^{n} \frac{\varepsilon_k(\,predicted\,)}{\varepsilon_k(\,actual\,)}$

CEB mean relative deviation: $M_{CEB} = \frac{1}{K}\sum_{t=1}^{K} M_t$

A mean relative deviation close to 1 indicates a good model fit.

The disadvantage of these three measures is that the data are immediately grouped into drying time intervals, which means that, although time intervals are weighted equally, experiments are not weighted equally in the calculation of the overall mean statistics. Bažant's approach, described above, meets both of these objectives. The use of the mean of the shrinkage values for a given time period ($\bar{\varepsilon}_t$) rather than the grand mean of all the shrinkage values in the calculation of the CEB coefficient of variation results in large values of V_t for short drying time data and vice versa. This in turn means that the calculated value of the overall coefficient of variation is over-influenced by the short drying time data. With respect to both the Bažant and CEB coefficients of variation, it should also be noted that, for the WITS model at least, the model fit was obtained by minimising the variation, not by minimising the coefficient of variance. The use of the relative errors in the calculation of the CEB mean square error is prone to over-emphasize the errors in low shrinkage values, which are not so important, and vice versa. The CEB mean relative deviation is a useful statistic as it identifies the magnitude and direction of bias in the predicted values.

The results from all the above calculations are given in Table 7.6. The CEB coefficient of variation statistics are shown in Figure 7.17. Results for only five time intervals are shown since there were no data at drying times longer than 1095 days in the data set used in this study. The WITS model exhibited the lowest overall CEB coefficient of variation, followed by the ACI 209R-92 and RILEM B3 models. Over the different time intervals, the WITS model had the lowest coefficient of variation, except for the longest drying time interval (731-1095 days) where it was superseded by the SANS 10100-1 model. It appeared that the SANS 10100-1 model was quite a good predictor of shrinkage at longer drying times, but under-predicted at shorter drying times, perhaps due to inadequate prediction of the 6-month shrinkage, as illustrated in Figure 7.18 for an experiment which includes data at very long drying times.

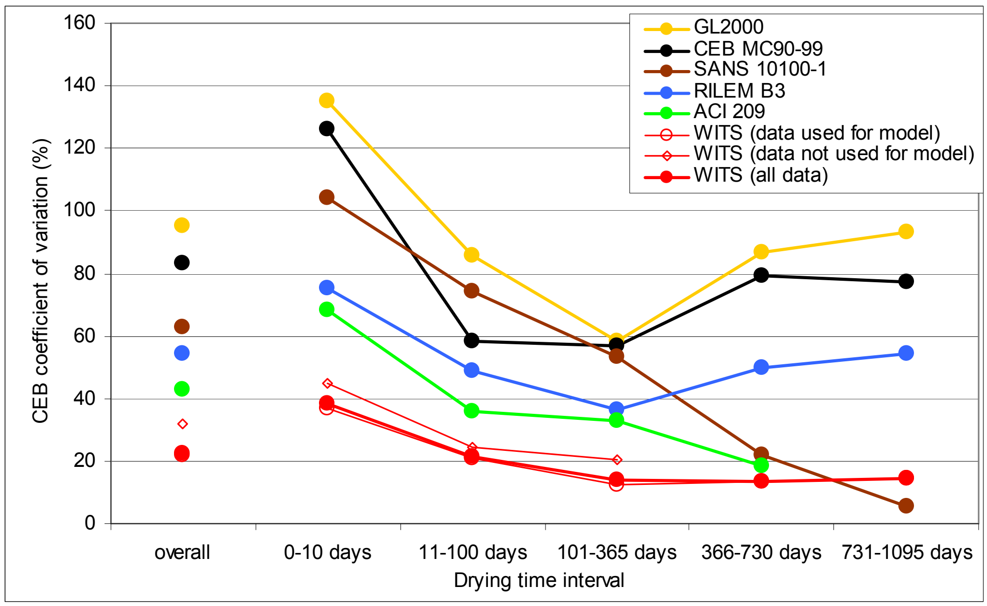


Figure 7.17 The CEB coefficient of variation: overall and for different drying time intervals

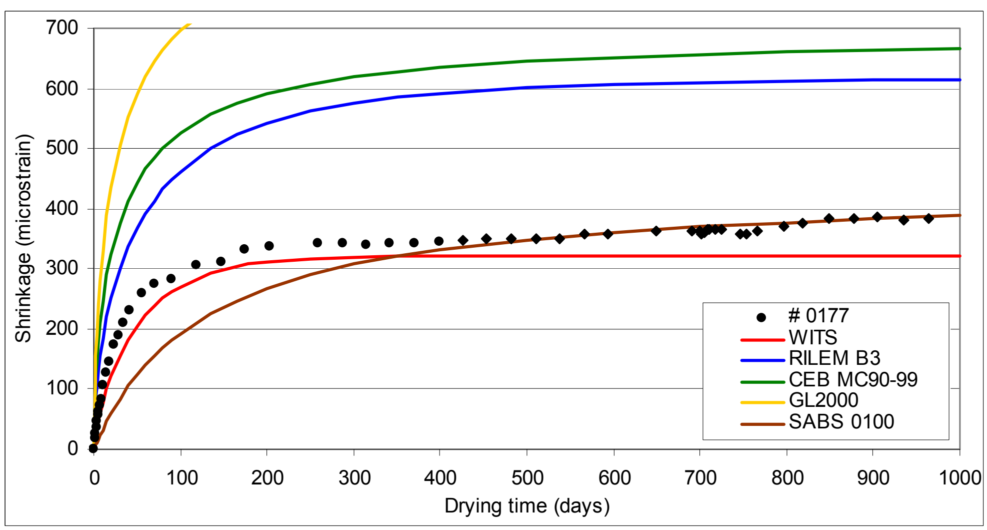


Figure 7.18 Illustration of the fit of the models to an experiment with data at long drying times

The CEB mean square error statistics are shown in Figure 7.19. The over-emphasis of the errors in low shrinkage values (and vice versa), as discussed above, can clearly be seen. The WITS model, closely followed by the ACI 209R-92 and SANS 10100-1 models, performed better than the other models on this measure.

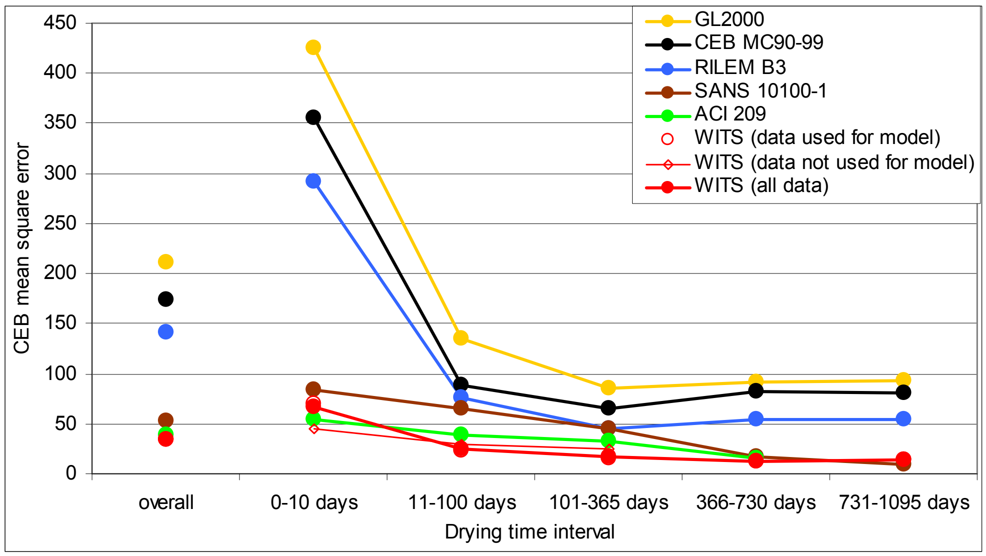


Figure 7.19 The CEB mean square error: overall and for different drying time intervals

The CEB mean relative deviation statistics are shown in Figure 7.20. The magnitude and direction of the bias in the different models (discussed previously), as a function of drying time, can be seen clearly. The WITS model showed the least bias across all time intervals, closely followed by the ACI 209R-92 model. The narrowing difference between the actual and predicted values of the SANS 10100-1 model with increasing drying time is shown clearly here, indicating that the longer-term (30-year) predictions of this model are more in line with the measured data than the shorter-term (6-month) predictions.

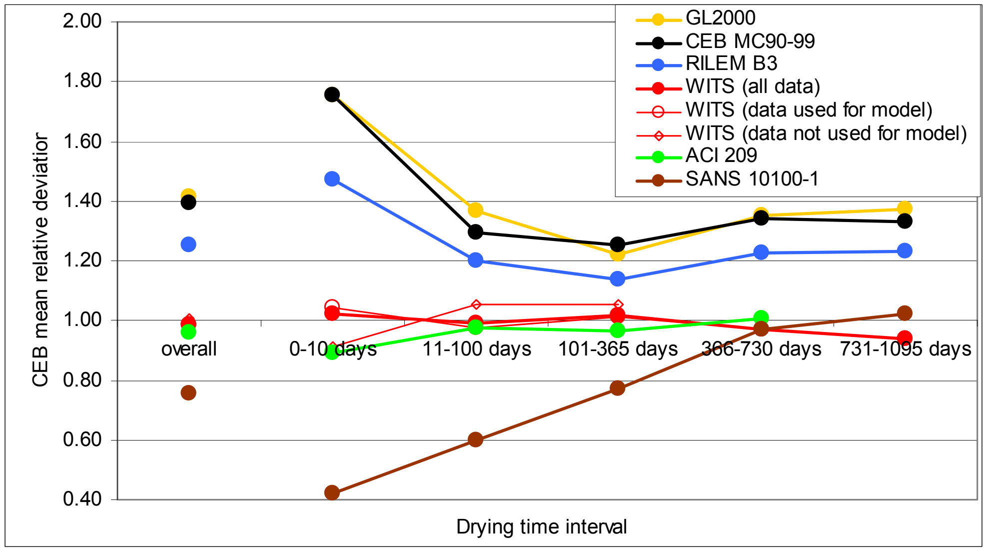


Figure 7.20 The CEB mean relative deviation: overall and for different drying time intervals

In the method used by **Gardner**[13], the observations are grouped into half-log decades (starting at a drying time of 3 days). Within each time interval, the average and the root mean square (RMS) values of the differences between the actual and predicted values (residuals) are calculated:

Average residual for time period *t*: $$\Delta_t = \frac{1}{n}\sum_{k=1}^{n}\Delta_k$$

RMS of residuals for time period *t*: $$S_t = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}\Delta_k^2}$$

The trend of the average residual with time will show whether there is over-or under-estimation (bias) with time by the model, while the trend of the RMS values with time will show if the deviation of the model increases with time.

Next, the RMS values are simply averaged. The average RMS values are then normalised by dividing by the average of the average shrinkage values (for the different time intervals) to produce a measure analogous to a **coefficient of variation**.

$$\bar{\varepsilon}_G = \frac{1}{K}\sum_{t=1}^{K}\bar{\varepsilon}_t \qquad S_G = \frac{1}{K}\sum_{t=1}^{K}S_t \qquad \omega_G = \frac{S_G}{\bar{\varepsilon}_G}$$

The formula for S_G given by Gardner is however incorrect, since the mean squares, not RMS values, should be averaged: $S_G = \sqrt{\frac{1}{K}\sum_{t=1}^{K} S_t^2}$. In this study, the corrected formula for S_G was applied. As was the case for the overall CEB statistical indicator, ω_G also suffers from the disadvantage that the data are immediately grouped into drying time intervals, which means that, although time intervals are weighted equally, experiments are not weighted equally in the calculation of this overall mean statistic.

The Gardner average residuals and the root mean squares of the residuals for the different time periods for all the models are shown in Figures 7.21 and 7.22 respectively. The mean residual plot showed the same information as in Figure 7.8 albeit in more illuminating detail as a result of the subdivision of the data into more time intervals. The WITS model had the lowest mean residuals, followed by the ACI 209R-92 model. The bias towards over- or under-prediction by the other models can clearly be seen. The Gardner coefficient of variation (Figure 7.22) also identified the WITS model as the best model according to this criterion, as did the root mean squares of the residuals for the different drying time intervals. The latter statistic showed that the deviations of all the models increased with time, with the exception of the WITS model and the SANS 10100-1 model, the latter at long drying times.

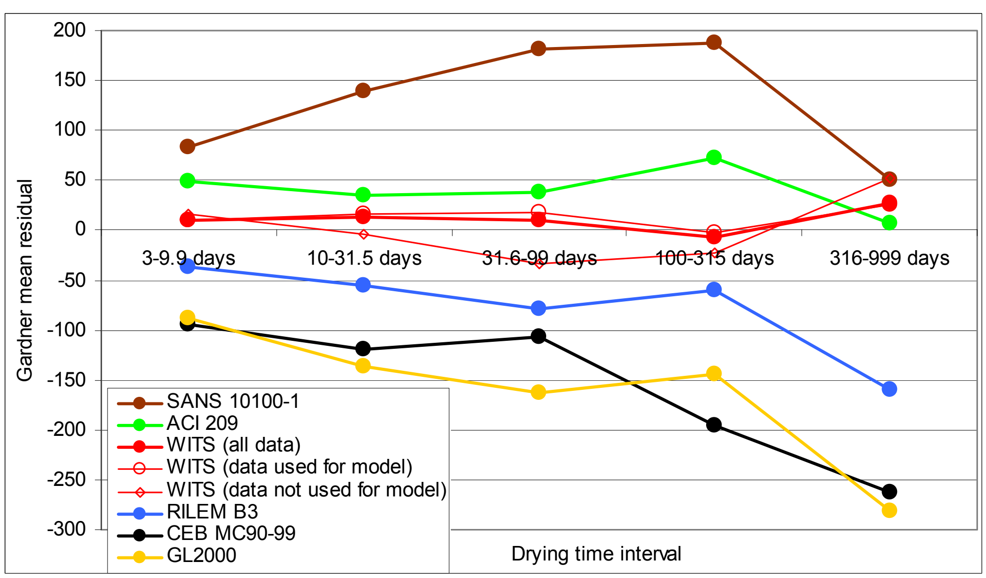


Figure 7.21 The Gardner average residuals for different drying time intervals

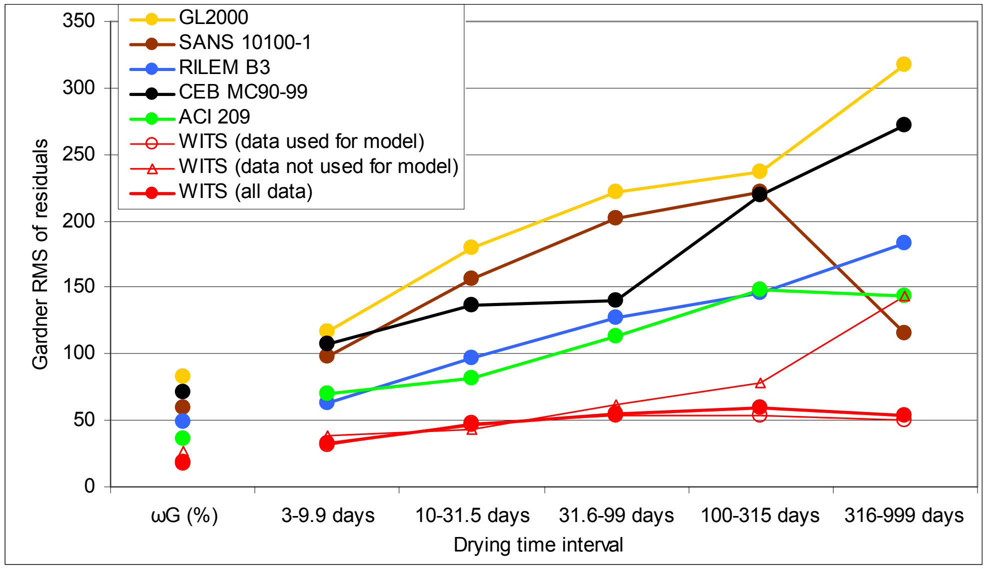


Figure 7.22 The Gardner coefficient of variation and root mean squares of the residuals for different drying time intervals

McDonald[(82)] gives an alternative formulation for the coefficient of variation, but this is considered to be inferior to the approaches outlined above. Goel[(83)] uses the ratio of the actual to predicted average shrinkage strain as well as the standard deviation of this ratio as a measure of goodness-of-fit but this is a very simplistic approach and does not convey as much information as the methods described above.

Vonesh[(50)] set out some properties which a goodness-of-fit statistic should possess. It should

1. have an intuitively reasonable interpretation,
2. be independent of the units of measurement,
3. have well defined endpoints corresponding to a perfect fit and a complete lack of fit,
4. be applicable to any type of model regardless of underlying distributional properties,
5. be comparable across different models fitted to the same data and
6. have positive and negative residuals weighted equally.

The mean absolute deviation, mean square prediction error, Gardner's average residuals and Gardner's RMS of the differences between actual and predicted values are not independent of the units of measurement. None of the statistical measures discussed above have defined endpoints corresponding to a complete lack of fit. In all other respects, the statistical measures discussed above exhibited the desired properties outlined by Vonesh.

The discussion above has shown that there is no one goodness-of-fit statistic which can adequately capture all aspects of the performance of a model for the shrinkage of concrete. In order to summarise the information contained in the different overall numerical goodness-of-fit measures discussed above, their calculated values for the different models are represented in Figure 7.23 by means of a radar chart. This plot showed that, across all the goodness-of-fit measures, the WITS model performed the best. This was perhaps to be expected since the model was derived from the data. However, even the performance of the model on the data which were not used in the development of the model was excellent, although, as was mentioned earlier, this small data set should not be regarded as a true validation data set (but on the other hand might be considered a worst-case validation set!). The ACI 209R-92 model exhibited the second-best performance, but here it must be noted that this model could be used to predict only 36% of the data set. The third best performance was shown by the RILEM B3 model, which, taken together with the fact that it was used to predict 69% of the data set, indicates that it is arguably the best alternative model to the WITS model for the South African data set. The SANS 10100-1 model performed poorly relative to the other models, but as discussed above, its predictive ability at longer drying times was better than that at shorter drying times. This may indicate that its 30-year predictions are more suited to the South African data set than its 6-month predictions, i.e. that the time development of shrinkage of the model may need to be adjusted relative to the British Standard from which the model was directly taken.

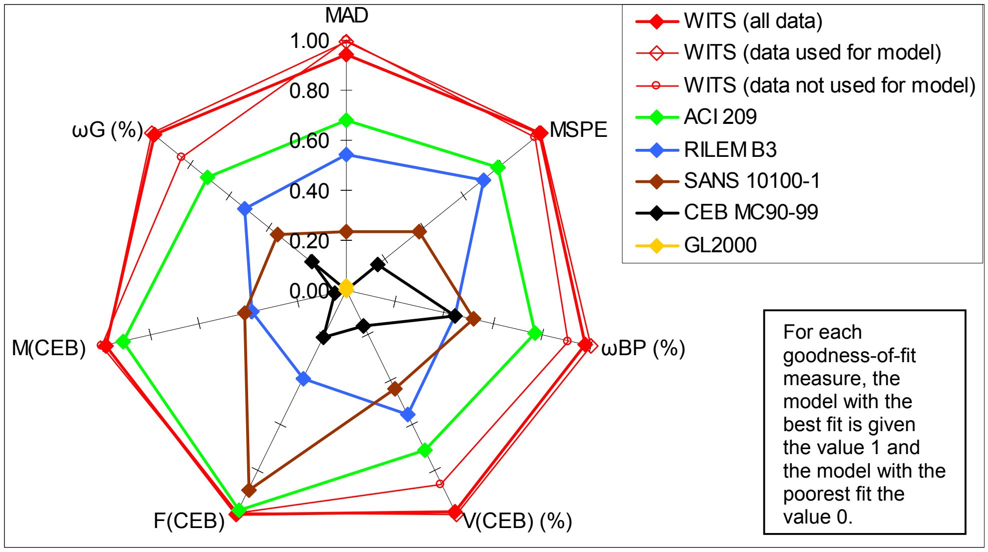


Figure 7.23 Comparison of the goodness-of-fit measures for the different concrete shrinkage models

In this chapter, five published models for concrete shrinkage were presented in detail. These models were compared to the WITS model regarding their predictive ability with respect to the South African data set used in this study. A wide variety of graphical and numerical goodness-of-fit measures were used to assess the relative performance of the models. It was found that the WITS model performed the best on balance across the different goodness-of-fit criteria. This was perhaps to be expected since the model was derived from the data. However, even the performance of the model on the data which were not used in the development of the model was excellent. The ACI 209R-92 model exhibited the second-best performance, but this model could be used to predict only just over one-third of the data set. The third best performance was shown by the RILEM B3 model. This model could be used to predict just over two-thirds of the data set and was thus arguably the best alternative model to the WITS model for the South African data set. The SANS 10100-1 model performed poorly relative to the other models. However, it appeared that its predictive ability at longer drying times was better than that at shorter drying times which may indicate that the time development of shrinkage of the model may need to be adjusted relative to the British Standard from which the model was taken.

8 Conclusions

The **aim of this study** was to **consider the possibility that a general approach towards modelling the shrinkage profiles of concrete could be found and to apply it to historical laboratory data on South African concretes.** The different steps in the research were:

- compiling a database of all the available results of laboratory studies on concrete shrinkage carried out in South Africa,
- identifying the most appropriate growth curve model for individual shrinkage profiles and
- modelling the parameters of this growth curve model, fitted to each shrinkage profile individually, in terms of suitable covariates, viz. the composition of the concrete, its other engineering properties as well as shrinkage test conditions.

A final aim of the study was to compare the model developed in this study to several models for shrinkage already in use in the concrete industry, with respect to the data in the South African database.

In Chapter 2, the study was firstly placed into context by introducing existing models for the prediction of concrete shrinkage, their inherent shortcomings when applied to South African data, and attempts by South African researchers to adapt these models to improve their predictive ability for local data. The factors known to affect concrete shrinkage were reviewed.

The sources of the data were described in Chapter 3 and the issues associated with the analysis of historical data were highlighted. A detailed description of the covariate and shrinkage data was given, along with the steps taken to prepare the data for analysis. An analysis of comparable experiments within and between individual studies making up the overall data set was conducted. It was found that the overall data set was sufficiently consistent for analysis as a whole, but that the laboratory in which the specimens were tested was confounded with other covariates specifying the concrete raw materials.

In Chapter 4, the requirements for a reliable shrinkage model were discussed relative to what is possible to achieve with the data set used in this study. The nonlinear hierarchical modelling approach was presented in detail and it was shown how this technique suited the nature of the data and the modelling problem.

The first step in the hierarchical modelling process was to choose the most appropriate growth curve for modelling the nonlinear shrinkage-time relationship (shrinkage profile). In the second step, the parameters of this nonlinear growth curve, fitted to each shrinkage profile, became the dependent variables in a multiple regression model to form a link between the covariates and the shrinkage profiles.

In Chapter 5, candidate growth curve models, thought to be suitable for modelling the shrinkage-time relationship, were reviewed. Then, the heterogeneous variance of the response variable (shrinkage) was modelled in order to define a weighting scheme for the shrinkage-time data. This weighting scheme was subsequently evaluated to determine the best nonlinear least squares fitting method appropriate to the data. It was found that OLS sufficed for fitting of the growth curve models to the shrinkage-time data. The autocorrelation of the shrinkage-time data was also considered. It was decided to assume an uncorrelated intra-experiment error structure which was subsequently found to be a reasonable assumption. Ten "key" shrinkage profiles for the evaluation of fifteen candidate growth curve models were selected and used to evaluate these models. MCDA was used to identify four models with the best fit for the "key" shrinkage profiles. The four models were then fitted to all the shrinkage profiles in order to identify the best growth curve model, chosen on the basis of practical considerations as well as the application of MCDA.

The second stage in the hierarchical non-linear modelling process was presented in Chapter 6. The three parameters of the chosen growth curve model were to be modelled in terms of the covariates (which described the concrete raw materials, concrete composition and the shrinkage testing conditions) by multivariate multiple regression. Preparation of both the dependent variables and the covariates for regression analysis was outlined. Then the large number of covariates were divided into meaningful groups for variable selection during the regression step, in order to avoid multicollinearity. Thereafter, the multivariate multiple regressions were carried out, during which different weighting schemes for the cases (experiments) were explored. MCDA was used to identify the best regression model, using appropriate goodness-of-fit measures. The best model (described in Figure 7.6), now known as the WITS model, included covariates relating to concrete raw materials (sand, stone and cement type), concrete composition (water content, stone content and aggregate/binder ratio) and testing conditions (temperature and volume to surface area ratio of the specimen). The model was described in terms of its range of applicability, significant terms and a 95% confidence interval was derived for the predicted shrinkage profiles. The model was interpreted in terms of what was already known about the factors affecting concrete shrinkage and was largely found to conform to expectations. Finally, the findings of a sensitivity analysis to investigate the effect of changes in the input parameters on the predicted growth curve parameters were presented.

In Chapter 7, five published models for concrete shrinkage were presented in detail. These models were compared to the WITS model regarding their predictive ability with respect to the South African data set used in this study. A wide variety of graphical and numerical goodness-of-fit measures were used to assess the relative performance of the models. It was found that the WITS model performed the best across the different goodness-of-fit criteria. The ACI 209R-92 model exhibited the second-best performance, but this model could be used to predict only just over one-third of the data set. The third best performance was shown by the RILEM B3 model. This model could be used to predict just over two-thirds of the data set and was thus arguably the best alternative model to the WITS model for the South African data set. The SANS 10100-1 model performed poorly relative to the other models. However, it appeared that its predictive ability at longer drying times was better than that at shorter drying times which may indicate that the time development of shrinkage of the model may need to be adjusted relative to the British Standard from which the model was taken.

The aims of the study were thus achieved. The **importance of the study** is two-fold:

- The concept of hierarchical nonlinear modelling was applied for the first time to the modelling of concrete shrinkage. This approach could prove useful to other researchers seeking to model concrete shrinkage and related time-dependent properties such as creep.
- The model derived in this study is the first comprehensive model to bring together laboratory data on the shrinkage of concrete generated in South

Africa over a span of 30 years, identifying which covariates are the most important contributors to both the magnitude and rate of concrete shrinkage. Recognising the limitations of the study (outlined in Chapter 1), the model provides a starting point for further, statistically designed, experiments to explore the effect of key variables more fully.

Based on the findings of this study a number of **recommendations** are made for further work:

- Shrinkage data collection should, where possible, be continued as long as possible, at least until the shrinkage has approached its ultimate value (i.e. until the shrinkage has 'lined out'). This will reduce the amount of data which has to be discarded in future analyses.
- A 'round robin' study to determine the differences between shrinkage tests carried out in different laboratories should be carried out. This should include specimens prepared in different laboratories (using the same raw materials).
- A 'control' concrete mix should be prepared and tested as part of each study, so that the long-term reproducibility of shrinkage specimen preparation and testing can be assessed. This 'control' mix should also be tested several times in a study to establish the short-term reproducibility of the shrinkage specimen preparation and testing.
- In future studies, information on covariates which currently suffer from large amounts of missing data (viz. the stone and sand numerical properties, the 28-day compressive strength and the 28-day elastic modulus of the concrete as well as the slump of the concrete) should be gathered where at all possible. Future modelling studies should re-evaluate the possible importance of these covariates once more data has become available.
- A series of designed experiments should be carried out to validate the developed model which will cover the range of the model in a balanced way. Key variables, such as humidity and temperature, should not be controlled at fixed values but should rather be included in the experimental design so that their effects on shrinkage may be measured.
- Exploration of the following covariates should be considered in future designed experiments to develop the WITS model further (these covariates were used in one or more of the published models, but could not be included in the WITS model, as a result of high levels of missing data or the limited range covered by the covariate):

- humidity,
- concrete properties (28-day compressive strength, 28-day elastic modulus and slump),
- curing method
- age at first drying and
- specimen shape.

- The WITS model should be developed further using other, preferably international, data sets.

References

1 Addis, B.J. and Owens, G. ed. *Fundamentals of Concrete*, South Africa: Cement and Concrete Institute, 2005, pp 35-49.

2 Owens, G. ed. *Fulton's Concrete Technology,* 9th ed. South Africa: Cement and Concrete Institute, 2009, pp. 131-138.

3 American Concrete Institute, *Guide for Modeling and Calculating Shrinkage and Creep in Hardened Concrete*, Detroit: American Concrete Institute, 2008 (ACI 209.2R-08).

4 Bažant, Z.P. Prediction of concrete creep and shrinkage: past, present and future, *Nuclear Engineering and Design*, vol. 203, 2001, pp. 27–38.

5 Kovler, K. and Zhutovsky, S. Overview and future trends of shrinkage research, *Materials and Structures*, vol. 39, 2006, pp. 827–847.

6 American Society of Testing Materials, *Standard Test Method for Length Change of Hardened Hydraulic Cement – Mortar and Concrete,* Philadelphia: American Society of Testing Materials, 1999 (ASTM C157/C157M-99).

7 South African National Standard, *Concrete Tests — Initial drying shrinkage and wetting expansion of concrete*; Pretoria: South African Bureau of Standards, 2006 (SANS 6085:2006).

8 Davidian, M. and Giltinan D.M. *Nonlinear Models for Repeated Measurement Data*, 1st ed. London: Chapman & Hall, 1995.

9 American Concrete Institute, *Prediction of creep, shrinkage and temperature effects in concrete structures, in "Designing for creep and shrinkage in concrete structures"*, Detroit: American Concrete Institute, 1982 (ACI 209R-82, pp. 193-300).

10 Bažant, Z.P. and Baweja, S. Creep and shrinkage model for analysis and design of concrete structures - model B3, *Materials and Structures*, vol. 28, 1995, pp. 357–365 with errata in vol. 29, 1996, p. 126.

11 Bažant, Z.P. Creep and shrinkage prediction model for analysis and design of concrete structures: Model B3, *Adam Neville Symposium: Creep and Shrinkage - Structural Design Effects,* 2000, pp.1-84.

12 Federation Internationale du Beton, *Structural Concrete – Textbook on Behaviour, Design and Performance. Updated Knowledge of the CEB/FIP Model Code 1990,* Lausanne, Switzerland: Federation Internationale du Beton, 1999 (FIB Bulletin, vol. 2, pp. 37-52).

13 Gardner, N.J. Comparison of prediction provisions for drying shrinkage and creep of normal-strength concretes, *Canadian Journal of Civil Engineering*, vol. 31(5), 2004, pp. 767-775.

14 South African National Standard, *The structural use of concrete, Part 1: Design*; Pretoria: South African Bureau of Standards, 2000 (SANS 10100-1).

15 Al-Manaseer, A. Predicting drying shrinkage of concrete, *Concrete International,* vol. 26(8), 2004, pp. 79-83.

16 Al-Sugair, F.H. and Almudaiheem, J.A. Further modification of the Ross equation to predict the ultimate drying shrinkage of concrete, *ACI Materials Journal*, vol. 82(2), 1990, pp. 237-240.

17 Basma, A.A, and Abdel-Jawad, Y. Probability model for the drying shrinkage of concrete, *ACI Materials Journal,* vol. 92(3), 1995, pp.246-251.

18 Bažant, Z.P. and Kim, J.-K. Improved prediction model for time dependent deformations of concrete Part 1: Shrinkage, *Materials and Structures*, vol. 24, 1991, pp. 327-345.

19 Eguchi, K. and Teranishi, K. Prediction equation of drying shrinkage of concrete based on composite model, *Cement and Concrete Research,* vol. 35(3), 2005, pp. 483-493.

20 Gardner, N.J. and Zhao, J.W. Creep and shrinkage revisited, *ACI Materials Journa*l, vol. 90, 1993, pp. 236–246.

21 Han, M.Y. Theoretical prediction of drying shrinkage of concrete, *Journal of Materials in Civil Engineering,* vol. 7(4), 1995, pp. 204-207.

22 Comité Euro-International du Béton, *CEB-FIP Model Code 1990,* Lausanne, Switzerland: Comité Euro-International du Béton, 1990 (CEB Bulletin d'Information No. 213.214).

23 Comité Euro-International du Béton, *Model Code for concrete structures*, Paris: Comité Euro-International du Béton, 1978 (Volume 2, Appendix E).

24 Bažant, Z.P. and Li, G.-H. Comprehensive database on concrete creep and shrinkage, *ACI Materials Journal,* vol. 105(6), 2008, pp.635-637.

25 Marques, E.T. *Models for Predicting Creep and Shrinkage of Concrete Using South African Aggregates,* MSc Dissertation, University of Pretoria, 1992.

26 British Standards Institution, *Structural Use of Concrete, Part 2: Code of Practice for Design and Construction*; London: British Standards Institution, 1985.

27 Alexander, M. G. *Properties of Aggregates in Concrete: Report on Phase 2 Testing of Concretes Made With Aggregates From a Further 10 Different Quarries, and Associated Design Recommendations*, South Africa: Department of Civil Engineering, University of Cape Town, March 1993.

28 Ballim, Y. The effect of shale in quartzite aggregate on the creep and shrinkage of concrete - A comparison with RILEM Model B3, *Materials and Structures*, vol. 33, 2000, pp. 235-242.

29 Ballim, Y. Localising international concrete models - the case of creep and shrinkage prediction, *Fifth International Conference on Concrete Technology for Developing Countries*, New Delhi, 17-19 November 1999.

30 Badenhorst, S. *A critical review of review of concrete drying shrinkage: test methods, prediction models and factors influencing shrinkage*, MSc (Eng) Dissertation, University of the Witwatersrand, Johannesburg, 2003.

31 Videla, C., Covarrubias, J.P. and Masana, C. Updating concrete drying-shrinkage prediction models for local materials, *ACI Materials Journal*, vol. 101(3), 2004, pp. 187-198.

32 Addis, B.J. ed. *Fulton's Concrete Technology*, 6th ed. South Africa: Portland Cement Institute (now Cement and Concrete Institute), 1986.

33 Alexander, M. and Mindess, S. *Aggregates in Concrete,* United Kingdom: Taylor & Francis, 2005.

34 Tia, M., Subramanian, R., Brown, D. and Broward, C. *Evaluation of Shrinkage Cracking Potential of Concrete Used in Bridge Decks in Florida*, Florida: Florida Department of Transportation, 2005 (Final Report BC 354 RPWO #26).

35 Ballim, Y. *The Concrete Making Properties of the Andesite Lavas from the Langgeleven Formation of the Ventersdorp Supergrou*p, MSc Dissertation, University of the Witwatersrand, Johannesburg, 1983.

36 Davis, D.E. and Alexander, M.G. *Properties of Aggregates in Concrete (Part 2),* South Africa: Hippo Quarries Technical Publication, October 1994.

37 Mehta, P.K. and Monteiro, P.J.M. *Concrete: Microstructure, Properties and Materials*, 3rd ed. USA: McGraw-Hill, 2006.

38 South African Bureau of Standards, *Portland cement (ordinary, rapid hardening, and sulphate resisting);* Pretoria: South African Bureau of Standards, 1971 (SABS 471–1971).

39 European Committee for Standardization, *Cement Part 1: Composition, Specifications and Conformity Criteria for Common Cements*; Brussels: European Committee for Standardization , 1996 (EN 197-1).

40 South African Bureau of Standards, *Cement Part 1: Composition, Specification and Conformity Criteria for Common Cements*; Pretoria: South African Bureau of Standards, 1994 (SANS 50197-1).

41 Holcim South Africa, *Holcim Materials Handbook,* South Africa: Holcim South Africa, date of publication unknown.

42 American Society of Testing Materials, *Standard Test Method for Length Change of Hardened Hydraulic Cement – Mortar and Concrete,* Philadelphia: American Society of Testing Materials, 1999 (ASTM C157/C157M-99).

43 Weiss, R.E. *Modeling Longitudinal Data,* 1st ed. USA: Springer, 2005.

44 RILEM Technical Committee TC-107, Guidelines for characterizing concrete creep and shrinkage in structural design codes or recommendations, *Materials and Structures*, vol. 28, 1995, pp. 52–55.

45 Bažant Z.P. Criteria for rational prediction of creep and shrinkage of concrete, *Adam Neville Symposium: Creep and Shrinkage - Structural Design Effects,* 2000, pp.237-260.

46 Davidian, M. and Giltinan D.M. Nonlinear models for repeated measurement data: an overview and update, *Journal of Agricultural, Biological, and Environmental Statistics*, vol. 8(4), Dec. 2003, pp. 387-419.

47 Yeap, B.Y., Catalano, P.J., Ryan, L.M. and Davidian, M. Robust two-stage approach to repeated measurements analysis of chronic ozone exposure in rats, *Journal of Agricultural, Biological and Environmental Statistics*, vol. 8(4), Dec. 2003, pp. 438-454.

48 Nummi, T. Analysis of growth curves under measurement errors, *Journal of Applied Statistics,* vol. 27(2), 2000, pp. 235-243.

49 Ratkowsky, D.A. *Handbook of Nonlinear Regression Models,* 1st ed. New York: Marcel Dekker, 1990.

50 Vonesh, E.F. and Carter, R.L. Mixed-effects nonlinear regression for unbalanced repeated measures, *Biometrics,* vol. 48, 1992, pp. 1-17.

51 Garcia, O. Visualization of a general family of growth functions and probability distributions – The Growth-curve Explorer, *Environmental Modelling & Software,* vol. 23, 2008, pp. 1474–1475.

52 Piegorsch, W.W. and Bailer, A.J. *Analyzing Environmental Data,* 1st ed. Chichester: John Wiley & Sons, 2005.

53 British Standards Institution, *Common Rules for Buildings and Civil Engineering Structures*, London: British Standards Institution, 2004 (BS EN 1992-1-1:2004(E)).

54 StatSoft, Inc., STATISTICA (data analysis software system), version 7.1. www.statsoft.com (Notes on nonlinear regression in Help documentation), StatSoft, Inc., 2005.

55 Hair, J.F., Anderson R.E., Tatham R.L. and Black W.C. *Multivariate Data Analysis,* 5th ed. New Jersey: Prentice-Hall Inc., 1998.

56 Erdman, D. and Little, M. Nonlinear regression analysis and nonlinear simulation models, SAS Institute Inc., Cary, NC, USA, http://support.sas.com/rnd/app/papers/nonlinearmodels.pdf, accessed on 20 November 2008.

57 Burnham, K.P. and Anderson, D.R. *Model selection and Multimodel Inference: A Practical Information-Theoretic Approach,* 2nd ed. New York: Springer, 2002.

58 SAS Institute Inc., *SAS Software, version 9.1.3 for Windows,* Cary, NC, USA: SAS Institute Inc., 2002-2003.

59 UCLA Academic Technology Services, *SAS library: nonlinear regression in SAS,* http://www.ats.ucla.edu/stat/sas/library/SASNLin_os.htm, accessed on 20 November 2008.

60 Vonesh, E.F., Chinchilli, V.M. and Pu, K. Goodness-of-fit in generalized nonlinear mixed-effects models, *Biometrics,* vol. 52, 1996, pp. 572-587.

61 Motulsky, H.J., and Ransnas, H.J. Fitting curves to data using nonlinear regression: a practical and nonmathematical review, *FASEB Journal*, vol. 1, 1987, pp. 365-374.

62 Belton, V. and Stewart, T.J. *Multiple Criteria Decision Analysis,* Boston: Kluwer Academic Publishers, 2002.

63 Roy B. *Multicriteria Methodology for Decision Aiding*, Dordrecht: Kluwer Academic Publishers, 1996.

64 Johnson, R.A. and Wichern, D.W. *Applied Multivariate Statistical Analysis,* 2nd ed., New Jersey: Prentice Hall, 1988.

65 Rencher, A.C. *Multivariate Statistical Inference and Applications*, USA: John Wiley & Sons, Inc., 1998.

66 Cox, T.F. *An Introduction to Multivariate Data Analysis*, London: Hodder Arnold, 2005.

67 Ballim, Y. and Alexander, M.G. Personal communication.

68 Al-Subaihi, A.A. Variable selection in multivariable regression using SAS/IML, *American Statistical Association Journal of Statistical Software* vol. 7(12), 2002.

69 Draper, N.R. and Smith, H. *Applied Regression Analysis,* 2nd ed. New York: John Wiley & Sons, 1981.

70 StatSoft, Inc., STATISTICA (data analysis software system), version 7.1. www.statsoft.com, 2005.

71 Alexander, M.G. Deformation properties of blended cement concretes containing blast furnace slag and condensed silica fume, *Advances in Cement Research,* vol. 6(22), 1994, pp. 73-81.

72 Kim, K. and Timm, N. *Univariate and Multivariate General Linear Models*, Boca Raton, Florida: Chapman & Hall / CRC, 2007.

73 McQuarrie, A.D.R. and Tsai, C.-L., *Regression and Time Series Model Selection*, Singapore: World Scientific, 1998.

74 Weisberg, S. *Applied Linear Regression*, Hoboken, New Jersey: Wiley-Interscience, 2005.

75 Chernick, M.R. *Bootstrap Methods: A Guide for Practitioners and Researchers*, 2nd ed. Hoboken, New Jersey: John Wiley & Sons, Inc., 2008.

76 Law, A.M. and Kelton W.D. *Simulation Modeling and Analysis*, 2nd ed. New York: McGraw-Hill, Inc., 1991.

77 Ballim, Y. Personal communication.

78 Imamoto, K. and Arai, M. Specific surface area of aggregate and its relation to concrete drying shrinkage, *Materials and Structures*, vol. 41, 2008, pp. 323-333.

79 Gilbert, R.I. *Time Effects in Concrete Structures,* Amsterdam: Elsevier, 1988.

80 Bažant Z.P. and Baweja S. Justification and refinement of Model B3 for concrete creep and shrinkage. 1. Statistics and sensitivity, *Materials and Structures*, vol. 28, 1995, pp. 415–430.

81 Al-Manaseer, A. and Lam, J.-P. Statistical Evaluation of Shrinkage and Creep Models, *ACI Materials Journal,* vol. 102(3), 2005, pp. 170-176.

82 McDonald, D.B. and Roper, H. Accuracy of prediction models for shrinkage of concrete, *ACI Materials Journal,* vol. 90(3), 1993, pp. 265-271.

83 Goel, R., Kumar, R. and Paul, D.K. Comparative study of various creep and shrinkage prediction models for concrete, *Journal of Materials in Civil Engineering*, Mar. 2007, pp. 249-260.

Appendices

Appendix A
Partial listing of data set

Table A.1 gives a listing of the sources of the 25 data sets used in this study.

Table A.2 gives a partial listing (first 18 experiments) of the data set, starting with the covariates and then giving the shrinkage data. The full data set may be found on the accompanying CD in the spreadsheet: “Data – concrete.xls”.

Table A.1 Sources of the 25 data sets included in the study

Data Set	Reference	Year of study	Number of experiments
1	Ballim, Y. The effect of shale in quartzite aggregate on the creep and shrinkage of concrete - A comparison with RILEM Model B3, *Materials and Structures*, vol. 33, 2000, pp 235-242. Ballim, Y. Localising international concrete models - the case of creep and shrinkage prediction, *Fifth International Conference on Concrete Technology for Developing Countries*, New Delhi, 17-19 November 1999.	1997	10
2	Theodosiou, G. *Performance of Selected Cements in Concrete*, South Africa: Cement and Concrete Institute, 26 July 2001 (C&CI Technical Report).	2000	20
3	Theodosiou, G. *Performance of Selected Cements in Concrete*, South Africa: Cement and Concrete Institute, 26 July 2001 (C&CI Technical Report).	2001	24
4	Data files held at the University of the Witwatersrand, Johannesburg. The work was done for the Cement and Concrete Institute but the corresponding final report could not be traced.	2002	22
5	Goodman, H.J. *Report on 2004 Cement Performance Project*, South Africa: Cement and Concrete Institute, 20 July 2005 (C&CI Technical Report).	2004	22
6	Data files held at the University of the Witwatersrand, Johannesburg. The work was done for the Cement and Concrete Institute but the corresponding final report could not be traced.	1995	4
7	Ballim, Y. *The Concrete Making Properties of the Andesite Lavas from the Langgeleven Formation of the Ventersdorp Supergroup*, MSc Dissertation, University of the Witwatersrand, Johannesburg, 1983. Alexander, M. G. and Ballim, Y. *The concrete-making properties of andesite aggregates*, The Civil Engineer in South Africa, vol. 28(2), 1986, pp. 49–57.	1983	6
8	Jaufeerally, H. *Performance and Properties of Structural Concrete Made with Corex Slag*, MSc Dissertation, University of Cape Town, Cape Town, 2001. Jaufeerally, H., Mackechnie, J. R. and Alexander, M.G. Creep, shrinkage and durability properties of Corex slag concrete, *Creep, Shrinkage and Durability Mechanics of Concrete and Other Quasi-Brittle Materials. Proceedings of the Sixth International Conference, CONREEP-6@MIT*, Cambridge, USA, Aug. 2001.	2001	9
9	Alexander, M. G. *Properties of Aggregates in Concrete: Report on Phase 1 Testing of Concretes Made With Aggregates From 13 Different Quarries, and Associated Design Recommendations,* South Africa: Department of Civil Engineering, University of the Witwatersrand, August 1990. Alexander, M. G. *Properties of Aggregates in Concrete: Report on Phase 2 Testing of Concretes Made With Aggregates From a Further 10 Different Quarries, and Associated Design Recommendations*, South Africa: Department of Civil Engineering, University of Cape Town, March 1993.	1993	21
10	Booyse, D. *Shrinkage and Curing of High Performance Concrete*, D Booyse, Fourth-year Civil Engineering Thesis, University of Cape Town, Cape Town, 2000.	2000	4
11	Data file held at the University of the Witwatersrand, Johannesburg. Publication or dissertation could not be traced. Work carried out by H. Jaufeerally.	1999	4

Table A.1 Sources of the 25 data sets included in the study (continued)

Data Set	Reference	Year of study	Number of experiments
12	Alexander, M. G. *Properties of Aggregates in Concrete: Report on Phase 1 Testing of Concretes Made With Aggregates From 13 Different Quarries, and Associated Design Recommendations,* South Africa: Department of Civil Engineering, University of the Witwatersrand, August 1990. Alexander, M. G. *Properties of Aggregates in Concrete: Report on Phase 2 Testing of Concretes Made With Aggregates From a Further 10 Different Quarries, and Associated Design Recommendations*, South Africa: Department of Civil Engineering, University of Cape Town, March 1993.	1990	26
13	Hoppe, G.E. *Creep, Shrinkage and Elastic Strains in a Pre-Stressed Concrete Bridge – Theoretical and Actual*, MSc Dissertation, University of Cape Town, Cape Town, 1981.	1978	1
14	Uys, A.J. *The Influence of Fly Ash on the Strength, Elastic Modulus, Shrinkage and Creep of Concrete Made with Cape Town Aggregates with Particular Reference to Alkali-Aggregate Reaction*, MSc Dissertation, University of Cape Town, Cape Town, 1983.	1979	8
15	Grieve, G. R. H. *The Influence of Two South African Fly Ashes on the Engineering Properties of Concrete*, PhD Thesis, University of the Witwatersrand, Johannesburg, 1991.	1990	6
16	Badenhorst, S. *A critical review of review of concrete drying shrinkage: test methods, prediction models and factors influencing shrinkage*, MSc (Eng) Dissertation, University of the Witwatersrand, Johannesburg, 2003.	2003	28
17	Fanourakis, G.C. *The Influence of Aggregate Stiffness on the Measured and Predicted Creep Behaviour of Concrete*, MSc Dissertation, University of the Witwatersrand, Johannesburg, 1998.	1998	6
18	Hsu, C.-W.P. *Alternative Aggregates for Concrete*, Fourth-year Civil Engineering Thesis, University of Cape Town, Cape Town, 2000.	2000	3
19	Confidential data source.	2007	30
20	Burmeister, N. R. *Evaluation of Shrinkage Prediction Models for Concrete*, Fourth-year Civil Engineering Thesis, University of Cape Town, Cape Town, 2007.	2007	18
21	Kutegeza, B. *Performance of Concrete made with Commercially Produced Recycled Coarse and Fine Aggregates in the Cape Peninsula*, MSc Dissertation, University of Cape Town, Cape Town, 2004.	2004	2
22	Alexander, M. G. *Deformation properties of blended cement concretes containing blast furnace slag and condensed silica fume*, Advances in Cement Research, vol. 6(22), 1994, pp. 73-81.	1994	12
23	Confidential data source.	2009	2
24	Confidential data source.	2006	1
25	Confidential data source.	2001	2

Table A.2 Partial listing of the data set (first 18 experiments)

Data Set	Notes	Experiment identification	Test Year	Cement					
				Cement Type as given in data source	CEM I / II / III	Supplementary cementitious material included in cement (%) derived from specification	Type of supplementary cementitious material	Strength Class	Early Strength Development
1		# 0001	1997	OPC	CEM I	0	-	str class not in use	str class not in use
1		# 0002	1997	OPC	CEM I	0	-	str class not in use	str class not in use
1		# 0003	1997	OPC	CEM I	0	-	str class not in use	str class not in use
1		# 0004	1997	OPC	CEM I	0	-	str class not in use	str class not in use
1		# 0005	1997	OPC	CEM I	0	-	str class not in use	str class not in use
1		# 0006	1997	OPC	CEM I	0	-	str class not in use	str class not in use
1		# 0007	1997	OPC	CEM I	0	-	str class not in use	str class not in use
1		# 0008	1997	OPC	CEM I	0	-	str class not in use	str class not in use
1		# 0009	1997	OPC	CEM I	0	-	str class not in use	str class not in use
1		# 0010	1997	OPC	CEM I	0	-	str class not in use	str class not in use
2		# 0011	2000	CEM I 52,5N Lichtenburg (Lafarge)	CEM I	0	-	52.5	N
2		# 0012	2000	CEM I 52,5N Lichtenburg (Lafarge)	CEM I	0	-	52.5	N
2		# 0013	2000	CEM II/A-L 32,5N Port Elizabeth (PPC)	CEM II	6-20	limestone	32.5	N
2		# 0014	2000	CEM II/A-L 32,5N Port Elizabeth (PPC)	CEM II	6-20	limestone	32.5	N
2		# 0015	2000	CEM I 42,5R Durban (NPC)	CEM I	0	-	42.5	R
2		# 0016	2000	CEM I 42,5R Durban (NPC)	CEM I	0	-	42.5	R
2		# 0017	2000	CEM II/A-V 32,5N Dudfield/Brakpan (Alpha)	CEM II	6-20	FA	32.5	N
2		# 0018	2000	CEM II/A-V 32,5N Dudfield/Brakpan (Alpha)	CEM II	6-20	FA	32.5	N

Table A.2 Partial listing of the data set (first 18 experiments) (continued)

Cement				Stone (black = value given in data source, blue = calculated value, red = value estimated from general data)							
Producer	**Milling / blending unit**	**Production unit**	**CEM type if separately added supplementary cementitious material included**	**Stone type as given in data source**	**Stone type (reference = andesite)**	**Stone source**	**Nominal stone size (mm)**	**Relative Density**	**Loose bulk density (kg/m^3)**	**Consolidated bulk density (kg/m^3)**	**Water absorption (%)**
Alpha	Roodepoort	Roodepoort	CEM III A	19 mm quartzite	Wits quartzite	Grootvlei (Springs)	19	2.67	1365	1490	0.23
Alpha	Roodepoort	Roodepoort	CEM III A	19 mm quartzite	Wits quartzite	Grootvlei (Springs)	19	2.67	1365	1490	0.23
Alpha	Roodepoort	Roodepoort	CEM III A	19 mm quartzite and shale	Wits quartzite/ shale 50/50	Grootvlei (Springs)	19	2.71			0.36
Alpha	Roodepoort	Roodepoort	CEM III A	19 mm quartzite and shale	Wits quartzite/ shale 50/50	Grootvlei (Springs)	19	2.71			0.36
Alpha	Roodepoort	Roodepoort	CEM III A	19 mm shale	shale	Grootvlei (Springs)	19	2.75			0.49
Alpha	Roodepoort	Roodepoort	CEM III A	19 mm shale	shale	Grootvlei (Springs)	19	2.75			0.49
Alpha	Roodepoort	Roodepoort	CEM III A	19 mm quartzite	Wits quartzite	Grootvlei (Springs)	19	2.67	1365	1490	0.23
Alpha	Roodepoort	Roodepoort	CEM III A	19 mm quartzite	Wits quartzite	Grootvlei (Springs)	19	2.67	1365	1490	0.23
Alpha	Roodepoort	Roodepoort	CEM III A	19 mm quartzite	Wits quartzite	Grootvlei (Springs)	19	2.67	1365	1490	0.23
Alpha	Roodepoort	Roodepoort	CEM III A	19 mm quartzite	Wits quartzite	Grootvlei (Springs)	19	2.67	1365	1490	0.23
Lafarge		Lichtenburg	CEM I	19 mm dolerite ex Lafarge, Rooikraal	dolerite	Rooikraal	19	2.98	1578	1693	0.53
Lafarge		Lichtenburg	CEM I	19 mm dolerite ex Lafarge, Rooikraal	dolerite	Rooikraal	19	2.98	1578	1693	0.53
PPC		Port Elizabeth	CEM II A-L	19 mm dolerite ex Lafarge, Rooikraal	dolerite	Rooikraal	19	2.98	1578	1693	0.53
PPC		Port Elizabeth	CEM II A-L	19 mm dolerite ex Lafarge, Rooikraal	dolerite	Rooikraal	19	2.98	1578	1693	0.53
NPC	Durban	Durban	CEM I	19 mm dolerite ex Lafarge, Rooikraal	dolerite	Rooikraal	19	2.98	1578	1693	0.53
NPC	Durban	Durban	CEM I	19 mm dolerite ex Lafarge, Rooikraal	dolerite	Rooikraal	19	2.98	1578	1693	0.53
Alpha	Brakpan	Dudfield	CEM II A-V	19 mm dolerite ex Lafarge, Rooikraal	dolerite	Rooikraal	19	2.98	1578	1693	0.53
Alpha	Brakpan	Dudfield	CEM II A-V	19 mm dolerite ex Lafarge, Rooikraal	dolerite	Rooikraal	19	2.98	1578	1693	0.53

Table A.2 Partial listing of the data set (first 18 experiments) (continued)

Stone (black = value given in data source, blue = calculated value, red = value estimated from general data)					Sand (black = value given in data source, blue = calculated value, red = value estimated from general data)					
Coefficient of thermal expansion ($x10^{-6}/^oC$)	Unconfined compression (MPa)	10% Fines Aggregate Crushing Value (kN)	Elastic Modulus (GPa)	Flakiness Index	Sand type as given in data source	Sand type (reference = dolomite)	Sand source	Relative Density	Loose bulk density (kg/m^3)	Consolidated bulk density (kg/m^3)
11.3	135	220	72	22	Quartzite	Wits quartzite	Grootvlei (Springs)	2.67		
11.3	135	220	72	22	Quartzite	Wits quartzite	Grootvlei (Springs)	2.67		
			86		Quartzite	Wits quartzite	Grootvlei (Springs)	2.67		
			86		Quartzite	Wits quartzite	Grootvlei (Springs)	2.67		
			99		Quartzite	Wits quartzite	Grootvlei (Springs)	2.67		
			99		Quartzite	Wits quartzite	Grootvlei (Springs)	2.67		
11.3	135	220	72	22	Quartzite and shale 50/50	Wits quartzite / shale 50/50	Grootvlei (Springs)	2.71		
11.3	135	220	72	22	Quartzite and shale 50/50	Wits quartzite / shale 50/50	Grootvlei (Springs)	2.71		
11.3	135	220	72	22	Shale	shale	Grootvlei (Springs)	2.75		
11.3	135	220	72	22	Shale	shale	Grootvlei (Springs)	2.75		
7.8	285	344	86	20	Dolomite ex Supersand ex Alpha, Olifantsfontein	dolomite	Olifantsfontein (Alpha)	2.85	1916	2111
7.8	285	344	86	20	Dolomite ex Supersand ex Alpha, Olifantsfontein	dolomite	Olifantsfontein (Alpha)	2.85	1916	2111
7.8	285	344	86	20	Dolomite ex Supersand ex Alpha, Olifantsfontein	dolomite	Olifantsfontein (Alpha)	2.85	1916	2111
7.8	285	344	86	20	Dolomite ex Supersand ex Alpha, Olifantsfontein	dolomite	Olifantsfontein (Alpha)	2.85	1916	2111
7.8	285	344	86	20	Dolomite ex Supersand ex Alpha, Olifantsfontein	dolomite	Olifantsfontein (Alpha)	2.85	1916	2111
7.8	285	344	86	20	Dolomite ex Supersand ex Alpha, Olifantsfontein	dolomite	Olifantsfontein (Alpha)	2.85	1916	2111
7.8	285	344	86	20	Dolomite ex Supersand ex Alpha, Olifantsfontein	dolomite	Olifantsfontein (Alpha)	2.85	1916	2111
7.8	285	344	86	20	Dolomite ex Supersand ex Alpha, Olifantsfontein	dolomite	Olifantsfontein (Alpha)	2.85	1916	2111

Table A.2 Partial listing of the data set (first 18 experiments) (continued)

Sand (black = value given in data source, blue = calculated value, red = value estimated from general data)				Concrete mix: quantities										
Fineness Modulus	Fines (< 75µm) (%)	Elastic Modulus (GPa)	Admixture type	Cement Content (kg/m^3)	Condensed Silica Fume Content (kg/m^3)	Fly Ash Content (kg/m^3)	GGBS (Slag) Content (kg/m^3)	GGCS (Corex Slag) Content (kg/m^3)	GGFS (Ferro-manganese Slag) Content (kg/m^3)	Limestone Content (kg/m^3)	Water Content (kg/m^3)	Sand Content (kg/m^3)	Stone Content (kg/m^3)	Admixture Content (kg/m^3)
3.15	6.3	72	-	146	0	0	146	0	0	0	195	926	1000	0
3.15	6.3	72	-	227	0	0	227	0	0	0	195	779	1000	0
3.15	6.3	72	-	146	0	0	146	0	0	0	195	926	1000	0
3.15	6.3	72	-	227	0	0	227	0	0	0	195	779	1000	0
3.15	6.3	72	-	146	0	0	146	0	0	0	195	926	1000	0
3.15	6.3	72	-	227	0	0	227	0	0	0	195	779	1000	0
		86	-	146	0	0	146	0	0	0	195	926	1000	0
		86	-	227	0	0	227	0	0	0	195	780	1000	0
		99	-	146	0	0	146	0	0	0	195	926	1000	0
		99	-	227	0	0	227	0	0	0	195	779	1000	0
2.93	16.4	114	-	524	0	0	0	0	0	0	220	690	1106	0
2.93	16.4	114	-	317	0	0	0	0	0	0	200	934	1106	0
2.93	16.4	114	-	435	0	0	0	0	0	65	210	719	1106	0
2.93	16.4	114	-	263	0	0	0	0	0	39	190	965	1106	0
2.93	16.4	114	-	536	0	0	0	0	0	0	225	665	1106	0
2.93	16.4	114	-	352	0	0	0	0	0	0	205	913	1106	0
2.93	16.4	114	-	425	0	63	0	0	0	0	205	742	1106	0
2.93	16.4	114	-	263	0	39	0	0	0	0	190	963	1106	0

Table A.2 Partial listing of the data set (first 18 experiments) (continued)

Concrete mix: ratios										**Test conditions**			
Paste Volume Fraction	**Condensed Silica Fume Fraction**	**Fly Ash Fraction**	**GGBS Fraction**	**GGCS Fraction**	**GGFS Fraction**	**Limestone Fraction**	**Water / Binder Ratio**	**Aggregate / Binder Ratio**	**Stone / Sand Ratio**	**Laboratory**	**Age at first drying (days)**	**Specimen Size (mm)** (square prism unless otherwise indicated)	**2 * volume / surface area of specimen (mm)**
0.292	0.00	0.00	0.50	0.00	0.00	0.00	0.67	6.6	1.1	Wits	28	100 x 100 x 200	50
0.346	0.00	0.00	0.50	0.00	0.00	0.00	0.43	3.9	1.3	Wits	28	100 x 100 x 200	50
0.292	0.00	0.00	0.50	0.00	0.00	0.00	0.67	6.6	1.1	Wits	28	100 x 100 x 200	50
0.346	0.00	0.00	0.50	0.00	0.00	0.00	0.43	3.9	1.3	Wits	28	100 x 100 x 200	50
0.292	0.00	0.00	0.50	0.00	0.00	0.00	0.67	6.6	1.1	Wits	28	100 x 100 x 200	50
0.346	0.00	0.00	0.50	0.00	0.00	0.00	0.43	3.9	1.3	Wits	28	100 x 100 x 200	50
0.292	0.00	0.00	0.50	0.00	0.00	0.00	0.67	6.6	1.1	Wits	28	100 x 100 x 200	50
0.346	0.00	0.00	0.50	0.00	0.00	0.00	0.43	3.9	1.3	Wits	28	100 x 100 x 200	50
0.292	0.00	0.00	0.50	0.00	0.00	0.00	0.67	6.6	1.1	Wits	28	100 x 100 x 200	50
0.346	0.00	0.00	0.50	0.00	0.00	0.00	0.43	3.9	1.3	Wits	28	100 x 100 x 200	50
0.387	0.00	0.00	0.00	0.00	0.00	0.00	0.42	3.4	1.6	Wits	31	100 x 100 x 200	50
0.301	0.00	0.00	0.00	0.00	0.00	0.00	0.63	6.4	1.2	Wits	31	100 x 100 x 200	50
0.373	0.00	0.00	0.00	0.00	0.00	0.13	0.42	3.7	1.5	Wits	28	100 x 100 x 200	50
0.288	0.00	0.00	0.00	0.00	0.00	0.13	0.63	6.9	1.1	Wits	28	100 x 100 x 200	50
0.396	0.00	0.00	0.00	0.00	0.00	0.00	0.42	3.3	1.7	Wits	30	100 x 100 x 200	50
0.317	0.00	0.00	0.00	0.00	0.00	0.00	0.58	5.7	1.2	Wits	30	100 x 100 x 200	50
0.368	0.00	0.13	0.00	0.00	0.00	0.00	0.42	3.8	1.5	Wits	29	100 x 100 x 200	50
0.291	0.00	0.13	0.00	0.00	0.00	0.00	0.63	6.9	1.1	Wits	28	100 x 100 x 200	50

Table A.2 Partial listing of the data set (first 18 experiments) (continued)

Test conditions			**Concrete properties**	
Average temperature (ºC)	**Average humidity (%)**	**Slump (mm)**	**28-day cube compressive strength (MPa)**	**28-day Elastic Modulus (GPa)**
23	65	80	29	31
23	65	55	50	32
23	65	75	28	27
23	65	60	47	31
23	65	75	30	25
23	65	50	49	27
23	65	50	31	26
23	65	35	51	26
23	65	25	32	24
23	65	10	50	
23	65	65	79	45
23	65	70	50	43
23	65	70	64	42
23	65	65	36	36
23	65	70	72	43
23	65	100	47	39
23	65	65	72	45
23	65	65	43	40

Table A.2 Partial listing of the data set (first 18 experiments) (continued)

Data set	Experiment identification	Measured shrinkage strain ($x10^{-6}$m/m) (time in days)																						
		0	1	2	3	4	5	6	7	8	9	10	12	14	15	18	21	25	27	28	34	56	57	84
1	# 0001	0	56	70	95	114			100	114	117	134	167	173			200	206				278		
1	# 0002	0	28	50	72	92			95	100	109	134	159	175			200	203				278		
1	# 0003	0	58	72	114	136			139	159	164	175	203	223			245	253				348		
1	# 0004	0	81	107	145	179			190	212	218	237	271	290			326	335				421		
1	# 0005	0	28	56	95		131		167					223		234				270			387	409
1	# 0006	0	17	109	159		206		237					289		317				362			465	482
1	# 0007	0	45	75	92		134		170					214		225				259			367	379
1	# 0008	0	22	61	84		139		175					223		234				267			367	379
1	# 0009	0	42	75	109		145	164	167						251				295		315	392		
1	# 0010	0	67	104	140		177	190	194						281				331		357	441		
2	# 0011	0	61	86	103				122					153			200			239		284		337
2	# 0012	0	31	47	84				92					103			136			189		225		278
2	# 0013	0	92	100	120				145					170			195			237		259		306
2	# 0014	0	31	33	53				64					72			89			125		150		198
2	# 0015	0		103	156				186					181			214			231		289		328
2	# 0016	0	95	128	150				178					192			223			262		317		370
2	# 0017	0	86	89	106				139					156			184			228		242		289
2	# 0018	0	58	64	70				81					86			109			150		178		220

Table A.2 Partial listing of the data set (first 18 experiments) (continued)

Data set	Experiment identification	Measured shrinkage strain ($x10^{-6}$m/m) (time in days)																							
		85	**88**	**112**	**140**	**141**	**142**	**148**	**161**	**163**	**168**	**170**	**196**	**198**	**199**	**227**	**252**	**253**	**281**	**282**	**290**	**308**	**338**	**339**	**364**
1	# 0001		303	362			401					412	415					462			490				487
1	# 0002		301	345			394					401	404					431			448				451
1	# 0003		376	426			456					482	487					543			560				562
1	# 0004		438	493			521					549	551					601			607				604
1	# 0005			454		473				518					562	585		579		615					
1	# 0006			534		554				590					635			640				643		643	
1	# 0007			431		448				490					537	551		559		568					
1	# 0008			412		431				468					504			500				503		504	
1	# 0009	431		482				537	551					607			601		646				649		
1	# 0010	468		501				558	568					595			585		615				618		
2	# 0011			345	370						409														
2	# 0012			298	317						356														
2	# 0013			340	362						392														
2	# 0014			223	239						267														
2	# 0015			365	406						459														
2	# 0016			398	470						518														
2	# 0017			317	331						362														
2	# 0018			237	256						289														

Appendix B

Comparable experiments in the data set

Comparable experiments (shrinkage profiles) were identified by matching cement, stone and sand type, followed by identifying experiments with similar water/binder ratio, aggregate/binder ratio, cement content, water content, sand content, stone content, stone/sand ratio, paste volume fraction and supplementary cementitious material fraction, in that order.

Table B.1 lists the comparable experiments identified within each of the 25 data sets making up the overall data set.

Table B.2 lists the comparable experiments identified between the 25 data sets.

Full graphical comparisons may be found in the spreadsheet file “Comparable experiments.xls” on the enclosed CD.

Table B.1 Comparable experiments within data sets

Comparison #	Matching criteria											
	Cement type	**StoneType**	**SandType**	**Water / binder ratio**	**Aggregate / binder ratio**	**Cement content (kg/m^3)**	**Water content (kg/m^3)**	**Sand content (kg/m^3)**	**Stone content (kg/m^3)**	**Stone / sand ratio**	**Paste volume fraction**	**Supplementary cementitious material fraction**
Comparisons within data sets												
# 1	CEM I	Greywacke	Cape Flats Sand	0.61-0.65	6.2	300	181-194	760-773	1100	1.4	0.278-0.290	0
# 2	CEM II B-V	Greywacke	Cape Flats sand	0.63	6.6-7.0	205-215	172-180	748-769	1150-1160	1.5	0.267-0.280	0.25
# 3	CEM II A-V	Greywacke	Cape Flats sand	0.64-0.65	6.8-7.1	230-240	174-183	754-778	1150	1.5	0.265-0.278	0.15
# 4	CEM I	Dolerite	Ecca grit	0.66-0.67	6.4-6.9	294-308	195-205	843-884	1090 and 1180	1.2-1.4	0.289-0.303	0
# 5	CEM II A-V	Dolerite	Ecca grit	0.63-0.71	6.3-7.0	236-268	198	868-878	1090-1100	1.3	0.291-0.303	0.15
# 6	CEM I	Wits Quartzite	Wits Quartzite / river sand	0.51-0.54	5.1-5.3	360-370	190-195	740-745	1150	1.5-1.6	0.308-0.310	0
# 7	CEM I	Wits Quartzite	Wits Quartzite / river sand	0.49-0.50	4.6-4.7	195-200	196	585-597	1250	2.1	0.325-0.329	0.5
# 8	CEM I	Wits Quartzite	Wits Quartzite / river sand	0.58	5.8	330	190	778-780	1150	1.5	0.295	0
# 9	CEM I	Wits Quartzite	Wits Quartzite / river sand	0.55-0.56	5.4-5.5	172-175	192	645-649	1250	1.9	0.306-0.308	0.5
# 10	CEM I	Andesite	Granite	0.42	3.3-3.5	500-524	210-220	674-721	1050	1.5-1.6	0.369-0.387	0
# 11	CEM I	Andesite	Granite	0.63	6.2-6.6	302-317	190-200	902-942	1050	1.1-1.2	0.286-0.301	0
# 12	CEM I	Greywacke	Pit sand Klipheuwel	0.40	3.5-4.1	500-525	200-210	630 and 775	1100-1400	1.7-1.8	0.359-0.377	0
# 13	CEM I	Greywacke	Pit sand Klipheuwel	0.50	5.2-5.6	380-400	190-200	735 and 868	1200-1400	1.4-1.9	0.311-0.327	0
# 14	CEM I	Greywacke	Pit sand Klipheuwel	0.60	7.5-7.6	267-283	160-170	733-939	1100-1400	1.2-1.9	0.245-0.260	0
# 15	CEM I	Greywacke	Pit sand Klipheuwel	0.60	6.2-6.4	317-350	190-210	1070-1249	900-1000	0.8	0.291-0.321	0

Table B.1 Comparable experiments within data sets (continued)

Comparison #	Admixture	Laboratory	Test year	Data set(s)	Number of experiments	Line # in data file	Comparison of shrinkage profiles	Relative standard deviation of final shrinkage value (%)
Comparisons within data sets								
# 1	1 with, 2 without	UCT	1979	14	3	176, 178, 179	Good	5.0
# 2	1 with, 1 without	UCT	1979	14	3	177, 182, 183	Quite variable	15.3
# 3	1 with, 1 without	UCT	1979	14	2	180, 181	Good	
# 4	-	Wits	1990	15	2	184, 185	Good	
# 5	-	Wits	1990	15	2	187, 188	Fair	
# 6	-	UCT	1994	22	2	277,281	Variable	
# 7	-	UCT	1994	22	2	279, 283	Variable	
# 8	-	UCT	1994	22	3	278, 282, 285	Very variable	32.9
# 9	-	UCT	1994	22	3	280, 284, 286	Fair	11.9
# 10	-	Wits	2004	5	4	81,83,85,87	Quite variable	15.2
# 11	-	Wits	2004	5	3	84,86,88	Quite variable	16.7
# 12	-	UCT	2007	20	2	269, 272	Good	
# 13	-	UCT	2007	20	2	267, 270	Good	
# 14	both with	UCT	2007	20	2	259, 262	Good	
# 15	-	UCT	2007	20	2	268, 274	Fair	

Table B.2 Comparable experiments between data sets

Comparison #	Matching criteria											
	Cement type	**StoneType**	**SandType**	**Water / binder ratio**	**Aggregate / binder ratio**	**Cement content (kg/m^3)**	**Water content (kg/m^3)**	**Sand content (kg/m^3)**	**Stone content (kg/m^3)**	**Stone / sand ratio**	**Paste volume fraction**	**Supplementary cementitious material fraction**
Comparisons between data sets / test years												
# 16	CEM I	Wits Quartzite	Wits Quartzite	0.42-0.43	3.4-3.9	465-504	198-210	644-664	1036-1147	1.6-1.8	0.346-0.371	0
# 17	CEM I	Andesite	Andesite	0.40-0.42	3.8-3.9	480-488	195-200	732 and 872	1005 and 1135	1.2-1.6	0.350-0.353	0
# 18	CEM I	Andesite	Andesite	0.56	5.5-5.7	348-360	195-200	860 and 983	1005 and 1135	1.0-1.3	0.306-0.315	0
# 19	CEM I	Granite	Granite	0.56-0.60	5.3-6.2	309-348	184-195	826-880	965-1076	1.1-1.3	0.282-0.306	0
# 20	CEM I	Greywacke	Cape Flats sand	0.60	5.7	333	200	787	1100	1.4	0.306	0
# 21	CEM I	Sandstone	Cape Flats sand	0.60	5.7	333	200	787	1100	1.4	0.306	0
# 22	CEM I	Dolerite	Dolomite	0.42	3.3-3.4	524-536	220-225	665-690	1106	1.6-1.7	0.387-0.396	0
# 23	CEM I	Dolerite	Dolomite	0.63	6.4-6.7	310-317	195-200	934-956	1106	1.2	0.294-0.301	0
# 24	CEM II A-L	Dolerite	Dolomite	0.42	3.7-3.8	425-435	205-210	719-745	1106	1.5	0.340-0.349	0.13
# 25	CEM II A-L	Dolerite	Dolomite	0.63-0.78	6.9-8.2	218-263	190-195	943-965	1106	1.1-1.2	0.264-0.274	0.13
# 26	CEM II A-V	Dolerite	Dolomite	0.42	3.5-3.8	425-445	205-215	694-742	1106	1.5-1.6	0.368-0.386	0.13
# 27	CEM II A-V	Dolerite	Dolomite	0.63	6.6-6.9	263-270	190-195	943-963	1106	1.1-1.2	0.291-0.298	0.13
# 28	CEM II B-V	Dolerite	Dolomite	0.42	3.7-3.9	333-351	200-205	721-772	1106	1.4-1.5	0.367-0.377	0.28-0.30
# 29	CEM II B-V	Dolerite	Dolomite	0.63	7.1-7.3	200-212	180-185	972-995	1106	1.1	0.281-0.289	0.28-0.30
# 30	CEM III A	Dolerite	Dolomite	0.42	3.5-3.7	250-259	210-215	671-720	1106	1.5-1.6	0.376-0.385	0.50-0.51
# 31	CEM III A	Dolerite	Dolomite	0.63	3.5-3.7	153-159	195-200	921-943	1106	1.2	0.298-0.305	0.50-0.51
# 32	CEM I	Greywacke	Pit sand Klipheuwel	0.60	6.4-7.0	300	180	820-889	1100-1200	1.3	0.276	0
# 33	CEM III A	Greywacke	Pit sand Klipheuwel	0.60	6.4-6.5	150	180	810-855	1100	1.3-1.4	0.228	0.5
# 34	CEM I	Greywacke	Pit sand Klipheuwel	0.40	4.1-4.4	425-450	170-180	676 and 740	1100-1200	1.5-1.8	0.305-0.323	0
# 35	CEM I	Greywacke	Pit sand Klipheuwel	0.50	5.3	360	180	804-815	1100	1.3-1.4	0.295	0

Table B.2 Comparable experiments between data sets (continued)

Comparison #	Admixture	Laboratory	Test year	Data set(s)	Number of experiments	Line # in data file	Comparison of shrinkage profiles	Relative standard deviation of final shrinkage value (%)
Comparisons within data sets								
# 16	-	Wits	1983, 1993	7, 9	2	110, 136	Good	
# 17	-	Wits	1983, 1998	7,17	2	107, 223	Good	
# 18	-	Wits	1983, 1998	7,17	2	106, 222	Quite variable	
# 19	-	Wits	1990, 1998	12,17	2	159, 220	Good	
# 20	1 with, 1 without	UCT	1999, 2000	11,18	2	146, 224	Good	
# 21	2 with, 2 without	UCT	1999, 2000	11,18	4	147,148, 225, 226	Three good, one different	24.9
# 22	-	Wits	2000, 2001	2,3	4	13,17,33,35	Good	15.3
# 23	-	Wits	2000, 2001	2,3	4	14,34,36,56	Good	8.2
# 24	-	Wits	2000, 2001	2,3	2	15, 39	Variable	
# 25	-	Wits	2000, 2001	2,3	2	16, 40	Good	
# 26	-	Wits	2000, 2001	2,3	2	19, 37	Variable	
# 27	-	Wits	2000, 2001	2,3	2	20, 38	Quite variable	
# 28	-	Wits	2000, 2001	2,3	4	23, 27, 29, 45	Good	4.0
# 29	-	Wits	2000, 2001	2,3	5	22, 24, 28, 30, 46	Good	14.3
# 30	-	Wits	2000, 2001	2,3	5	25, 31, 41, 43, 47	Quite variable	19.7
# 31	-	Wits	2000, 2001	2,3	5	26, 32, 42, 44, 48	Fair	16.5
# 32	-	UCT	2001, 2004, 2007	8, 20, 21	3	117, 265, 276	Two good, one different	
# 33	-	UCT	2001, 2004	8, 21	2	118, 275	Good	
# 34	both with, but different	UCT	2001, 2007	8,20	2	111, 260	Good	
# 35	1 with, 1 without	UCT	2001, 2007	8,20	2	114, 264	Good	

Appendix C

Comparison of OLS and WLS for the fitting of growth curve Model 1 to the shrinkage-time profile for the experiment: # 0173

This Appendix contains the following sections:

- Input File
- SAS Program Listing
- Output: OLS
- Output: WLS

Notes regarding the output:

- The output entitled 'The Means Procedure' was required to obtain the corrected sum of squares for the shrinkage, was then used to calculate the AIC and BIC statistics.
- The next two sections, 'Iterative Phase' and 'Estimation Summary' relate to the execution of the Gauss-Newton Algorithm to estimate the parameters of the growth curve, and show that the criteria for convergence were met. No intercept was specified for this model since the growth curve model used does not have an intercept. The 'missing' data refers to additional values placed in the input file for model prediction.
- The next table shows that the model fit is highly significant. This is followed by the parameter estimates and their approximate 95% confidence intervals, as well as their correlation coefficient.
- The final table gives the predicted values of shrinkage from the fitted model, the raw and studentised residuals, as well as the lagged residuals, which are used in the diagnostic plots which follow.

Input file "Mix0173.xls"

Time	Shrinkage	WeightOLS	WeightWLS
0	0	0.026316	0
3	0	0.026316	0
7	-11	0.026316	0
10	74	0.026316	0.115860
14	142	0.026316	0.059596
15	144	0.026316	0.058752
16	147	0.026316	0.057529
17	155	0.026316	0.054502
20	191	0.026316	0.044045
22	212	0.026316	0.039599
23	217	0.026316	0.038669
24	228	0.026316	0.036767
27	254	0.026316	0.032932
28	273	0.026316	0.030596
34	316	0.026316	0.026355
41	359	0.026316	0.023139
48	384	0.026316	0.021604
55	402	0.026316	0.020618
62	409	0.026316	0.020258
69	429	0.026316	0.019295
81	437	0.026316	0.018935
86	440	0.026316	0.018803
93	454	0.026316	0.018212
100	462	0.026316	0.017890
107	469	0.026316	0.017618
114	489	0.026316	0.016883
121	494	0.026316	0.016709
127	493	0.026316	0.016743
148	499	0.026316	0.016538
170	493	0.026316	0.016743
198	503	0.026316	0.016404
235	519	0.026316	0.015888
260	507	0.026316	0.016272
332	519	0.026316	0.015888
492	542	0.026316	0.015201
580	539	0.026316	0.015287
715	546	0.026316	0.015087
833	557	0.026316	0.014783

Program Listing

```
PROC IMPORT OUT= CONCRETE.Mix0173         /* Import data file */
      DATAFILE= "C:\Documents and Settings\Petra Gaylard\My
        Documents\Wits Masters\Research Report\Analyses\9-10
        Growth curves\Key shrinkage profiles raw data\
        Mix0173.xls"
     DBMS=EXCEL REPLACE;
     SHEET="Sheet1$";
     GETNAMES=YES;
RUN;

goptions reset=all;           /* Set up graphics options */
ods html;
ods graphics on;
symbol1 c=blue value=circle;
symbol2 c=red value=none interpol=join;
title 'Model 1 / Data: Mix 0173 / WLS';

data work;
   set concrete.Mix0173;
run;

proc means data=work css;
            /* Used to obtain corrected SS for Shrinkage */
run;

data predictions;
            /* Set up data set for predicted growth curve at fixed
            intervals of time */
      do time = 0 to 900 by 25;
            predict=1;
            output;
      end;
run;

data fitthis;
            /* concatenate raw data and prediction data sets */
      set work predictions;
run;

proc nlin data=fitthis maxiter=500;
            /* fit growth curve model to raw data */
      parameters a = 1 b = .99;  /* parameter starting values */
      bounds a>0;                /* parameter bounds */
      bounds b>0;
      bounds b<1;
      model Shrinkage = a*(1-(b**time));  /* fit model */
      _weight_= WeightWLS;    /* Statement amended for OLS fit */
output out=model_results predicted=pred residual=resid
            student=stu_resid weight=weight;
run;

data results;
set model_results;
stu_resid1=lag(stu_resid);
                  /* set up lagged residuals for later plots */
stu_resid2=lag(stu_resid1);
stu_resid3=lag(stu_resid2);
```

```
proc print data=results;
            /* print raw data and prediction data sets with
            predicted values, weights and residuals */
run;

proc sort data=results;
            /* sort output table by time in order to plot
            predicted growth curve */
by Time;
run;

proc gplot data=results;
plot (shrinkage pred)*time / overlay;
                  /* plot predicted growth curve and raw data vs.
                  time */
plot stu_resid*pred / vref=0;
                  /* plot of studentised residuals vs. predicted
                  values */
plot stu_resid * stu_resid1;  /* plots of lagged residuals */
plot stu_resid * stu_resid2;
plot stu_resid * stu_resid3;
plot (shrinkage pred) * pred / overlay;
                  /* plot of measured vs. predicted shrinkage with
                  unity line */
plot (stu_resid stu_resid) * time / overlay vref=0;
                  /* plot of studentised residuals vs. time */
run;

proc univariate data=results noprint;
      var stu_resid;
      probplot / normal (mu=0 sigma=1);
                  /* normal probability plot of errors */
quit;

ods graphics off;
ods html close;
```

Output: OLS

Model 1 / Data: Mix0173 / OLS

The MEANS Procedure

Variable	Label	Corrected SS
Time	Time	1469091.71
Shrinkage	Shrinkage	1138664.55
WeightOLS	WeightOLS	0
WeightWLS	WeightWLS	0.0166937

The NLIN Procedure
Dependent Variable Shrinkage
Method: Gauss-Newton

Iterative Phase

Iter	a	b	Sum of Squares
0	1.0000	0.9900	5767425
1	63.1851	0.4188	4252619
2	72.6414	0.7622	4045040
3	84.4227	0.8839	3808626
4	115.6	0.9790	3571135
5	248.0	0.9760	1605189
6	481.5	0.9776	66044.7
7	525.5	0.9768	19388.7
8	525.6	0.9769	19387.0
9	525.6	0.9769	19387.0

NOTE: Convergence criterion met.

Estimation Summary

Method	Gauss-Newton
Iterations	9
Subiterations	10
Average Subiterations	1.111111
R	6.271E-7
PPC(a)	3.356E-8
RPC(a)	4.245E-6
Object	4.644E-9
Objective	19387.05
Observations Read	75
Observations Used	38
Observations Missing	37

NOTE: An intercept was not specified for this model.

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Model	2	151715	75857.7	5352.71	<.0001
Error	36	510.2	14.1718		
Uncorrected Total	38	152226			

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits	
a	525.6	6.9047	511.5	539.6
b	0.9769	0.000951	0.9750	0.9788

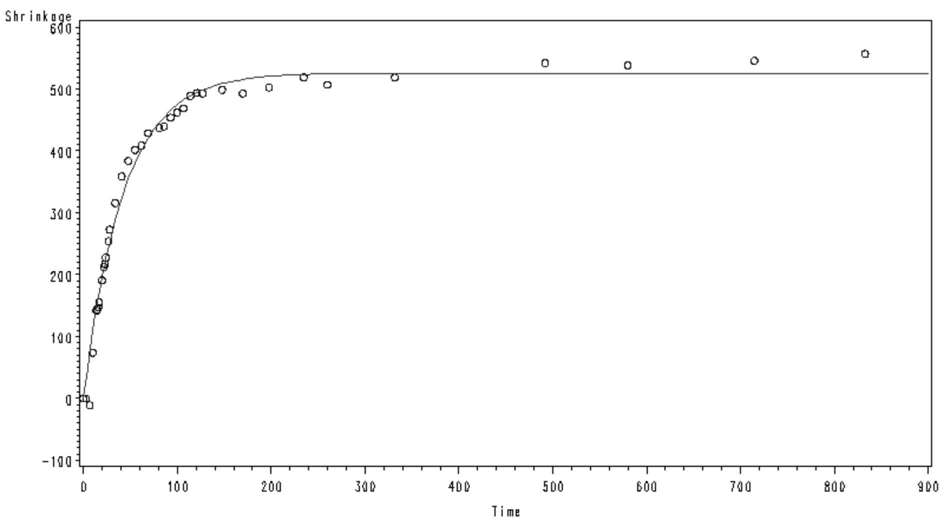
Approximate Correlation Matrix

	a	b
a	1.0000000	0.6773630
b	0.6773630	1.0000000

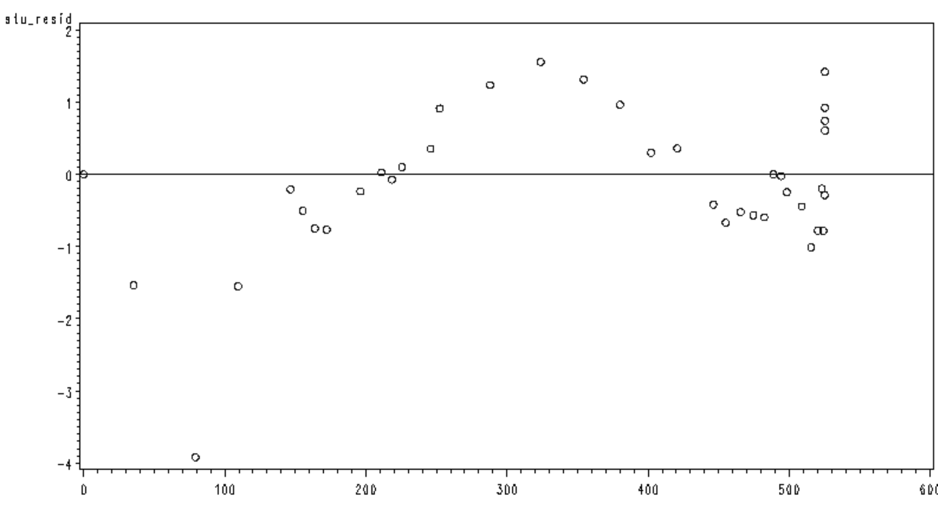
Obs	Time	Shrin- kage	Weight OLS	Weight WLS	pre- dict	pred	resid	stu_ resid	weight	stu_ resid1	stu_ resid2	stu_ resid3
1	0	0	0.026316	0.00000	.	0.000	0.0000	0.00000	0.026316	.	.	.
2	3	0	0.026316	0.00000	.	35.605	-35.6053	-1.53624	0.026316	0.00000	.	.
3	7	-11	0.026316	0.00000	.	79.355	-90.3553	-3.91560	0.026316	-1.53624	0.00000	.
4	10	74	0.026316	0.11586	.	109.584	-35.5844	-1.54868	0.026316	-3.91560	-1.53624	0.00000
5	14	142	0.026316	0.05960	.	146.728	-4.7284	-0.20703	0.026316	-1.54868	-3.91560	-1.53624
6	15	144	0.026316	0.05875	.	155.484	-11.4841	-0.50356	0.026316	-0.20703	-1.54868	-3.91560
7	16	147	0.026316	0.05753	.	164.037	-17.0375	-0.74814	0.026316	-0.50356	-0.20703	-1.54868
8	17	155	0.026316	0.05450	.	172.393	-17.3931	-0.76482	0.026316	-0.74814	-0.50356	-0.20703
9	20	191	0.026316	0.04404	.	196.319	-5.3191	-0.23480	0.026316	-0.76482	-0.74814	-0.50356
10	22	212	0.026316	0.03960	.	211.362	0.6378	0.02822	0.026316	-0.23480	-0.76482	-0.74814
11	23	217	0.026316	0.03867	.	218.624	-1.6241	-0.07193	0.026316	0.02822	-0.23480	-0.76482
12	24	228	0.026316	0.03677	.	225.718	2.2819	0.10116	0.026316	-0.07193	0.02822	-0.23480
13	27	254	0.026316	0.03293	.	246.031	7.9687	0.35416	0.026316	0.10116	-0.07193	0.02822
14	28	273	0.026316	0.03060	.	252.492	20.5081	0.91209	0.026316	0.35416	0.10116	-0.07193
15	34	316	0.026316	0.02636	.	288.237	27.7626	1.23815	0.026316	0.91209	0.35416	0.10116
16	41	359	0.026316	0.02314	.	324.071	34.9295	1.55882	0.026316	1.23815	0.91209	0.35416
17	48	384	0.026316	0.02160	.	354.493	29.5069	1.31516	0.026316	1.55882	1.23815	0.91209
18	55	402	0.026316	0.02062	.	380.322	21.6780	0.96397	0.026316	1.31516	1.55882	1.23815
19	62	409	0.026316	0.02026	.	402.251	6.7491	0.29930	0.026316	0.96397	1.31516	1.55882
20	69	429	0.026316	0.01929	.	420.869	8.1313	0.35964	0.026316	0.29930	0.96397	1.31516
21	81	437	0.026316	0.01893	.	446.482	-9.4824	-0.41793	0.026316	0.35964	0.29930	0.96397
22	86	440	0.026316	0.01880	.	455.207	-15.2073	-0.66960	0.026316	-0.41793	0.35964	0.29930
23	93	454	0.026316	0.01821	.	465.829	-11.8290	-0.52041	0.026316	-0.66960	-0.41793	0.35964
24	100	462	0.026316	0.01789	.	474.847	-12.8469	-0.56503	0.026316	-0.52041	-0.66960	-0.41793
25	107	469	0.026316	0.01762	.	482.503	-13.5031	-0.59401	0.026316	-0.56503	-0.52041	-0.66960
26	114	489	0.026316	0.01688	.	489.003	-0.0033	-0.00014	0.026316	-0.59401	-0.56503	-0.52041
27	121	494	0.026316	0.01671	.	494.522	-0.5220	-0.02300	0.026316	-0.00014	-0.59401	-0.56503
28	127	493	0.026316	0.01674	.	498.584	-5.5841	-0.24628	0.026316	-0.02300	-0.00014	-0.59401
29	148	499	0.026316	0.01654	.	509.048	-10.0485	-0.44509	0.026316	-0.24628	-0.02300	-0.00014
30	170	493	0.026316	0.01674	.	515.686	-22.6859	-1.00956	0.026316	-0.44509	-0.24628	-0.02300
31	198	503	0.026316	0.01640	.	520.426	-17.4260	-0.77935	0.026316	-1.00956	-0.44509	-0.24628
32	235	519	0.026316	0.01589	.	523.394	-4.3944	-0.19739	0.026316	-0.77935	-1.00956	-0.44509
33	260	507	0.026316	0.01627	.	524.350	-17.3496	-0.78074	0.026316	-0.19739	-0.77935	-1.00956
34	332	519	0.026316	0.01589	.	525.329	-6.3290	-0.28546	0.026316	-0.78074	-0.19739	-0.77935
35	492	542	0.026316	0.01520	.	525.547	16.4530	0.74260	0.026316	-0.28546	-0.78074	-0.19739
36	580	539	0.026316	0.01529	.	525.552	13.4483	0.60700	0.026316	0.74260	-0.28546	-0.78074
37	715	546	0.026316	0.01509	.	525.552	20.4477	0.92293	0.026316	0.60700	0.74260	-0.28546
38	833	557	0.026316	0.01478	.	525.552	31.4477	1.41942	0.026316	0.92293	0.60700	0.74260
39	0	.	.	.	1	0.000	.	.	.	1.41942	0.92293	0.60700
40	25	.	.	.	1	232.648	.	.	.	.	1.41942	0.92293
41	50	.	.	.	1	362.309	.	.	.	.	.	1.41942
42	75	.	.	.	1	434.573	.	.	.	.	.	.
43	100	.	.	.	1	474.847	.	.	.	.	.	.
44	125	.	.	.	1	497.293	.	.	.	.	.	.
45	150	.	.	.	1	509.803	.	.	.	.	.	.
46	175	.	.	.	1	516.775	.	.	.	.	.	.
47	200	.	.	.	1	520.660	.	.	.	.	.	.
48	225	.	.	.	1	522.826	.	.	.	.	.	.
49	250	.	.	.	1	524.033	.	.	.	.	.	.
50	275	.	.	.	1	524.705	.	.	.	.	.	.
51	300	.	.	.	1	525.080	.	.	.	.	.	.
52	325	.	.	.	1	525.289	.	.	.	.	.	.
53	350	.	.	.	1	525.406	.	.	.	.	.	.
54	375	.	.	.	1	525.471	.	.	.	.	.	.
55	400	.	.	.	1	525.507	.	.	.	.	.	.
56	425	.	.	.	1	525.527	.	.	.	.	.	.
57	450	.	.	.	1	525.538	.	.	.	.	.	.
58	475	.	.	.	1	525.544	.	.	.	.	.	.

59	500	.	.	.	1	525.548	.	.	.	.	.	.
60	525	.	.	.	1	525.550	.	.	.	.	.	.
61	550	.	.	.	1	525.551	.	.	.	.	.	.
62	575	.	.	.	1	525.552	.	.	.	.	.	.
63	600	.	.	.	1	525.552	.	.	.	.	.	.
64	625	.	.	.	1	525.552	.	.	.	.	.	.
65	650	.	.	.	1	525.552	.	.	.	.	.	.
66	675	.	.	.	1	525.552	.	.	.	.	.	.
67	700	.	.	.	1	525.552	.	.	.	.	.	.
68	725	.	.	.	1	525.552	.	.	.	.	.	.
69	750	.	.	.	1	525.552	.	.	.	.	.	.
70	775	.	.	.	1	525.552	.	.	.	.	.	.
71	800	.	.	.	1	525.552	.	.	.	.	.	.
72	825	.	.	.	1	525.552	.	.	.	.	.	.
73	850	.	.	.	1	525.552	.	.	.	.	.	.
74	875	.	.	.	1	525.552	.	.	.	.	.	.
75	900	.	.	.	1	525.552	.	.	.	.	.	.

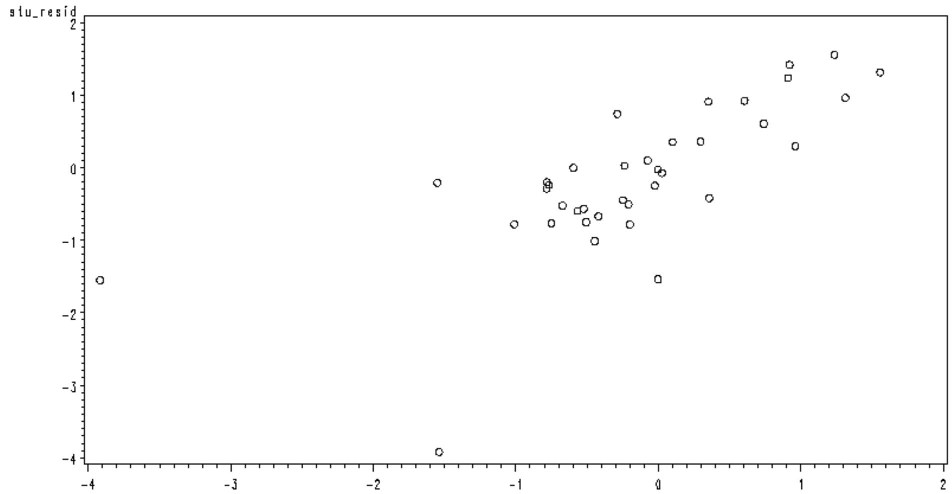
Model 1 / Data: Mix 0173 / OLS



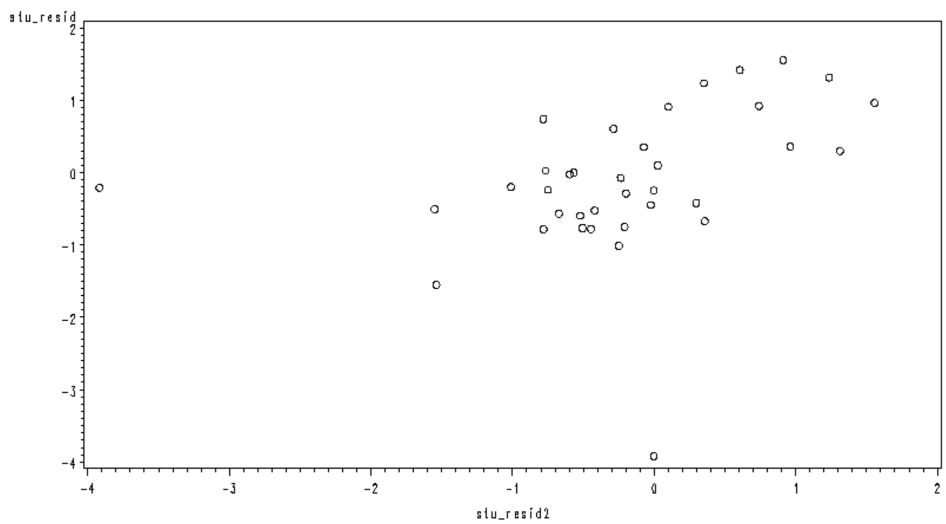
Model 1 / Data: Mix 0173 / OLS



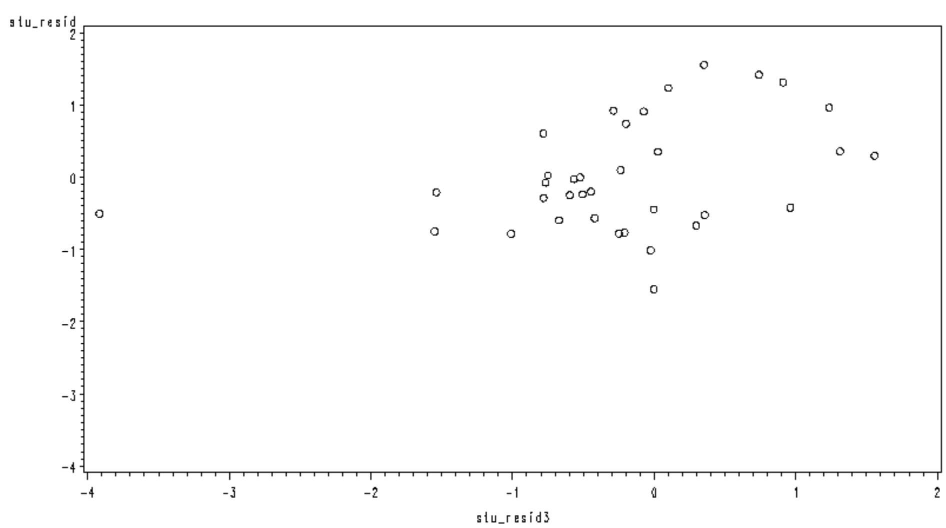
Model 1 / Data: Mix 0173 / OLS



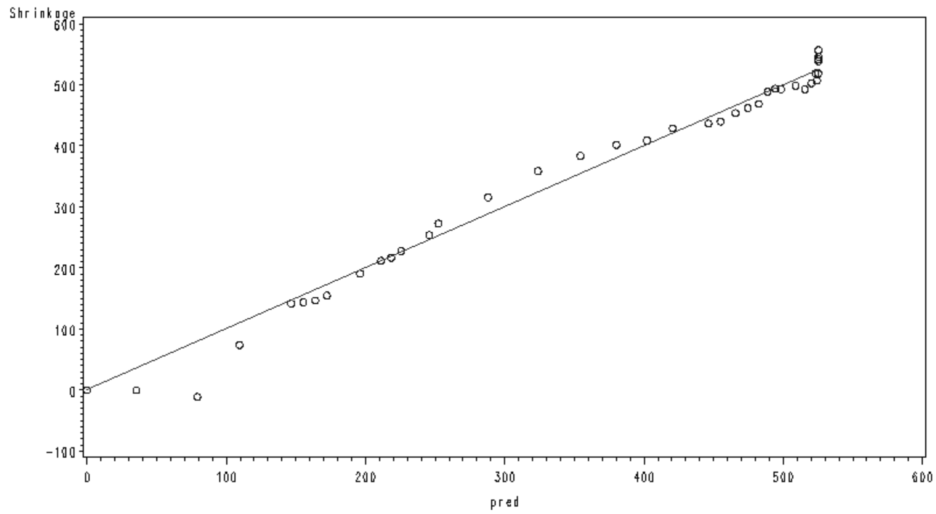
Model 1 / Data: Mix 0173 / OLS



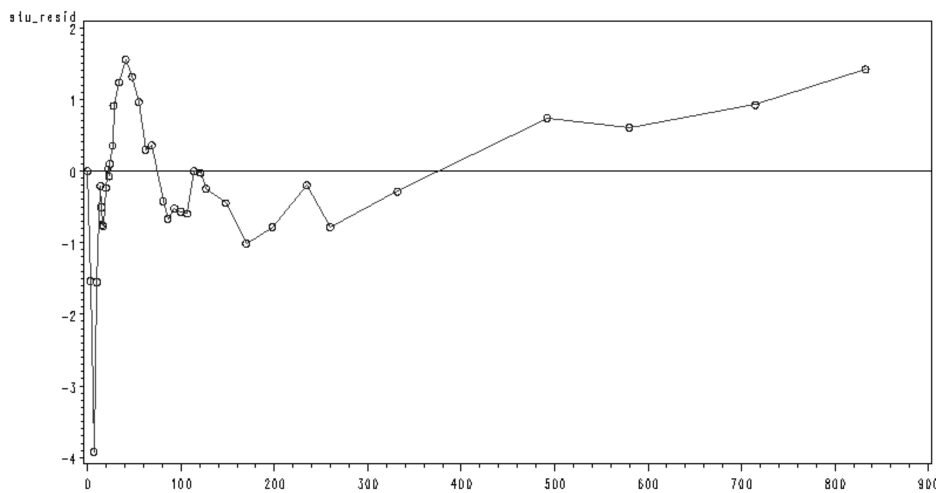
Model 1 / Data: Mix 0173 / OLS



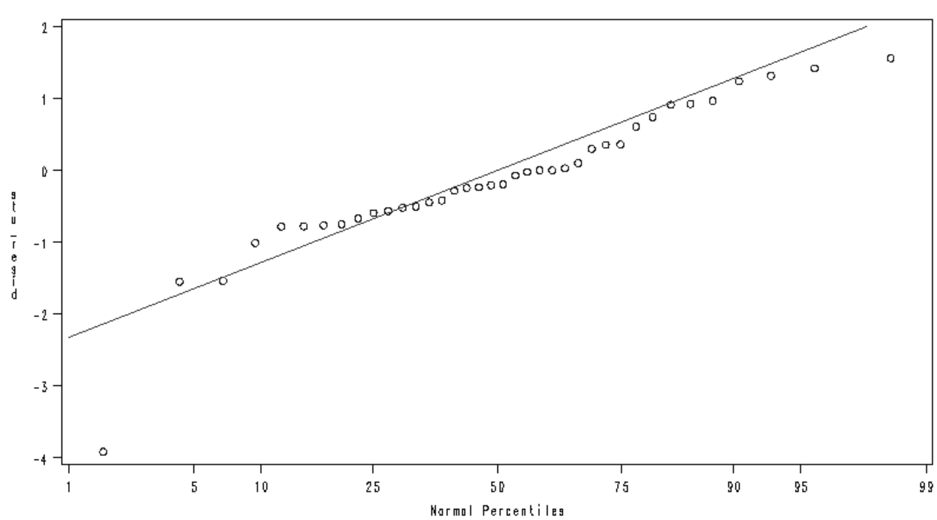
Model 1 / Data: Mix 0173 / OLS



Model 1 / Data: Mix 0173 / OLS



Model 1 / Data: Mix 0173 / OLS



Output: WLS

Model 1 / Data: Mix0173 / WLS

The MEANS Procedure

Variable	Label	Corrected SS
Time	Time	1469091.71
Shrinkage	Shrinkage	1138664.55
WeightOLS	WeightOLS	0
WeightWLS	WeightWLS	0.0166937

The NLIN Procedure
Dependent Variable Shrinkage
Method: Gauss-Newton

Iterative Phase

Iter	a	b	Sum of Squares
0	1.0000	0.9900	109851
1	61.7470	0.3584	78031.5
2	61.7628	0.6120	78027.0
3	61.8460	0.6240	77989.7
4	62.2637	0.6751	77804.9
5	64.4078	0.8079	76968.0
6	76.6531	0.9482	75033.7
7	147.0	0.9898	69234.9
8	324.3	0.9721	13810.2
9	518.2	0.9809	1073.6
10	525.1	0.9774	346.0
11	527.3	0.9774	343.4
12	527.3	0.9774	343.4

NOTE: Convergence criterion met.

Estimation Summary

Method	Gauss-Newton
Iterations	12
Subiterations	16
Average Subiterations	1.333333
R	3.365E-6
PPC(a)	1.952E-7
RPC(a)	0.000011
Object	4.446E-8
Objective	343.4071
Observations Read	75
Observations Used	35
Observations Missing	40

NOTE: An intercept was not specified for this model.

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Model	2	109817	54908.7	5276.50	<.0001
Error	33	343.4	10.4063		
Uncorrected Total	35	110161			

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits	
a	527.3	7.4041	512.3	542.4
b	0.9774	0.000778	0.9758	0.9790

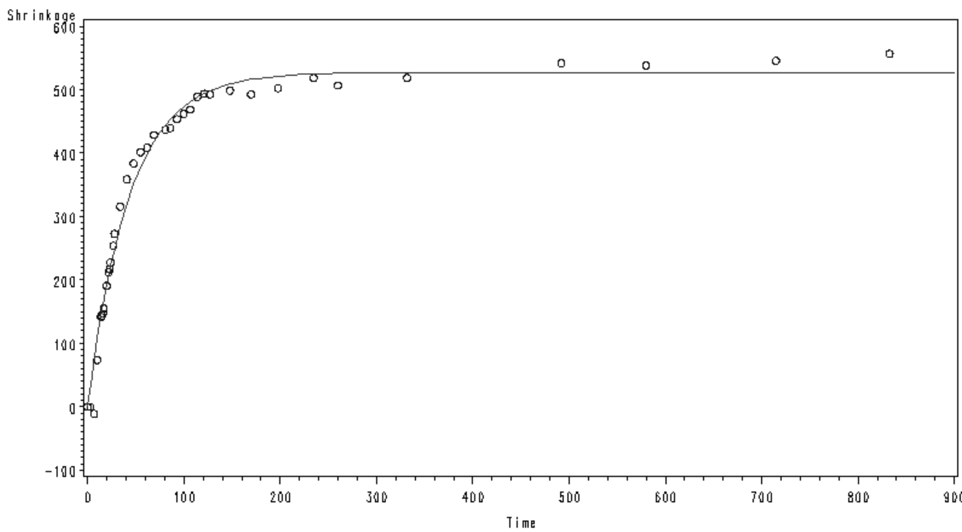
Approximate Correlation Matrix

	a	b
a	1.0000000	0.7206600
b	0.7206600	1.0000000

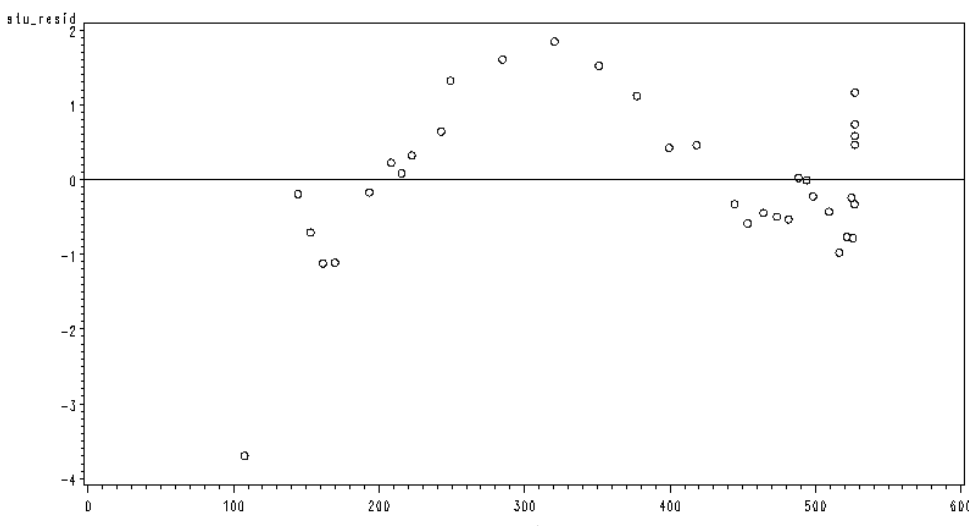
Obs	Time	Shrinkage	Weight OLS	Weight WLS	predict	pred	resid	stu_resid	weight	stu_resid1	stu_resid2	stu_resid3
1	0	0	0.026316	0.00000	.	0.000	0.0000	.	0.00000	.	.	.
2	3	0	0.026316	0.00000	.	34.957	-34.9575	.	0.00000	.	.	.
3	7	-11	0.026316	0.00000	.	77.989	-88.9894	.	0.00000	.	.	.
4	10	74	0.026316	0.11586	.	107.777	-33.7767	-3.69273	0.11586	.	.	.
5	14	142	0.026316	0.05960	.	144.444	-2.4443	-0.19042	0.05960	-3.69273	.	.
6	15	144	0.026316	0.05875	.	153.099	-9.0991	-0.70540	0.05875	-0.19042	-3.69273	.
7	16	147	0.026316	0.05753	.	161.558	-14.5582	-1.11893	0.05753	-0.70540	-0.19042	-3.69273
8	17	155	0.026316	0.05450	.	169.826	-14.8261	-1.10976	0.05450	-1.11893	-0.70540	-0.19042
9	20	191	0.026316	0.04404	.	193.525	-2.5254	-0.16969	0.04404	-1.10976	-1.11893	-0.70540
10	22	212	0.026316	0.03960	.	208.445	3.5546	0.22635	0.03960	-0.16969	-1.10976	-1.11893
11	23	217	0.026316	0.03867	.	215.653	1.3465	0.08477	0.03867	0.22635	-0.16969	-1.10976
12	24	228	0.026316	0.03677	.	222.699	5.3014	0.32527	0.03677	0.08477	0.22635	-0.16969
13	27	254	0.026316	0.03293	.	242.893	11.1073	0.64445	0.03293	0.32527	0.08477	0.22635
14	28	273	0.026316	0.03060	.	249.322	23.6779	1.32208	0.03060	0.64445	0.32527	0.08477
15	34	316	0.026316	0.02636	.	284.959	31.0412	1.60584	0.02636	1.32208	0.64445	0.32527
16	41	359	0.026316	0.02314	.	320.803	38.1966	1.84678	0.02314	1.60584	1.32208	0.64445
17	48	384	0.026316	0.02160	.	351.347	32.6534	1.52253	0.02160	1.84678	1.60584	1.32208
18	55	402	0.026316	0.02062	.	377.373	24.6274	1.11977	0.02062	1.52253	1.84678	1.60584
19	62	409	0.026316	0.02026	.	399.549	9.4506	0.42544	0.02026	1.11977	1.52253	1.84678
20	69	429	0.026316	0.01929	.	418.446	10.5537	0.46295	0.01929	0.42544	1.11977	1.52253
21	81	437	0.026316	0.01893	.	444.570	-7.5702	-0.32878	0.01893	0.46295	0.42544	1.11977
22	86	440	0.026316	0.01880	.	453.509	-13.5094	-0.58476	0.01880	-0.32878	0.46295	0.42544
23	93	454	0.026316	0.01821	.	464.426	-10.4257	-0.44417	0.01821	-0.58476	-0.32878	0.46295
24	100	462	0.026316	0.01789	.	473.728	-11.7275	-0.49552	0.01789	-0.44417	-0.58476	-0.32878
25	107	469	0.026316	0.01762	.	481.654	-12.6536	-0.53107	0.01762	-0.49552	-0.44417	-0.58476
26	114	489	0.026316	0.01688	.	488.407	0.5925	0.02436	0.01688	-0.53107	-0.49552	-0.44417
27	121	494	0.026316	0.01671	.	494.162	-0.1624	-0.00665	0.01671	0.02436	-0.53107	-0.49552
28	127	493	0.026316	0.01674	.	498.413	-5.4128	-0.22219	0.01674	-0.00665	0.02436	-0.53107
29	148	499	0.026316	0.01654	.	509.435	-10.4350	-0.42767	0.01654	-0.22219	-0.00665	0.02436
30	170	493	0.026316	0.01674	.	516.505	-23.5045	-0.97423	0.01674	-0.42767	-0.22219	-0.00665
31	198	503	0.026316	0.01640	.	521.618	-18.6178	-0.76687	0.01640	-0.97423	-0.42767	-0.22219
32	235	519	0.026316	0.01589	.	524.873	-5.8725	-0.23869	0.01589	-0.76687	-0.97423	-0.42767
33	260	507	0.026316	0.01627	.	525.938	-18.9377	-0.78101	0.01627	-0.23869	-0.76687	-0.97423
34	332	519	0.026316	0.01589	.	527.053	-8.0529	-0.32850	0.01589	-0.78101	-0.23869	-0.76687
35	492	542	0.026316	0.01520	.	527.312	14.6877	0.58527	0.01520	-0.32850	-0.78101	-0.23869
36	580	539	0.026316	0.01529	.	527.318	11.6817	0.46693	0.01529	0.58527	-0.32850	-0.78101
37	715	546	0.026316	0.01509	.	527.319	18.6808	0.74138	0.01509	0.46693	0.58527	-0.32850
38	833	557	0.026316	0.01478	.	527.319	29.6808	1.16499	0.01478	0.74138	0.46693	0.58527
39	0	.	.	.	1	0.000	.	.	.	1.16499	0.74138	0.46693
40	25	.	.	.	1	229.584	.	.	.	.	1.16499	0.74138
41	50	.	.	.	1	359.212	.	.	.	.	.	1.1649959
42	75	.	.	.	1	432.403	.	.	.	.	.	.
43	100	.	.	.	1	473.728	.	.	.	.	.	.
44	125	.	.	.	1	497.060	.	.	.	.	.	.
45	150	.	.	.	1	510.234	.	.	.	.	.	.
46	175	.	.	.	1	517.673	.	.	.	.	.	.
47	200	.	.	.	1	521.873	.	.	.	.	.	.
48	225	.	.	.	1	524.244	.	.	.	.	.	.
49	250	.	.	.	1	525.583	.	.	.	.	.	.
50	275	.	.	.	1	526.339	.	.	.	.	.	.
51	300	.	.	.	1	526.766	.	.	.	.	.	.
52	325	.	.	.	1	527.007	.	.	.	.	.	.
53	350	.	.	.	1	527.143	.	.	.	.	.	.
54	375	.	.	.	1	527.220	.	.	.	.	.	.
55	400	.	.	.	1	527.263	.	.	.	.	.	.
56	425	.	.	.	1	527.287	.	.	.	.	.	.
57	450	.	.	.	1	527.301	.	.	.	.	.	.
58	475	.	.	.	1	527.309	.	.	.	.	.	.

59	500	.	.	.	1	527.313	.	.	.	.	.	.
60	525	.	.	.	1	527.316	.	.	.	.	.	.
61	550	.	.	.	1	527.317	.	.	.	.	.	.
62	575	.	.	.	1	527.318	.	.	.	.	.	.
63	600	.	.	.	1	527.319	.	.	.	.	.	.
64	625	.	.	.	1	527.319	.	.	.	.	.	.
65	650	.	.	.	1	527.319	.	.	.	.	.	.
66	675	.	.	.	1	527.319	.	.	.	.	.	.
67	700	.	.	.	1	527.319	.	.	.	.	.	.
68	725	.	.	.	1	527.319	.	.	.	.	.	.
69	750	.	.	.	1	527.319	.	.	.	.	.	.
70	775	.	.	.	1	527.319	.	.	.	.	.	.
71	800	.	.	.	1	527.319	.	.	.	.	.	.
72	825	.	.	.	1	527.319	.	.	.	.	.	.
73	850	.	.	.	1	527.319	.	.	.	.	.	.
74	875	.	.	.	1	527.319	.	.	.	.	.	.
75	900	.	.	.	1	527.319	.	.	.	.	.	.

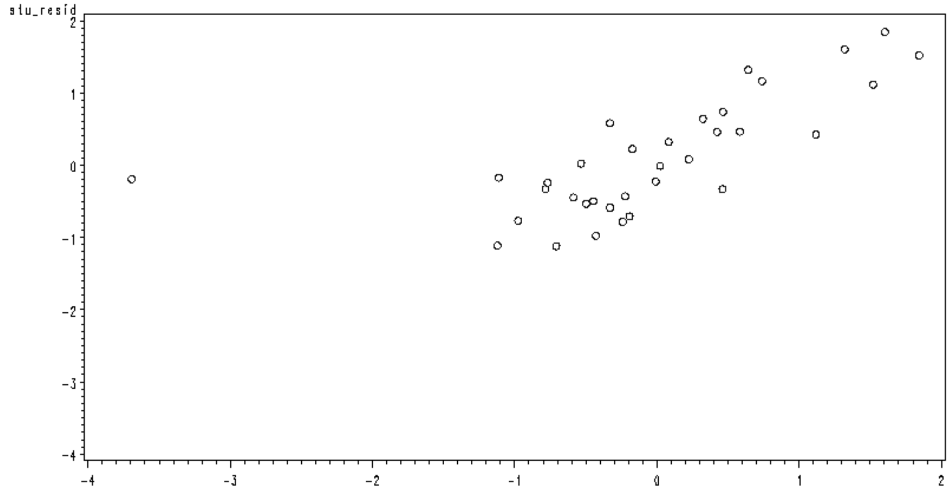
Model 1 / Data: Mix 0173 / WLS



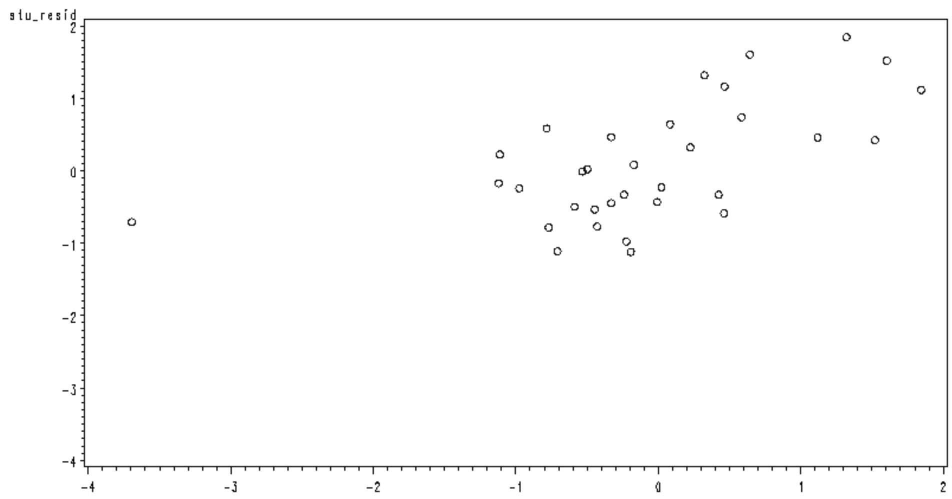
Model 1 / Data: Mix 0173 / WLS



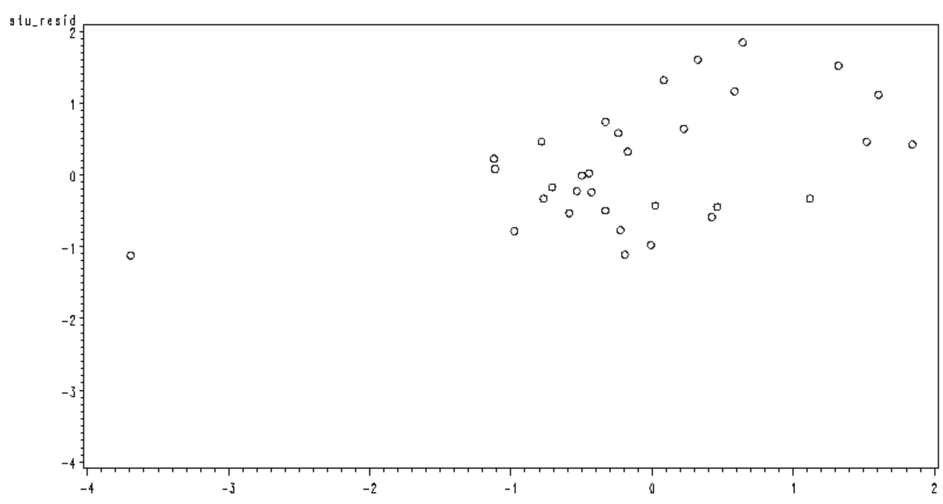
Model 1 / Data: Mix 0173 / WLS



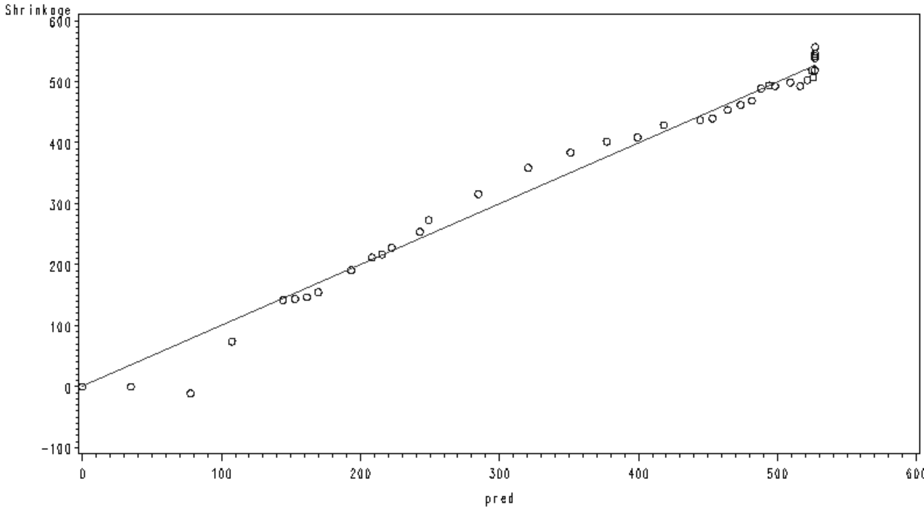
Model 1 / Data: Mix 0173 / WLS



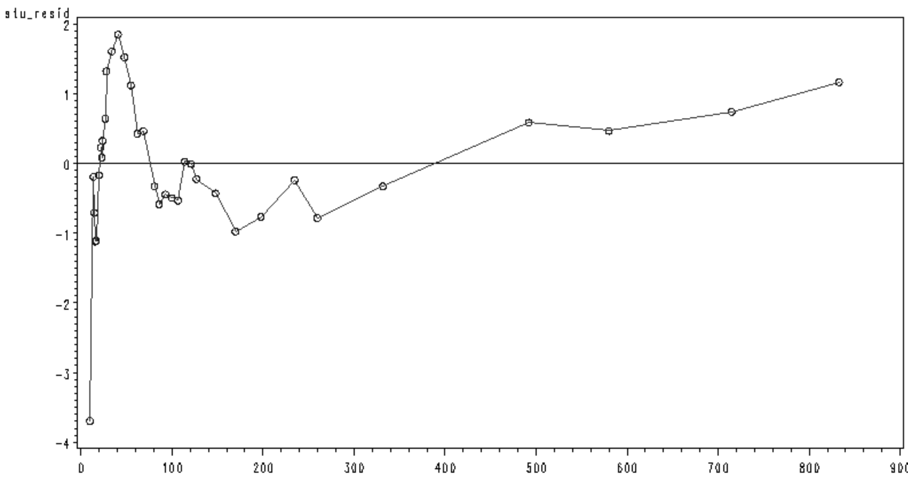
Model 1 / Data: Mix 0173 / WLS



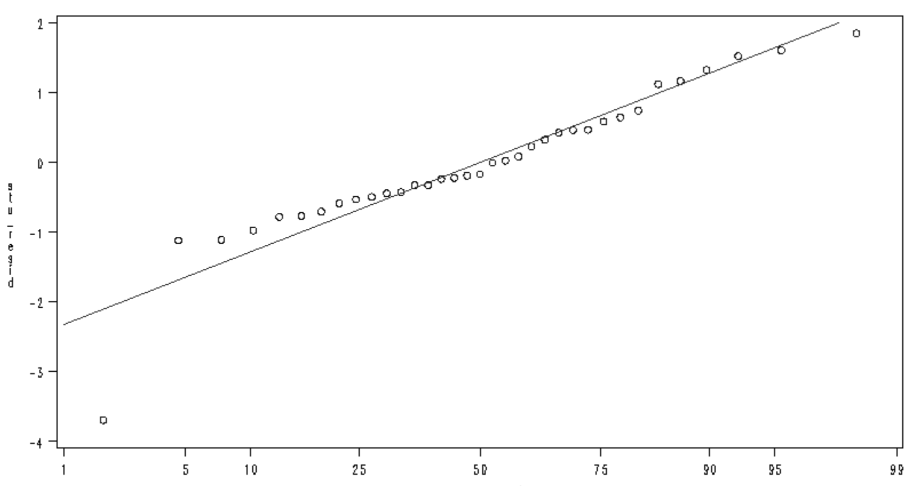
Model 1 / Data: Mix 0173 / WLS



Model 1 / Data: Mix 0173 / WLS



Model 1 / Data: Mix 0173 / WLS



Appendix D

SAS Program listing for fitting of the fifteen growth curve models to shrinkage profiles

```
PROC IMPORT OUT= CONCRETE.Mix0173         /* Import data file */
      DATAFILE= "C:\Documents and Settings\Petra Gaylard\My
      Documents\Wits Masters\Research Report\Analyses\9-10 Growth
      curves\Key shrinkage profiles raw data\Mix0173.xls"
     DBMS=EXCEL REPLACE;
     SHEET="Sheet1$";
     GETNAMES=YES;
RUN;

goptions reset=all;           /* Set up graphics options */
ods html;
ods graphics on;
symbol1 c=blue value=circle;
symbol2 c=red value=none interpol=join;
title 'Model 15 / Data: Mix0173;

data work;
   set concrete.Mix0173;
run;

proc means data=work css;
      /* Used to obtain corrected SS for Shrinkage */
run;

data predictions;
      /* Set up data set for predicted growth curve at fixed
      intervals of time */
      do time = 0 to 900 by 25;
            predict=1;
            output;
      end;
run;

data fitthis;/* concatenate raw data and prediction data sets */
      set work predictions;
run;

proc nlin data=fitthis best=10 hougaard smethod=golden
     maxiter=500;       /* fit growth curve model to raw data */
**parameters a = 1 b = .99;/* select parameter starting values */
**parameters a = 700 b = 25 g = .39;
**10 parameters a=100 to 800 by 50 b=.0010 to .0050 by .001
      g=.1 to 1 by 0.1;
**12 parameters a=300 to 1500 by 50 b=10 to 120 by 5
      g=.1 to 1 by 0.1;
**13,15; parameters a=400 to 500 by 10 b=700 to 800 by 5
      g=.7 to .7 by 0.1;
bounds a>0;                         /* select parameter bounds */
bounds b>0;
**bounds b<1;
bounds g>0;
**bounds g<1;
** 1 model Shrinkage = a*(1-(b**time));         /* select model */
** 2 model Shrinkage = a*(1-exp(-b*time));
** 4 model Shrinkage = (a*time) / (b+time);
** 5 model Shrinkage = (a*b*time) / (1+(b*time));
** 7 model Shrinkage = a*tanh(sqrt(time/b));
** 8 model Shrinkage = a*time**(b*time**(-g));
** 9 model Shrinkage =(a*b*time**(1-g)) / (1+(b*time**(1-g)));
** 10 model Shrinkage = a*(1-exp(-b*time))**g;
```

```
** 11 model Shrinkage = a*((time**b)/(1+(time**b)))**g;
** 12 model Shrinkage = a*(time**g)/(b+(time**g));
** 13 model Shrinkage = a*(time /(b+time))**g;
** 14 model Shrinkage = a*(1-(1/(time+1)**b))**g;
** 15; model Shrinkage = a*(1-(1/(1+(time/b)**g)))**g;
output out=model_results predicted=pred residual=resid
      student=stu_resid;
run;

data results;
set model_results;
stu_resid1=lag(stu_resid);
      /* set up lagged residuals for later plots */
stu_resid2=lag(stu_resid1);
stu_resid3=lag(stu_resid2);

proc print data=results;
      /* print raw data and prediction data sets with predicted
      values, weights and residuals */
run;

proc sort data=results;
      /* sort output table by time in order to plot predicted
      growth curve */
by Time;
run;

proc gplot data=results;
plot (shrinkage pred)*time / overlay;
      /* plot predicted growth curve and raw data vs. time */
plot stu_resid*pred / vref=0;
      /* plot of studentised residuals vs. predicted values */
plot stu_resid * stu_resid1;  /* plots of lagged residuals */
plot stu_resid * stu_resid2;
plot stu_resid * stu_resid3;
plot (shrinkage pred) * pred / overlay;
      /* plot of measured vs. predicted shrinkage with unity line
      */
plot (stu_resid stu_resid) * time / overlay vref=0;
      /* plot of studentised residuals vs. time */
run;

proc univariate data=results noprint;
      var stu_resid;
      probplot / normal (mu=0 sigma=1);
            /* normal probability plot of errors */
quit;

ods graphics off;
ods html close;
```

Appendix E

Results of the fitting of the fifteen growth curve models to one shrinkage profile (# 0173)

In the table which follows, highlighting has been used to identify poor model fits easily. Relative standard errors of parameter estimates greater than 10% are highlighted in pale orange, while those greater than 100% are highlighted in bright orange. Value of skewness of the estimated parameters greater than 0.25 are highlighted in pink. Poor diagnostic findings are given in red font.

Correlations between the parameters exceeding 0.90, which are important to note, are highlighted in yellow.

The results of the fitting of the 15 growth curve models to all ten "key" shrinkage profiles may be found in the spreadsheet "Growth curve models – key shrinkage profiles.xls" on the accompanying CD.

Table E.1 Results of the fitting of the fifteen growth curve models to one shrinkage profile (# 0173)

Model	Number of obser-vations	Number of para-meters	Fitting algorithm	Parameter estimates			Standard error of parameter estimate			Relative standard error of parameter estimate (values > 10% highlighted)			Skewness (values > 0.25 highlighted)		
				α	β	γ	α	β	γ	α	β	γ	α	β	γ
Model 1	38	2	Gauss-Newton	525.6	0.977		6.90	0.000951		1.3%	0.1%		0.036	-0.136	
Model 2	38	2	Gauss-Newton	525.6	0.023		6.90	0.000974		1.3%	4.2%		0.036	0.139	
Model 3	38	2	Gauss-Newton	421.6	16700159		20.00	9631640		4.7%	57.7%		0.001	2.032	
Model 4	38	2	Gauss-Newton	604.8	36.351		13.89	2.88		2.3%	7.9%		0.075	0.202	
Model 5	38	2	Gauss-Newton	604.8	0.028		13.89	0.00218		2.3%	7.9%		0.075	0.273	
Model 6	38	2	Gauss-Newton	0.0074	-0.151		0.00	0.00802		8.7%	-5.3%		0.267	-0.570	
Model 7	35	2	Gauss-Newton	579.0	108.900		23.30	16.81		4.0%	15.4%		0.000	0.488	
Model 8	37	3	Gauss-Newton	1.2	2.839	0.171	0.40	0.19	0.00265	32.7%	6.6%	1.6%	0.900	0.076	-0.208
Model 9	37	3	Gauss-Newton	590.0	0.030	0.0001	24.23	0.00869	0.100	4.1%	29.0%	100100%	0.443	0.596	-0.219
Model 10	37	3	Gauss-Newton	515.0	0.031	1.304	6.49	0.00286	0.122	1.3%	9.3%	9.4%	0.063	0.269	0.410
Model 11	37	3	Gauss-Newton	553.2	1.122	30.314	5.28	0.0349	3.06	1.0%	3.1%	10.1%	0.093	0.099	0.412
Model 12	37	3	Gauss-Newton	534.2	240.700	1.634	5.78	49.21	0.0654	1.1%	20.4%	4.0%	0.090	0.748	0.137
Model 13	38	3	Gauss-Newton	569.1	0.179	116.100	4.84	1.547	999.3	0.9%	864.7%	861%	0.057	0.303	.-
Model 14	38	3	Gauss-Newton	552.6	1.138	32.504	5.38	0.0391	3.87	1.0%	3.4%	11.9%	0.105	0.080	0.437
Model 15	38	3	Gauss-Newton	538.0	20.168	1.476	5.19	0.676	0.0435	1.0%	3.4%	2.9%	0.074	0.224	0.122

Table E.1 Results of the fitting of the fifteen growth curve models to one shrinkage profile (# 0173) (continued)

Model	SS (model)	df (model)	SS (error)	df (error)	Correlations between parameters (absolute values > 0.90 highlighted)			Corrected SS	Studentised residuals vs. predicted values	Lagged residuals vs. residuals	MSE	AIC	BIC
					corr(α,β)	corr(α,γ)	corr(β,γ)		(no trend, no fanning out)	(no correlation)			
Model 1	5765182	2	19387	36	0.68			1138665	sinusoidal trend, outlier at -3sd	some correlation for lag=1	538.5	10.3	246.3
Model 2	5765182	2	19387	36	-0.68			1138665	sinusoidal trend, outlier at -3sd	correlation at lag=1,2	538.5	10.3	246.3
Model 3	5366254	2	418315	36	0.10			1138665	terrible	correlation at lag=1,2,3	11619.9	13.4	363.0
Model 4	5746372	2	38197	36	0.81			1138665	sinusoidal trend, outlier at -3sd	correlation at lag=1,2,3	1061.0	11.0	272.0
Model 5	5746372	2	38197	36	-0.81			1138665	sinusoidal trend, outlier at -3sd	correlation at lag=1,2,3	1061.0	11.0	272.0
Model 6	5057251	2	727318	36	-0.98			1138665	terrible	correlation at lag=1,2	20203.3	13.9	384.0
Model 7	5698526	2	86043	33	0.86			1138665	sinusoidal trend, outlier at -3sd	correlation at lag=1,2	2607.4	11.9	282.4
Model 8	5766509	3	18060	34	-0.99	-0.76	0.84	1138665	sinusoidal trend, outlier at -3sd	some correlation at lag=1	531.2	12.3	243.0
Model 9	5744766	3	39803	34	0.64	0.81	0.96	1138665	sinusoidal trend, outlier at -3sd	correlation at lag=1,2,3	1170.7	13.1	272.2
Model 10	5768617	3	15952	34	-0.60	-0.43	0.93	1138665	sinusoidal trend, outlier at -3sd	some correlation at lag=1	469.2	12.2	238.4
Model 11	5780715	3	3854	34	-0.81	-0.73	0.98	1138665	good	good	113.4	10.7	185.9
Model 12	5777260	3	7309	34	-0.59	-0.69	0.98	1138665	good	good	215.0	11.4	209.5
Model 13	5779479	3	5090	34	0.61	-0.61	0.99	1138665	sinusoidal trend, outlier at -3sd	some correlation at lag=1	149.7	11.0	201.2
Model 14	5780695	3	3874	34	-0.81	-0.75	0.99	1138665	sinusoidal trend, outlier at -3sd	some correlation at lag=1	113.9	10.7	190.9
Model 15	5778788	3	5781	34	0.83	-0.69	-0.80	1138665	sinusoidal trend, outlier at -3sd	some correlation at lag=1	170.0	11.1	206.1

Table E.1 Results of the fitting of the fifteen growth curve models to one shrinkage profile (# 0173) (continued)

Model	Model curve on data points (good fit)	Observed vs. predicted values (close to linear)	Residuals vs. time (no trend)	Normal plot of residuals	Pseudo-R^2	CCC	RMS	Standard errors of parameters	Correlation between parameters	Skewness of parameters	Comment
Model 1	OK, but data oscillates about curve	OK, but oscillates about unity line	Oscillating trend about 0	OK (oscillates about line)	0.983	0.991	23.21	OK	OK	OK	Model not good
Model 2	OK, but data oscillates about curve	OK, but oscillates about unity line	Oscillating trend about 0	OK (oscillates about line)	0.983	0.991	23.21	OK	OK	OK	Model not good
Model 3	very poor	very poor	very poor	very bad oscillation about line	0.633	0.803	107.80	β high	OK	β high	Model no good at all
Model 4	OK, but data oscillates about curve	OK, but oscillates about unity line	Oscillating trend about 0	Oscillates about line, deviates at high end	0.966	0.982	32.57	β high	high ~ 0.8	OK	Model not good
Model 5	OK, but data oscillates about curve	OK, but oscillates about unity line	Oscillating trend about 0	Oscillates about line, deviates at high end	0.966	0.982	32.57	β high	high ~ 0.8	β high	Model not good
Model 6	very poor	very poor	very poor	very poor	0.361	0.771	142.14	α, β high	high	α, β high	Model no good at all
Model 7	OK, but data oscillates about curve	OK, but oscillates about unity line	Oscillating trend about 0	OK (tails deviate)	0.924	0.956	51.06	β high	high	β high	Model not good
Model 8	Data oscillates about curve; curve does not line out	OK	Oscillating trend about 0	OK	0.984	0.992	23.05	α, β high	αβ high	α high	Model no good at all
Model 9	Data oscillates about curve; curve does not line out	OK, but oscillates about unity line	Oscillating trend about 0	Oscillates about line, deviates at high end	0.965	0.981	34.21	β, γ high	βγ high	α, β high	Reduces to Model 5 since γ very small
Model 10	OK, but data oscillates about curve	OK, but oscillates about unity line	Oscillating trend about 0	OK	0.986	0.993	21.66	β, γ high	βγ high	β,γ high	Model not good
Model 11	excellent	very good	Oscillates, but much better than above models	good	0.997	0.998	10.65	g high	βγ high	γ high	Model OK
Model 12	good	good	Oscillates, but much better than above models	good	0.994	0.997	14.66	β high	βγ high	β high	Model OK
Model 13	good	good	Oscillating trend about 0	Oscillates about line, deviates at high end	0.996	0.998	12.24	β, γ VERY high	βγ high	β high, γ nd	A good fit, but γ = 116 seems unreasonable
Model 14	good	good	Oscillates, but much better than above models	tails deviate	0.997	0.998	10.67	γ high	βγ high	γ high	A good fit, but γ = 30 seems unreasonable
Model 15	good	good	Oscillating trend about 0	Oscillates about line, deviates at high end	0.995	0.997	13.04	OK	OK	OK	Good model

Appendix F

Results of bivariate plots of the covariates versus one another

Table F.1 Results of bivariate plots of the covariates versus one another

<table>
<tr><th>Variables</th><th>Cement Type</th><th>Cement Early Strength Gain Class</th><th>Stone Type</th><th>Stone Source</th><th>Stone Size</th><th>Stone Relative Density</th><th>Stone Loose Bulk Density</th><th>Stone Consolidated Bulk Density</th><th>In(Stone Water Absorption)</th><th>Stone Coeff of Thermal Expansion</th></tr>
<tr><td>Cement Type</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Cement Strength Class</td><td rowspan="12">Good spread across Cement Type</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Cement Early Strength Gain Class</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Stone Type</td><td rowspan="10">Good spread across Cement Early Strength Gain Class</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Stone Source</td><td>Highly correlated</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Stone Size</td><td>Non-19mm stone sizes linked to a particular stone type</td><td rowspan="8">Similar observations as for Stone Type, since these variables are highly correlated.</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Stone Relative Density</td><td rowspan="7">Correlated to some extent as expected</td><td rowspan="7">OK, given the limitations of this variable</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Stone Loose Bulk Density</td><td>Positive correlation</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Stone Consolidated Bulk Density</td><td>Positive correlation</td><td>Positive correlation</td><td></td><td></td><td></td></tr>
<tr><td>In (Stone Water Absorption)</td><td>No trend</td><td>No trend</td><td>No trend</td><td></td><td></td></tr>
<tr><td>Stone Coeff of Thermal Expansion</td><td>Negative correlation</td><td>Negative correlation</td><td>Negative correlation</td><td rowspan="3">No trend</td><td></td></tr>
<tr><td>Stone Unconfined Compression</td><td>No trend</td><td>No trend</td><td>Mild positive correlation</td><td>Mild negative correlation</td></tr>
<tr><td>Stone 10% FACT</td><td>Positive correlation</td><td>Mild positive correlation</td><td>Positive correlation</td><td>Negative correlation</td></tr>
</table>

Table F.1 Results of bivariate plots of the covariates versus one another (continued)

<table>
<tr><th>Variables</th><th>Cement Type</th><th>Cement Early Strength Gain Class</th><th>Stone Type</th><th>Stone Source</th><th>Stone Size</th><th>Stone Relative Density</th><th>Stone Loose Bulk Density</th><th>Stone Consolidated Bulk Density</th><th>ln(Stone Water Absorption)</th><th>Stone Coeff of Thermal Expansion</th></tr>
<tr><td>Stone Elastic Modulus</td><td rowspan="8">Good spread across Cement Type</td><td rowspan="19">Good spread across Cement Early Strength Gain Class</td><td rowspan="2">Correlated to some extent as expected</td><td rowspan="19">Similar observations as for Stone Type, since these variables are highly correlated.</td><td rowspan="19">OK, given the limitations of this variable</td><td>Positive correlation</td><td>Mild positive correlation</td><td>Positive correlation</td><td rowspan="2">No trend</td><td>Mild negative correlation</td></tr>
<tr><td>Stone Flakiness Index</td><td>No trend</td><td>No trend</td><td>No trend</td><td>No trend</td></tr>
<tr><td>Sand Type</td><td rowspan="6">Good spread across Stone Type</td><td rowspan="4">Not correlated, as expected</td><td rowspan="4">Not correlated, as expected</td><td rowspan="4">Not correlated, as expected</td><td rowspan="4">Not correlated, as expected</td><td rowspan="4">Not correlated, as expected</td></tr>
<tr><td>Sand Source</td></tr>
<tr><td>Sand Relative Density</td></tr>
<tr><td>Sand Fineness Modulus</td></tr>
<tr><td>Admixture Type</td><td rowspan="13">Good spread across Stone Relative Density</td><td rowspan="13">Good spread across Stone Loose Bulk Density</td><td rowspan="13">Good spread across Stone Consolidated Bulk Density</td><td rowspan="13">Good spread across ln(Stone Water Absorption)</td><td rowspan="13">Good spread across Stone Coefficient of Thermal Expansion</td></tr>
<tr><td>ln (Cement Content)</td></tr>
<tr><td>CSF Content</td><td rowspan="6">Correlated with Cement Type by definition</td><td>Non-zero only for dolerite & Wits Quartzite</td></tr>
<tr><td>FA Content</td><td rowspan="2">Good spread across StoneType</td></tr>
<tr><td>GGBS Content</td></tr>
<tr><td>GGCS Content</td><td>Non-zero only for andesite & greywacke</td></tr>
<tr><td>GGFS Content</td><td>Non-zero only for andesite</td></tr>
<tr><td>Limestone Content</td><td>Non-zero only for dolerite and andesite</td></tr>
<tr><td>Water Content</td><td rowspan="5">Good spread across Cement Type</td><td rowspan="5">Good spread across Stone Type</td></tr>
<tr><td>Sand Content</td></tr>
<tr><td>Stone Content</td></tr>
<tr><td>Admix Content</td></tr>
<tr><td>PVF</td></tr>
</table>

Table F.1 Results of bivariate plots of the covariates versus one another (continued)

<table>
<tr><th>Variables</th><th>Cement Type</th><th>Cement Early Strength Gain Class</th><th>Stone Type</th><th>Stone Source</th><th>Stone Size</th><th>Stone Relative Density</th><th>Stone Loose Bulk Density</th><th>Stone Consolidated Bulk Density</th><th>ln(Stone Water Absorption)</th><th>Stone Coeff of Thermal Expansion</th></tr>
<tr><td>CSF Fraction</td><td rowspan="6">Correlated with Cement Type by definition</td><td rowspan="19">Good spread across Cement Early Strength Gain Class</td><td>as above</td><td rowspan="19">Similar observations as for Stone Type, since these variables are highly correlated.</td><td rowspan="11">OK, given the limitations of this variable</td><td rowspan="14">Good spread across Stone Relative Density</td><td rowspan="14">Good spread across Stone Loose Bulk Density</td><td rowspan="14">Good spread across Stone Consolidated Bulk Density</td><td rowspan="14">Good spread across Stone Water Absorption</td><td rowspan="14">Good spread across Stone Coefficient of Thermal Expansion</td></tr>
<tr><td>FA Fraction</td><td rowspan="2">Good spread across Stone Type</td></tr>
<tr><td>GGBS Fraction</td></tr>
<tr><td>GGCS Fraction</td><td>as above</td></tr>
<tr><td>GGFS Fraction</td><td>as above</td></tr>
<tr><td>Limestone Fraction</td><td>as above</td></tr>
<tr><td>Water/Binder Ratio</td><td rowspan="13">Good spread across Cement Type</td><td rowspan="5">Good spread across Stone Type</td></tr>
<tr><td>Aggregate/Binder Ratio</td></tr>
<tr><td>Stone/Sand Ratio</td></tr>
<tr><td>Laboratory</td></tr>
<tr><td>Age at first drying</td></tr>
<tr><td>ln (2* Volume To Surface Area)</td><td>Values > 60 only for dolerite</td><td>Correlated</td></tr>
<tr><td>Temperature</td><td rowspan="7">Good spread across Stone Type</td><td rowspan="2">OK</td></tr>
<tr><td>Humidity</td></tr>
<tr><td>Slump</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="2">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td></tr>
<tr><td>28-day Compressive Strength</td></tr>
<tr><td>α</td><td>Mild negative correlation?</td></tr>
<tr><td>ln(β)</td><td rowspan="2">No trend</td></tr>
<tr><td>ln(γ)</td></tr>
</table>

Table F.1 Results of bivariate plots of the covariates versus one another (continued)

<table>
<tr><th>Variables</th><th>Stone 10% FACT</th><th>Stone Elastic Modulus</th><th>Stone Flakiness Index</th><th>Sand Type</th><th>Sand Source</th><th>Sand Relative Density</th><th>Sand Fineness Modulus</th></tr>
<tr><td>Stone 10% FACT</td><td></td><td rowspan="2"></td><td rowspan="3"></td><td rowspan="4"></td><td rowspan="5"></td><td rowspan="6"></td><td rowspan="7"></td></tr>
<tr><td>Stone Elastic Modulus</td><td>Positive correlation</td></tr>
<tr><td>Stone Flakiness Index</td><td>No trend</td><td>No trend</td></tr>
<tr><td>Sand Type</td><td rowspan="4">Not correlated, as expected</td><td rowspan="4">Not correlated, as expected</td><td rowspan="4">Not correlated, as expected</td></tr>
<tr><td>Sand Source</td><td>Highly correlated</td></tr>
<tr><td>Sand Relative Density</td><td rowspan="2">Correlated to some extent as expected</td><td rowspan="20">Similar observations as for Sand Type, since these variables are highly correlated.</td></tr>
<tr><td>Sand Fineness Modulus</td><td>No trend</td></tr>
<tr><td>Admixture Type</td><td rowspan="18">Good spread across Stone 10% FACT</td><td rowspan="18">Good spread across Stone Elastic Modulus</td><td rowspan="18">Good spread across Stone Flakiness Index</td><td>Correlated to some extent</td><td rowspan="18">Good spread across Sand Relative Density</td><td rowspan="18">Good spread across Sand Fineness Modulus</td></tr>
<tr><td>In (Cement Content)</td><td>Good spread across Sand Type</td></tr>
<tr><td>CSF Content</td><td>Non-zero only for dolerite and Wits quartzite/river sand 75/25</td></tr>
<tr><td>FA Content</td><td rowspan="2">Good spread across Sand Type</td></tr>
<tr><td>GGBS Content</td></tr>
<tr><td>GGCS Content</td><td>Non-zero only for Klipheuwel pit sand and natural sand</td></tr>
<tr><td>GGFS Content</td><td>Non-zero only for natural sand</td></tr>
<tr><td>Limestone Content</td><td>Non-zero only for dolomite & granite</td></tr>
<tr><td>Water Content</td><td rowspan="5">Good spread across Sand Type</td></tr>
<tr><td>Sand Content</td></tr>
<tr><td>Stone Content</td></tr>
<tr><td>Admix Content</td></tr>
<tr><td>PVF</td></tr>
<tr><td>CSF Fraction</td><td>as above</td></tr>
<tr><td>FA Fraction</td><td rowspan="2">Good spread across Sand Type</td></tr>
<tr><td>GGBS Fraction</td></tr>
<tr><td>GGCS Fraction</td><td>as above</td></tr>
<tr><td>GGFS Fraction</td><td>as above</td></tr>
</table>

Table F.1 Results of bivariate plots of the covariates versus one another (continued)

<table>
<tr><th>Variables</th><th>Stone Unconfined Compression</th><th>Stone 10% FACT</th><th>Stone Elastic Modulus</th><th>Stone Flakiness Index</th><th>Sand Type</th><th>Sand Source</th><th>Sand Relative Density</th><th>Sand Fineness Modulus</th></tr>
<tr><td>Limestone Content</td><td rowspan="20">Good spread across Stone Unconfined Compression</td><td rowspan="20">Good spread across Stone 10% FACT</td><td rowspan="20">Good spread across Stone Elastic Modulus</td><td rowspan="20">Good spread across Stone Flakiness Index</td><td>Non-zero only for dolomite & granite</td><td rowspan="25">Similar observations as for Sand Type, since these variables are highly correlated.</td><td rowspan="20">Good spread across Sand Relative Density</td><td rowspan="20">Good spread across Sand Fineness Modulus</td></tr>
<tr><td>Water Content</td><td rowspan="5">Good spread across Sand Type</td></tr>
<tr><td>Sand Content</td></tr>
<tr><td>Stone Content</td></tr>
<tr><td>Admix Content</td></tr>
<tr><td>PVF</td></tr>
<tr><td>CSF Fraction</td><td>as above</td></tr>
<tr><td>FA Fraction</td><td rowspan="2">Good spread across Sand Type</td></tr>
<tr><td>GGBS Fraction</td></tr>
<tr><td>GGCS Fraction</td><td>as above</td></tr>
<tr><td>GGFS Fraction</td><td>as above</td></tr>
<tr><td>Limestone Fraction</td><td>as above</td></tr>
<tr><td>Water/Binder Ratio</td><td rowspan="8">Good spread across Sand Type</td></tr>
<tr><td>Aggregate/Binder Ratio</td></tr>
<tr><td>Stone/Sand Ratio</td></tr>
<tr><td>Laboratory</td></tr>
<tr><td>Age at first drying</td></tr>
<tr><td>ln (2* Volume To Surface Area)</td></tr>
<tr><td>Temperature</td></tr>
<tr><td>Humidity</td></tr>
<tr><td>Slump</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="2">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td></tr>
<tr><td>28-day Compressive Strength</td></tr>
<tr><td>α</td><td>Mild negative correlation?</td></tr>
<tr><td>ln(β)</td><td rowspan="2">No trend</td></tr>
<tr><td>ln(γ)</td></tr>
</table>

Table F.1 Results of bivariate plots of the covariates versus one another (continued)

Variables	Admixture Type	ln (Cement Content)	CSF Content	FA Content	GGBS Content	GGCS Content	GGFS Content	Limestone Content	Water Content	Sand Content
Cement Content	Good spread across Admixture Type									
CSF Content		Sub-groups of data are linearly correlated (by definition of recipe)								
FA Content			**Non-zeros mutually exclusive**							
GGBS Content			Good spread	Good spread						
GGCS Content			**Non-zeros mutually exclusive**	**Non-zeros mutually exclusive**	**Non-zeros mutually exclusive**					
GGFS Content						**Non-zeros mutually exclusive**				
Limestone Content							**Non-zeros mutually exclusive**			
Water Content		No trend	Good spread	Good spread	Good spread	Good spread	Good spread	Good spread		
Sand Content									No trend	
Stone Content							**Constant for non-zero**			**Negative correlation**
Admix Content							**Non-zeros mutually exclusive**	**Non-zeros mutually exclusive**		Good spread
PVF		**Positive correlation (by definition)**		**Positive correlation for non-zero (by definition)**	**Positive correlation for non-zero (by definition)**	**Positive correlation for non-zero (by definition)**	**Positive correlation for non-zero (by definition)**	**Positive correlation for non-zero (by definition)**	**Positive correlation (by definition)**	**Negative correlation**
CSF Fraction		Sub-groups of data are linearly correlated (by definition of recipe)	**Totally correlated as expected**	**Non-zeros mutually exclusive**	**Non-zeros mutually exclusive**	**Non-zeros mutually exclusive**		**Non-zeros are mutually exclusive**	Good spread	Good spread
FA Fraction			**Non-zeros mutually exclusive**	**Positive correlation (by definition)**			**Non-zeros mutually exclusive**			
GGBS Fraction			Good spread	Good spread	**Positive correlation (by definition)**	**Non-zeros mutually exclusive**	**Non-zeros mutually exclusive**			
GGCS Fraction			**Non-zeros mutually exclusive**	**Non-zeros mutually exclusive**	**Non-zeros mutually exclusive**	**Positive correlation (by definition)**				

Table F.1 Results of bivariate plots of the covariates versus one another (continued)

<table>
<tr><th>Variables</th><th>Admixture Type</th><th>ln (Cement Content)</th><th>CSF Content</th><th>FA Content</th><th>GGBS Content</th><th>GGCS Content</th><th>GGFS Content</th><th>Limestone Content</th><th>Water Content</th></tr>
<tr><td>GGFS Fraction</td><td rowspan="5">Good spread across Admixture Type</td><td rowspan="2">Sub-groups of data are linearly correlated (by definition of recipe)</td><td rowspan="2">Non-zeros mutually exclusive</td><td rowspan="2">Non-zeros mutually exclusive</td><td rowspan="2">Non-zeros mutually exclusive</td><td rowspan="2">Non-zeros mutually exclusive</td><td>Positive correlation (by definition)</td><td>Non-zeros mutually exclusive</td><td rowspan="2">Good spread</td></tr>
<tr><td>Limestone Fraction</td><td>Non-zeros mutually exclusive</td><td>Positive correlation (by definition)</td></tr>
<tr><td>Water/Binder Ratio</td><td>Negative correlation (by definition)</td><td rowspan="8">Good spread</td><td rowspan="8">Good spread</td><td rowspan="8">Good spread</td><td rowspan="3">Good spread</td><td rowspan="3">Good spread</td><td rowspan="3">Good spread</td><td rowspan="3">No trend</td></tr>
<tr><td>Aggregate/Binder Ratio</td><td>Negative correlation (by definition)</td></tr>
<tr><td>Stone/Sand Ratio</td><td rowspan="7">No trend</td></tr>
<tr><td>Laboratory</td><td>Mutually exclusive except for '-'</td><td>UCT only</td><td>UCT only</td><td>Non-zero for WITS only</td><td rowspan="5">Good spread</td></tr>
<tr><td>Age at first drying</td><td>Good spread across Admixture Type</td><td rowspan="4">Good spread</td><td rowspan="4">Constant for non-zero</td><td rowspan="2">Constant for non-zero</td></tr>
<tr><td>ln (2* Volume To Surface Area)</td><td>High values correlated with one admix type</td></tr>
<tr><td>Temperature</td><td rowspan="2">Good spread across Admixture Type</td><td rowspan="2">Good spread</td></tr>
<tr><td>Humidity</td></tr>
<tr><td>Slump</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td></tr>
<tr><td>28-day Compressive Strength</td><td>Positive correlation</td></tr>
<tr><td>α</td><td rowspan="3">No trend</td></tr>
<tr><td>ln(β)</td></tr>
<tr><td>ln(γ)</td></tr>
</table>

Table F.1 Results of bivariate plots of the covariates versus one another (continued)

<table>
<tr><th>Variables</th><th>Stone Content</th><th>Admix Content</th><th>PVF</th><th>CSF Fraction</th><th>FA Fraction</th><th>GGBS Fraction</th><th>GGCS Fraction</th><th>GGFS Fraction</th><th>Lime-stone Fraction</th><th>Water/ Binder Ratio</th><th>Aggregate/ Binder Ratio</th></tr>
<tr><td>Admix Content</td><td>Good spread</td><td></td><td rowspan="2"></td><td rowspan="3"></td><td rowspan="4"></td><td rowspan="5"></td><td rowspan="6"></td><td rowspan="7"></td><td rowspan="8"></td><td rowspan="9"></td><td rowspan="10"></td></tr>
<tr><td>PVF</td><td>No trend</td><td rowspan="15">Good spread</td></tr>
<tr><td>CSF Fraction</td><td rowspan="7">Good spread</td><td rowspan="6">Good spread</td></tr>
<tr><td>FA Fraction</td><td>Non-zeros mutually exclusive</td></tr>
<tr><td>GGBS Fraction</td><td>Good spread</td><td>Good spread</td></tr>
<tr><td>GGCS Fraction</td><td rowspan="3">Non-zeros mutually exclusive</td><td rowspan="3">Non-zeros mutually exclusive</td><td rowspan="3">Non-zeros mutually exclusive</td></tr>
<tr><td>GGFS Fraction</td><td rowspan="2">Non-zeros mutually exclusive</td></tr>
<tr><td>Limestone Fraction</td><td>Non-zeros mutually exclusive</td></tr>
<tr><td>Water/Binder Ratio</td><td>Negative correlation</td><td rowspan="8">Good spread</td><td rowspan="8">Good spread</td><td rowspan="8">Good spread</td><td rowspan="3">Good spread</td><td rowspan="3">Good spread</td><td rowspan="8">Good spread</td></tr>
<tr><td>Aggregate/Binder Ratio</td><td>No trend</td><td>Negative correlation</td><td>Positive correlation as expected</td></tr>
<tr><td>Stone/Sand Ratio</td><td>Positive correlation as expected</td><td>No trend</td><td rowspan="6">Good spread</td><td rowspan="6">Good spread</td></tr>
<tr><td>Laboratory</td><td rowspan="5">Good spread</td><td rowspan="5">Good spread</td><td>UCT only</td><td>UCT only</td></tr>
<tr><td>Age at first drying</td><td rowspan="4">Good spread</td><td rowspan="4">Constant for non-zero</td></tr>
<tr><td>ln (2* Volume To Surface Area)</td></tr>
<tr><td>Temperature</td></tr>
<tr><td>Humidity</td></tr>
<tr><td>Slump</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td>No trend</td><td>No trend</td></tr>
<tr><td>28-day Compressive Strength</td><td>Negative correlation</td><td>Negative correlation</td></tr>
<tr><td>α</td><td rowspan="3">No trend</td><td rowspan="3">No trend</td></tr>
<tr><td>ln(β)</td></tr>
<tr><td>ln(γ)</td></tr>
</table>

Table F.1 Results of bivariate plots of the covariates versus one another (continued)

<table>
<tr><th>Variable</th><th>Stone/Sand Ratio</th><th>Laboratory</th><th>Age at first drying</th><th>ln (2* Volume To Surface Area)</th><th>Temperature</th><th>Humidity</th><th>Slump</th><th>28-day Compressive Strength</th></tr>
<tr><td>Laboratory</td><td rowspan="5">Good spread</td><td></td><td rowspan="2"></td><td rowspan="3"></td><td rowspan="4"></td><td rowspan="5"></td><td rowspan="6"></td><td rowspan="7"></td></tr>
<tr><td>Age at first drying</td><td rowspan="4">Good spread</td></tr>
<tr><td>ln (2* Volume To Surface Area)</td><td rowspan="3">Good spread</td></tr>
<tr><td>Temperature</td><td rowspan="2">Good spread</td></tr>
<tr><td>Humidity</td><td>Good spread</td></tr>
<tr><td>Slump</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td><td rowspan="5">No trend</td></tr>
<tr><td>28-day Compressive Strength</td><td rowspan="4">No trend</td></tr>
<tr><td>α</td><td rowspan="3">No trend</td></tr>
<tr><td>ln(β)</td></tr>
<tr><td>ln(γ)</td></tr>
</table>

Appendix G

Correlation coefficients for the covariates versus one another and versus the three dependent variables, α, ln(β) and ln(γ)

Correlation coefficients with absolute values of 0.70 or greater are highlighted in the table.

Table G.1 Correlation coefficients for the covariates versus one another and versus the three dependent variables, α, ln(β) and ln(γ)

Variables	Stone Size	Stone Relative Density	Stone Loose Bulk Density	Stone Consolidated Bulk Density	ln (Stone Water Absorption)	Stone Coeff of Thermal Expansion	Stone Unconfined Compression	Stone 10% FACT	Stone Elastic Modulus	Stone Flakiness Index	Sand Relative Density	Sand Fineness Modulus	ln (Cement Content)
Stone Size	1.00												
Stone Relative Density	0.13	1.00											
Stone Loose Bulk Density	-0.10	0.92	1.00										
Stone Consolidated Bulk Density	-0.13	0.92	0.83	1.00									
ln (Stone Water Absorption)	0.12	0.25	0.33	0.53	1.00								
Stone Coeff of Thermal Expansion	-0.10	-0.72	-0.72	-0.80	-0.51	1.00							
Stone Unconfined Compression	-0.06	0.50	0.37	0.68	0.34	-0.67	1.00						
Stone 10% FACT	0.05	0.76	0.67	0.87	0.27	-0.72	0.85	1.00					
Stone Elastic Modulus	0.02	0.72	0.66	0.71	-0.08	-0.66	0.74	0.72	1.00				
Stone Flakiness Index	-0.05	0.35	0.12	0.49	-0.17	-0.06	0.48	0.55	0.34	1.00			
Sand Relative Density	0.31	0.25	0.23	-0.08	0.18	0.01	-0.56	-0.26	-0.15	-0.25	1.00		
Sand Fineness Modulus	0.18	0.46	0.56	0.56	0.19	-0.80	0.60	0.64	0.57	0.07	-0.14	1.00	
ln (Cement Content)	-0.06	-0.08	-0.10	-0.14	0.17	0.06	-0.06	-0.14	-0.13	-0.16	0.16	-0.22	1.00
CSF Content	0.32	0.00	-0.08	-0.13	-0.11	0.07	-0.16	-0.05	-0.06	0.00	0.16	0.08	-0.13
FA Content	0.02	0.21	0.18	0.19	0.22	-0.13	-0.02	0.08	0.01	-0.03	0.21	-0.10	-0.03
GGBS Content	0.00	0.05	0.10	0.01	-0.15	0.00	-0.20	0.00	0.01	0.05	0.13	0.16	-0.44
GGCS Content	-0.03	0.02	0.01	0.11	-0.04	-0.02	0.19	0.14	0.08	0.16	-0.23	0.04	-0.22
GGFS Content	-0.03	0.16	0.16	0.23	-0.02	-0.19	0.30	0.29	0.24	0.17	-0.32	0.25	-0.14
Limestone Content	-0.02	0.14	0.13	0.12	0.08	-0.10	0.04	0.10	0.07	0.02	0.09	0.01	0.11
Water Content	-0.10	0.28	0.28	0.23	0.26	-0.27	-0.06	0.16	-0.02	-0.11	0.28	0.14	0.39
Sand Content	0.28	0.08	0.02	-0.02	0.07	-0.06	0.01	-0.03	0.08	-0.02	0.30	0.02	-0.34
Stone Content	0.12	0.06	-0.02	0.00	-0.10	0.15	-0.19	-0.10	-0.10	0.14	0.02	-0.19	0.01
Admix Content	0.10	-0.38	-0.38	-0.39	0.03	0.14	-0.17	-0.37	-0.25	-0.25	0.09	-0.13	0.06
PVF	-0.11	0.22	0.22	0.20	0.19	-0.16	-0.04	0.13	0.01	-0.04	0.21	0.03	0.57
CSF Fraction	0.34	0.01	-0.08	-0.13	-0.11	0.06	-0.16	-0.05	-0.06	0.00	0.16	0.08	-0.13
FA Fraction	0.03	0.17	0.13	0.16	0.21	-0.10	-0.03	0.05	-0.01	-0.04	0.19	-0.13	-0.08
GGBS Fraction	0.02	0.03	0.08	0.00	-0.18	0.01	-0.20	0.00	0.01	0.05	0.11	0.18	-0.54
GGCS Fraction	-0.03	0.03	0.02	0.12	-0.04	-0.03	0.20	0.16	0.09	0.16	-0.24	0.05	-0.28
GGFS Fraction	-0.03	0.16	0.17	0.24	-0.02	-0.20	0.32	0.31	0.26	0.18	-0.34	0.26	-0.20
Limestone Fraction	-0.02	0.15	0.15	0.13	0.10	-0.10	0.02	0.09	0.06	0.01	0.11	0.01	0.08
Water/Binder Ratio	0.07	0.00	0.01	0.02	-0.04	-0.07	0.11	0.07	0.07	0.02	-0.11	0.16	-0.64
Aggregate/Binder Ratio	0.17	0.00	-0.02	-0.01	-0.08	-0.01	0.10	0.03	0.10	0.07	-0.08	0.08	-0.63
Stone/Sand Ratio	-0.16	-0.04	-0.02	0.01	-0.10	0.12	-0.10	-0.03	-0.10	0.08	-0.22	-0.09	0.30

Table G.1 Correlation coefficients for the covariates versus one another and versus the three dependent variables, α, ln(β) and ln(γ) (continued)

Variables	Stone Size	Stone Relative Density	Stone Loose Bulk Density	Stone Consolidated Bulk Density	ln (Stone Water Absorption)	Stone Coeff of Thermal Expansion	Stone Unconfined Compression	Stone 10% FACT	Stone Elastic Modulus	Stone Flakiness Index	Sand Relative Density	Sand Fineness Modulus	ln (Cement Content)
Age at first drying	0.17	0.50	0.60	0.36	0.05	-0.62	0.16	0.43	0.44	-0.18	0.22	0.79	-0.14
ln (2* Volume To Surface Area)	0.74	0.30	0.37	0.23	0.24	-0.25	-0.09	0.11	0.08	-0.19	0.53	0.19	-0.07
Temperature	-0.18	-0.22	-0.33	-0.02	-0.20	0.21	0.32	0.10	0.07	0.40	-0.69	-0.15	0.02
Humidity	0.12	0.03	0.10	-0.14	0.14	-0.16	-0.30	-0.17	-0.07	-0.46	0.60	0.10	0.08
Slump	-0.20	0.20	0.29	0.37	0.27	-0.43	0.61	0.44	0.39	0.08	-0.34	0.34	-0.14
28-day Compressive Strength	-0.15	-0.05	-0.03	-0.02	0.04	-0.01	0.08	-0.02	0.00	-0.02	-0.10	-0.01	0.71
α	0.18	-0.22	-0.40	-0.11	-0.12	-0.06	0.18	0.09	-0.01	0.15	-0.37	0.21	-0.03
ln(β)	-0.15	-0.23	-0.16	-0.29	0.18	0.25	-0.26	-0.36	-0.37	-0.26	0.26	-0.44	0.36
ln(γ)	0.03	0.03	0.09	0.04	0.25	0.09	-0.03	-0.06	-0.15	-0.03	0.09	-0.25	0.11

Table G.1 Correlation coefficients for the covariates versus one another and versus the three dependent variables, α, ln(β) and ln(γ) (continued)

Variables	CSF Content	FA Content	GGBS Content	GGCS Content	GGFS Content	Lime-stone Content	Water Content	Sand Content	Stone Content	Admix Content	PVF	CSF Fraction	FA Fraction	GGBS Fraction
CSF Content	1.00													
FA Content	-0.05	1.00												
GGBS Content	0.08	-0.15	1.00											
GGCS Content	-0.03	-0.09	-0.15	1.00										
GGFS Content	-0.03	-0.07	-0.12	-0.05	1.00									
Limestone Content	-0.02	-0.04	-0.07	-0.03	-0.03	1.00								
Water Content	0.01	0.03	0.22	-0.14	-0.07	0.17	1.00							
Sand Content	0.01	0.02	-0.14	-0.09	-0.11	-0.04	-0.35	1.00						
Stone Content	0.19	0.00	-0.04	-0.01	0.00	-0.03	-0.10	-0.37	1.00					
Admix Content	-0.02	-0.07	-0.13	0.01	-0.11	-0.07	-0.34	0.20	-0.24	1.00				
PVF	-0.06	0.15	0.24	-0.02	-0.01	0.15	0.81	-0.58	-0.04	-0.25	1.00			
CSF Fraction	1.00	-0.05	0.07	-0.03	-0.03	-0.02	0.01	0.02	0.19	-0.02	-0.06	1.00		
FA Fraction	-0.05	0.97	-0.16	-0.09	-0.07	-0.05	-0.04	0.07	0.01	-0.05	0.06	-0.05	1.00	
GGBS Fraction	0.11	-0.16	0.97	-0.15	-0.12	-0.08	0.15	-0.06	-0.02	-0.12	0.12	0.10	-0.17	1.00
GGCS Fraction	-0.04	-0.09	-0.15	0.96	-0.06	-0.03	-0.17	-0.04	-0.01	-0.02	-0.09	-0.04	-0.10	-0.16
GGFS Fraction	-0.03	-0.08	-0.12	-0.06	0.95	-0.03	-0.11	-0.06	0.00	-0.12	-0.09	-0.03	-0.08	-0.13
Limestone Fraction	-0.02	-0.05	-0.08	-0.03	-0.03	0.96	0.15	-0.01	-0.03	-0.07	0.11	-0.02	-0.05	-0.08
Water/Binder Ratio	0.10	-0.09	-0.15	-0.04	-0.01	-0.07	-0.39	0.60	-0.04	0.02	-0.83	0.11	-0.02	-0.03
Aggregate/Binder Ratio	0.10	-0.06	-0.20	-0.02	-0.01	-0.09	-0.59	0.68	0.05	0.08	-0.92	0.11	0.01	-0.08
Stone/Sand Ratio	0.08	-0.03	0.10	0.06	0.09	0.01	0.26	-0.91	0.68	-0.24	0.47	0.07	-0.07	0.04
Age at first drying	0.10	0.06	0.19	-0.03	0.08	0.06	0.36	0.02	-0.22	-0.18	0.18	0.10	0.02	0.19
ln (2* Volume To Surface Area)	0.23	0.22	0.10	-0.04	-0.11	0.05	0.13	0.22	0.04	0.03	0.09	0.25	0.23	0.09
Temperature	-0.07	-0.15	-0.15	0.10	0.20	-0.11	-0.33	-0.11	0.19	-0.14	-0.15	-0.08	-0.15	-0.14
Humidity	0.05	0.16	0.20	-0.15	-0.27	0.11	0.31	0.09	-0.25	0.19	0.23	0.05	0.13	0.19
Slump	-0.13	0.11	-0.19	0.07	0.21	-0.03	0.02	0.04	-0.26	-0.17	-0.10	-0.13	0.11	-0.17
28-day Compressive Strength	-0.13	-0.14	-0.12	0.19	0.04	0.04	0.30	-0.59	-0.01	0.16	0.68	-0.13	-0.22	-0.20
α	0.10	-0.16	-0.03	0.02	-0.02	-0.07	0.05	0.05	-0.10	0.04	-0.07	0.10	-0.15	0.00
ln(β)	-0.17	0.05	-0.26	-0.01	-0.17	0.06	0.02	0.04	-0.07	0.35	0.09	-0.16	0.04	-0.31
ln(γ)	-0.11	-0.05	-0.26	-0.03	0.00	-0.02	-0.12	0.20	0.04	-0.09	-0.16	-0.10	-0.02	-0.27

Table G.1 Correlation coefficients for the covariates versus one another and versus the three dependent variables, α, ln(β) and ln(γ) (continued)

Variables	GGCS Fraction	GGFS Fraction	Limestone Fraction	Water/ Binder Ratio	Aggregate/ Binder Ratio	Stone/ Sand Ratio	Age at first drying	ln (2* Volume To Surface Area)	Tempe- rature	Humidity	Slump	28-day Compressive Strength
GGCS Fraction	1.00											
GGFS Fraction	-0.06	1.00										
Limestone Fraction	-0.04	-0.03	1.00									
Water/Binder Ratio	0.05	0.09	-0.01	1.00								
Aggregate/Binder Ratio	0.07	0.09	-0.04	0.95	1.00							
Stone/Sand Ratio	0.01	0.04	-0.01	-0.50	-0.53	1.00						
Age at first drying	-0.02	0.08	0.07	0.11	-0.01	-0.12	1.00					
ln (2* Volume To Surface Area)	-0.05	-0.12	0.06	0.00	0.03	-0.20	0.42	1.00				
Temperature	0.11	0.21	-0.13	-0.04	0.05	0.18	-0.52	-0.57	1.00			
Humidity	-0.17	-0.28	0.12	-0.14	-0.20	-0.20	0.38	0.48	-0.71	1.00		
Slump	0.11	0.24	-0.02	0.25	0.16	-0.17	0.17	-0.11	0.01	-0.09	1.00	
28-day Compressive Strength	0.10	-0.05	0.00	-0.84	-0.81	0.47	-0.10	-0.22	0.14	-0.03	-0.12	1.00
α	0.05	-0.01	-0.11	0.13	0.07	-0.06	0.08	0.02	0.37	-0.15	-0.09	0.00
ln(β)	-0.02	-0.19	0.08	-0.17	-0.14	-0.04	-0.33	-0.20	-0.20	0.16	-0.02	0.22
ln(γ)	0.00	0.02	0.01	0.15	0.20	-0.13	-0.28	0.03	-0.02	-0.19	0.07	-0.08

Appendix H

Partial listing of the coded data set used for the multivariate multiple regression analysis

Table H.1 gives a partial listing (first 27 experiments) of the coded data set used in the multivariate multiple regression, starting with the covariates and then giving the parameters of the fitted growth curve model. The full data set may be found on the accompanying CD in the spreadsheet: 'Data for multivariate multiple regression.xls'.

Table H.1 Partial listing of the coded data set used for the multivariate multiple regression analysis

Data Set	Experiment Identification	Cement Early Strength Gain Class	Cement Type	CEM II A-D	CEM II A-L	CEM II A-M(L)	CEM II A-S GGBS	CEM II A-S GGCS	CEM II A-S GGFS
1	# 0002	0	CEM III A	0	0	0	0	0	0
1	# 0003	0	CEM III A	0	0	0	0	0	0
1	# 0004	0	CEM III A	0	0	0	0	0	0
1	# 0005	0	CEM III A	0	0	0	0	0	0
1	# 0006	0	CEM III A	0	0	0	0	0	0
1	# 0008	0	CEM III A	0	0	0	0	0	0
1	# 0009	0	CEM III A	0	0	0	0	0	0
1	# 0010	0	CEM III A	0	0	0	0	0	0
2	# 0019	0	CEM II B-V	0	0	0	0	0	0
2	# 0021	0	CEM II B-V	0	0	0	0	0	0
2	# 0023	0	CEM III A	0	0	0	0	0	0
3	# 0031	1	CEM I	0	0	0	0	0	0
3	# 0032	1	CEM I	0	0	0	0	0	0
3	# 0033	0	CEM I	0	0	0	0	0	0
3	# 0035	0	CEM II A-V	0	0	0	0	0	0
3	# 0036	0	CEM II A-V	0	0	0	0	0	0
3	# 0037	0	CEM II A-L	0	1	0	0	0	0
3	# 0038	0	CEM II A-L	0	1	0	0	0	0
3	# 0039	0	CEM III A	0	0	0	0	0	0
3	# 0040	0	CEM III A	0	0	0	0	0	0
3	# 0041	0	CEM III A	0	0	0	0	0	0
3	# 0042	0	CEM III A	0	0	0	0	0	0
3	# 0043	0	CEM II B-V	0	0	0	0	0	0
3	# 0044	0	CEM II B-V	0	0	0	0	0	0
3	# 0045	0	CEM III A	0	0	0	0	0	0
3	# 0046	0	CEM III A	0	0	0	0	0	0
3	# 0047	0	CEM II A-S	0	0	0	1	0	0

Table H.1 Partial listing of the coded data set used for the multivariate multiple regression analysis (continued)

CEM II A-V	CEM II B-M(V/L)	CEM II B-S GGBS	CEM II B-S GGCS	CEM II B-S GGFS	CEM II B-V	CEM III A GGBS	CEM III A GGCS	CEM III A GGFS	CEM V A	Stone Type	Stone Wits Quartzite
0	0	0	0	0	0	1	0	0	0	Wits quartzite	1
0	0	0	0	0	0	1	0	0	0	Wits quartzite/shale 50/50	0.5
0	0	0	0	0	0	1	0	0	0	Wits quartzite/shale 50/50	0.5
0	0	0	0	0	0	1	0	0	0	shale	0
0	0	0	0	0	0	1	0	0	0	shale	0
0	0	0	0	0	0	1	0	0	0	Wits quartzite	1
0	0	0	0	0	0	1	0	0	0	Wits quartzite	1
0	0	0	0	0	0	1	0	0	0	Wits quartzite	1
0	0	0	0	0	1	0	0	0	0	dolerite	0
0	0	0	0	0	1	0	0	0	0	dolerite	0
0	0	0	0	0	0	1	0	0	0	dolerite	0
0	0	0	0	0	0	0	0	0	0	dolerite	0
0	0	0	0	0	0	0	0	0	0	dolerite	0
0	0	0	0	0	0	0	0	0	0	dolerite	0
1	0	0	0	0	0	0	0	0	0	dolerite	0
1	0	0	0	0	0	0	0	0	0	dolerite	0
0	0	0	0	0	0	0	0	0	0	dolerite	0
0	0	0	0	0	0	0	0	0	0	dolerite	0
0	0	0	0	0	0	1	0	0	0	dolerite	0
0	0	0	0	0	0	1	0	0	0	dolerite	0
0	0	0	0	0	0	1	0	0	0	dolerite	0
0	0	0	0	0	0	1	0	0	0	dolerite	0
0	0	0	0	0	1	0	0	0	0	dolerite	0
0	0	0	0	0	1	0	0	0	0	dolerite	0
0	0	0	0	0	0	1	0	0	0	dolerite	0
0	0	0	0	0	0	1	0	0	0	dolerite	0
0	0	0	0	0	0	0	0	0	0	dolerite	0

Table H.1 Partial listing of the coded data set used for the multivariate multiple regression analysis (continued)

Stone Shale	Stone Dolerite	Stone Greywacke	Stone Dolomite	Stone Granite	Stone Quartz alluvial	Stone Pretoria Quartzite	Stone Tillite	Stone Quartzite	Stone Size	Stone Relative Density	Stone Loose Bulk Density	Stone Consolidated Bulk Density	ln (Stone Water Absorption)	Stone Coeff of Thermal Expansion
0	0	0	0	0	0	0	0	0	19	2.67	1365	1490	-1.47	11.3
0.5	0	0	0	0	0	0	0	0	19	2.71			-1.02	
0.5	0	0	0	0	0	0	0	0	19	2.71			-1.02	
1	0	0	0	0	0	0	0	0	19	2.75			-0.71	
1	0	0	0	0	0	0	0	0	19	2.75			-0.71	
0	0	0	0	0	0	0	0	0	19	2.67	1365	1490	-1.47	11.3
0	0	0	0	0	0	0	0	0	19	2.67	1365	1490	-1.47	11.3
0	0	0	0	0	0	0	0	0	19	2.67	1365	1490	-1.47	11.3
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8
0	1	0	0	0	0	0	0	0	19	2.98	1578	1693	-0.63	7.8

Table H.1 Partial listing of the coded data set used for the multivariate multiple regression analysis (continued)

Stone Unconfined Compression	Stone 10% FACT	Stone Elastic Modulus	Stone Flakiness Index	Sand Type	Sand Wits Quartzite	Sand Shale	Sand Granite	Sand Dolerite	Sand Andesite	Sand Klipheuwel pit	Sand natural	Sand River Windhoek	Sand Cape Flats
135	220	72	22	Wits quartzite	1	0	0	0	0	0	0	0	0
		86		Wits quartzite	1	0	0	0	0	0	0	0	0
		86		Wits quartzite	1	0	0	0	0	0	0	0	0
		99		Wits quartzite	1	0	0	0	0	0	0	0	0
		99		Wits quartzite	1	0	0	0	0	0	0	0	0
135	220	72	22	Wits quartzite/shale 50/50	0.5	0.5	0	0	0	0	0	0	0
135	220	72	22	shale	0	1	0	0	0	0	0	0	0
135	220	72	22	shale	0	1	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0
285	344	86	20	dolomite	0	0	0	0	0	0	0	0	0

Table H.1 Partial listing of the coded data set used for the multivariate multiple regression analysis (continued)

Sand River Vaal	Sand Pretoria Quartzite	Sand Tillite	Sand Quartzite	Sand River	Sand Ecca Grit	Sand Relative Density	Sand Fineness Modulus	Admix Type	Admix Conplast	Admix ABE/Fosroc	Admix Pozzolith	Admix Embecon	Admix Chryso	Admix Sikament
0	0	0	0	0	0	2.67	3.15	-	0	0	0	0	0	0
0	0	0	0	0	0	2.67	3.15	-	0	0	0	0	0	0
0	0	0	0	0	0	2.67	3.15	-	0	0	0	0	0	0
0	0	0	0	0	0	2.67	3.15	-	0	0	0	0	0	0
0	0	0	0	0	0	2.67	3.15	-	0	0	0	0	0	0
0	0	0	0	0	0	2.71		-	0	0	0	0	0	0
0	0	0	0	0	0	2.75		-	0	0	0	0	0	0
0	0	0	0	0	0	2.75		-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0
0	0	0	0	0	0	2.85	2.93	-	0	0	0	0	0	0

Table H.1 Partial listing of the coded data set used for the multivariate multiple regression analysis (continued)

Admix Sikament NN Super-plasticiser	In (Cement Content)	CSF Content	FA Content	GGBS Content	GGCS Content	GGFS Content	Limestone Content	Water Content	Sand Content	Stone Content	Admix Content	PVF	CSF Fraction	FA Fraction	GGBS Fraction
0	5.42	0	0	227	0	0	0	195	779	1000	0	0.346	0.00	0.00	0.50
0	4.98	0	0	146	0	0	0	195	926	1000	0	0.292	0.00	0.00	0.50
0	5.42	0	0	227	0	0	0	195	779	1000	0	0.346	0.00	0.00	0.50
0	4.98	0	0	146	0	0	0	195	926	1000	0	0.292	0.00	0.00	0.50
0	5.42	0	0	227	0	0	0	195	779	1000	0	0.346	0.00	0.00	0.50
0	5.42	0	0	227	0	0	0	195	780	1000	0	0.346	0.00	0.00	0.50
0	4.98	0	0	146	0	0	0	195	926	1000	0	0.292	0.00	0.00	0.50
0	5.42	0	0	227	0	0	0	195	779	1000	0	0.346	0.00	0.00	0.50
0	5.78	0	139	0	0	0	0	195	773	1106	0	0.359	0.00	0.30	0.00
0	5.86	0	137	0	0	0	0	205	754	1106	0	0.376	0.00	0.28	0.00
0	5.55	0	0	256	0	0	0	215	694	1106	0	0.385	0.00	0.00	0.50
0	6.28	0	0	0	0	0	0	225	665	1106	0	0.396	0.00	0.00	0.00
0	5.76	0	0	0	0	0	0	200	934	1106	0	0.301	0.00	0.00	0.00
0	6.26	0	0	0	0	0	0	220	665	1106	0	0.387	0.00	0.00	0.00
0	6.10	0	67	0	0	0	0	215	694	1106	0	0.386	0.00	0.13	0.00
0	5.60	0	40	0	0	0	0	195	943	1106	0	0.298	0.00	0.13	0.00
0	6.05	0	0	0	0	0	63	205	745	1106	0	0.364	0.00	0.00	0.00
0	5.38	0	0	0	0	0	33	195	943	1106	0	0.276	0.00	0.00	0.00
0	5.55	0	0	256	0	0	0	215	694	1106	0	0.385	0.00	0.00	0.50
0	5.04	0	0	155	0	0	0	195	943	1106	0	0.298	0.00	0.00	0.50
0	5.52	0	0	250	0	0	0	210	720	1106	0	0.376	0.00	0.00	0.50
0	5.04	0	0	155	0	0	0	195	943	1106	0	0.298	0.00	0.00	0.50
0	5.84	0	133	0	0	0	0	200	772	1106	0	0.367	0.00	0.28	0.00
0	5.36	0	82	0	0	0	0	185	987	1106	0	0.288	0.00	0.28	0.00
0	5.54	0	0	259	0	0	0	215	671	1106	0	0.385	0.00	0.00	0.51
0	5.03	0	0	157	0	0	0	195	929	1106	0	0.298	0.00	0.00	0.51
0	6.12	0	0	68	0	0	0	220	682	1106	0	0.389	0.00	0.00	0.13

Table H.1 Partial listing of the coded data set used for the multivariate multiple regression analysis (continued)

GGCS Fraction	GGFS Fraction	Limestone Fraction	Water/ Binder Ratio	Aggregate/ Binder Ratio	Stone/ Sand Ratio	Lab	Lab Wits	Age at first drying	ln(2* Volume To Surface Area)	Tempe-rature	Humidity	Slump	28-day Compressive Strength
0.00	0.00	0.00	0.43	3.92	1.3	Wits	0	28	3.91	23	65	55	50
0.00	0.00	0.00	0.67	6.60	1.1	Wits	0	28	3.91	23	65	75	28
0.00	0.00	0.00	0.43	3.92	1.3	Wits	0	28	3.91	23	65	60	47
0.00	0.00	0.00	0.67	6.60	1.1	Wits	0	28	3.91	23	65	75	30
0.00	0.00	0.00	0.43	3.92	1.3	Wits	0	28	3.91	23	65	50	49
0.00	0.00	0.00	0.43	3.92	1.3	Wits	0	28	3.91	23	65	35	51
0.00	0.00	0.00	0.67	6.60	1.1	Wits	0	28	3.91	23	65	25	32
0.00	0.00	0.00	0.43	3.92	1.3	Wits	0	28	3.91	23	65	10	50
0.00	0.00	0.00	0.42	4.05	1.4	Wits	0	30	3.91	23	65	70	55
0.00	0.00	0.00	0.42	3.81	1.5	Wits	0	31	3.91	23	65	75	57
0.00	0.00	0.00	0.42	3.52	1.6	Wits	0	30	3.91	23	65	70	46
0.00	0.00	0.00	0.42	3.30	1.7	Wits	0	34	3.91	21	65	75	71
0.00	0.00	0.00	0.63	6.44	1.2	Wits	0	34	3.91	21	65	70	46
0.00	0.00	0.00	0.42	3.38	1.7	Wits	0	33	3.91	21	65	65	70
0.00	0.00	0.00	0.42	3.52	1.6	Wits	0	30	3.91	22	66	70	57
0.00	0.00	0.00	0.63	6.61	1.2	Wits	0	30	3.91	21	66	60	30
0.00	0.00	0.13	0.42	3.79	1.5	Wits	0	29	3.91	21	66	85	52
0.00	0.00	0.13	0.78	8.20	1.2	Wits	0	29	3.91	21	66	70	28
0.00	0.00	0.00	0.42	3.52	1.6	Wits	0	29	3.91	21	64	75	52
0.00	0.00	0.00	0.63	6.61	1.2	Wits	0	29	3.91	21	64	55	31
0.00	0.00	0.00	0.42	3.65	1.5	Wits	0	28	3.91	21	64	75	42
0.00	0.00	0.00	0.63	6.61	1.2	Wits	0	28	3.91	21	64	75	25
0.00	0.00	0.00	0.42	3.95	1.4	Wits	0	33	3.91	21	65	75	51
0.00	0.00	0.00	0.63	7.12	1.1	Wits	0	33	3.91	21	65	70	25
0.00	0.00	0.00	0.42	3.47	1.6	Wits	0	32	3.91	21	65	45	48
0.00	0.00	0.00	0.63	6.56	1.2	Wits	0	32	3.91	21	65	65	30
0.00	0.00	0.00	0.42	3.41	1.6	Wits	0	29	3.91	21	61	70	66

Table H.1 Partial listing of the coded data set used for the multivariate multiple regression analysis (continued)

α	β	γ	ln(β)	ln(γ)	weight from generalised variance	normalised weights	# data points (n)	normalised weight based on 'n'	scaled weight based on 'n'
469	0.0073	0.49	-4.91	-0.72	9,718	0.0026	21	0.0078	4.3
624	0.0044	0.40	-5.43	-0.92	24,467	0.0066	21	0.0078	4.3
627	0.0070	0.36	-4.97	-1.01	11,199	0.0030	21	0.0078	4.3
692	0.0047	0.45	-5.35	-0.80	4,529	0.0012	17	0.0063	3.5
658	0.0092	0.41	-4.69	-0.89	2,197	0.0006	17	0.0063	3.5
512	0.0104	0.47	-4.57	-0.76	2,909	0.0008	17	0.0063	3.5
711	0.0052	0.44	-5.27	-0.82	12,868	0.0034	18	0.0066	3.7
629	0.0091	0.42	-4.70	-0.87	22,248	0.0060	18	0.0066	3.7
325	0.0099	0.40	-4.62	-0.92	2,835	0.0008	11	0.0041	2.3
312	0.0081	0.41	-4.82	-0.88	3,258	0.0009	11	0.0041	2.3
356	0.0100	0.42	-4.61	-0.86	619	0.0002	12	0.0044	2.5
364	0.0263	0.88	-3.64	-0.13	3,386	0.0009	12	0.0044	2.5
427	0.0127	0.60	-4.37	-0.52	1,541	0.0004	12	0.0044	2.5
352	0.0194	0.76	-3.94	-0.27	5,796	0.0016	12	0.0044	2.5
210	0.0391	0.84	-3.24	-0.18	1,414	0.0004	12	0.0044	2.5
194	0.0241	0.75	-3.73	-0.28	780	0.0002	12	0.0044	2.5
255	0.0247	0.67	-3.70	-0.41	4,406	0.0012	12	0.0044	2.5
221	0.0279	0.89	-3.58	-0.12	3,358	0.0009	12	0.0044	2.5
256	0.0237	0.83	-3.74	-0.18	768	0.0002	13	0.0048	2.7
242	0.0113	0.42	-4.48	-0.87	1,674	0.0004	13	0.0048	2.7
266	0.0179	0.66	-4.02	-0.41	1,331	0.0004	13	0.0048	2.7
193	0.0196	0.78	-3.93	-0.25	638	0.0002	13	0.0048	2.7
298	0.0402	0.84	-3.21	-0.17	4,205	0.0011	12	0.0044	2.5
259	0.0289	1.20	-3.54	0.18	16,569	0.0044	12	0.0044	2.5
208	0.0287	1.21	-3.55	0.19	831	0.0002	12	0.0044	2.5
169	0.0155	0.95	-4.17	-0.05	1,197	0.0003	12	0.0044	2.5
274	0.0286	0.82	-3.55	-0.20	1,814	0.0005	12	0.0044	2.5

Appendix I

Results of the multivariate multiple regression analyses

Table I.1 gives the results of the multivariate multiple regression analyses from the forward and backward stepwise regressions using each of the covariate sets. The Table shows which variables were eliminated at the start in order to allow the design matrix to be inverted.

Regression models with an average adjusted R^2 of 0.62 or greater are highlighted in yellow. Where there were too many unique observations to calculate the lack of fit, this is indicated by an asterisk (*) in the relevant column.

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
					Forward stepwise regression					After completion of regression				
A1 (without Cement Early Strength Gain Class)	B1	C1	D1-1 (without Admixture variables, Slump)	E1	0.73	0.62	0.50	0.62	*	0.73	0.62	0.50	0.62	None
				E2	0.75	0.60	0.47	0.61	*					
			D1-2 (without Admixture variables, Slump)	E1	0.74	0.63	0.49	0.62	*	0.74	0.63	0.49	0.62	None
				E2	0.69	0.56	0.41	0.55	None					
			D1-3 (without Admixture variables, Slump)	E1	0.73	0.62	0.49	0.61	*					
				E2	0.75	0.60	0.47	0.61	None					
			D1-4 (without Admixture variables, Slump)	E1	0.75	0.62	0.49	0.62	None	0.75	0.62	0.49	0.62	None
				E2	0.69	0.56	0.41	0.55	None					
			D2-1 (without Admixture variables, Slump)	E1	0.71	0.62	0.47	0.60	*					
				E2	0.72	0.58	0.43	0.58	None					
			D2-2 (without Admixture variables, Slump)	E1	0.70	0.62	0.47	0.60	None					
				E2	0.61	0.54	0.36	0.50	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.70	0.63	0.45	0.59	*					
				E2	0.61	0.55	0.35	0.50	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.71	0.62	0.47	0.60	*					
				E2	0.73	0.60	0.44	0.59	None					
A1 (without Cement Early Strength Gain Class)	B1	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.74	0.63	0.36	0.58	*					
				E2 without Age at first drying	0.74	0.61	0.36	0.57	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.74	0.63	0.37	0.58	*					
				E2 without Age at first drying	0.74	0.61	0.37	0.57	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.74	0.63	0.36	0.58	None					
				E2 without Age at first drying	0.74	0.61	0.36	0.57	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.74	0.63	0.37	0.58	None					
				E2 without Age at first drying	0.74	0.61	0.37	0.57	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.72	0.64	0.33	0.56	None					
				E2 without Age at first drying	0.73	0.61	0.33	0.56	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.72	0.64	0.32	0.56	*					
				E2 without Age at first drying	0.73	0.62	0.32	0.56	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1 without Age at first drying	0.72	0.64	0.33	0.56	*					
				E2 without Age at first drying	0.70	0.62	0.34	0.55	None					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1 without Age at first drying	0.71	0.64	0.33	0.56	*					
				E2 without Age at first drying	0.70	0.63	0.33	0.55	None					
A1 (without Cement Early Strength Gain Class)	B2-1 (without Stone Uncon-fined Com-pression and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.73	0.65	0.50	0.63	*	0.74	0.63	0.50	0.62	None
				E2	0.72	0.58	0.47	0.59	None					
			D1-2 (without Admixture variables, Slump)	E1	0.75	0.63	0.51	0.63	*	0.75	0.61	0.51	0.62	None
				E2	0.74	0.57	0.49	0.60						
			D1-3 (without Admixture variables, Slump)	E1	0.75	0.62	0.51	0.63	*	0.75	0.61	0.50	0.62	None
				E2	0.73	0.58	0.49	0.60						
			D1-4 (without Admixture variables, Slump)	E1	0.75	0.62	0.51	0.63	None	0.75	0.61	0.50	0.62	None
				E2	0.75	0.57	0.49	0.60						
			D2-2 (without Admixture variables, Slump)	E1	0.71	0.63	0.49	0.61	None					

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.72	0.64	0.49	0.62	*	0.72	0.63	0.48	0.61	None
				E2	0.72	0.57	0.44	0.58	*					
A1	B2-1	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.78	0.63	0.47	0.63	None	0.43	0.56	0.31	0.43	None
				E2 without Age at first drying	0.65	0.64	0.44	0.58	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.79	0.62	0.48	0.63	None	0.42	0.55	0.33	0.43	None
				E2 without Age at first drying	0.67	0.62	0.45	0.58	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.78	0.63	0.47	0.63	*	reduces to same model as A1/B2-1/C2/D1-1/E1/f				
				E2 without Age at first drying	0.63	0.63	0.41	0.56	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.79	0.62	0.48	0.63	None	reduces to same model as A1/B2-1/C2/D1-2/E1/f				
				E2 without Age at first drying	0.65	0.61	0.40	0.55	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.75	0.65	0.43	0.61	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.73	0.62	0.43	0.59	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.78	0.61	0.43	0.61	*					
			D3-2	E1 without Age at first drying	0.78	0.61	0.43	0.61	*					
A1 (without Cement Early Strength Gain Class)	B2-4 without CTE	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.67	0.60	0.36	0.54	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.67	0.61	0.37	0.55	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.68	0.61	0.36	0.55	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.68	0.61	0.36	0.55	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.75	0.60	0.44	0.60	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.60	0.39	0.59	*					
A1	B2-5 (without Stone LBD and Stone Flakiness Index)	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.72	0.66	0.50	0.63	None	0.39	0.49	0.32	0.40	None
				E2 without Age at first drying	0.63	0.64	0.43	0.57	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.71	0.66	0.51	0.63	None	0.36	0.49	0.31	0.39	None
				E2 without Age at first drying	0.63	0.64	0.47	0.58	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.72	0.66	0.50	0.63	None	reduces to same model as A1/B2-5/C2/D1-1/E1/f				
				E2 without Age at first drying	0.61	0.64	0.43	0.56	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.71	0.66	0.51	0.63	None	reduces to same model as A1/B2-5/C2/D1-2/E1/f				
				E2 without Age at first drying	0.62	0.64	0.41	0.56	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.70	0.66	0.46	0.61	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.68	0.66	0.46	0.60	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.67	0.65	0.48	0.60	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.67	0.65	0.48	0.60	*					
A1	B2-6 (without Stone LBD and Stone Flakiness Index)	C2	D1-4 (without Admixture variables)	E1 without Age at first drying	0.76	0.62	0.45	0.61	None					
A1	B2-7 without CTE	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.78	0.63	0.47	0.63	None	reduces to same model as B2-1/C2/D1-1/E1/f				
				E2 without Age at first drying	0.67	0.63	0.43	0.58	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.79	0.62	0.48	0.63	None	reduces to same model as B2-1/C2/D1-2/E1/f				
				E2 without Age at first drying	0.67	0.62	0.45	0.58	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.78	0.63	0.47	0.63	None	reduces to same model as B2-1/C2/D1-1/E1/f				
				E2 without Age at first drying	0.62	0.62	0.39	0.54	None					

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.79	0.62	0.48	0.63	None	reduces to same model as B2-1/C2/D1-2/E1/f				
				E2 without Age at first drying	0.65	0.61	0.40	0.55	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.75	0.65	0.43	0.61	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.75	0.65	0.43	0.61	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.78	0.61	0.43	0.61	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.61	0.43	0.61	*					
A2	B2-1	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.76	0.56	0.46	0.59	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.77	0.56	0.48	0.60	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.76	0.56	0.46	0.59	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.77	0.56	0.48	0.60	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.76	0.58	0.45	0.60	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.75	0.56	0.43	0.58	None					
			D3-1 (without Admixture variables, Compressive Strength)	E1 without Age at first drying	0.79	0.55	0.44	0.59	*					
			D3-2 (without Admixture variables, Compressive Strength)	E1 without Age at first drying	0.79	0.54	0.43	0.59	*					
A3	B2-1	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.46	0.60	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.46	0.60	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.76	0.58	0.44	0.59	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.76	0.57	0.43	0.59	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.79	0.56	0.44	0.60	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.55	0.43	0.59	*					
A3	B2-2 (without Stone LBD)	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.79	0.58	0.46	0.61	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.79	0.57	0.46	0.61	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.79	0.58	0.43	0.60	None					
A3	B2-5 (without Stone LBD)	C2	D1-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	*					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.79	0.57	0.46	0.61	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
A3	B2-7 without CTE	C2	D1-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.79	0.57	0.46	0.61	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
					Backward stepwise regression					After completion of regression				
A1 (without Cement Early Strength Gain Class)	B1	C1	D1-1 (without Admixture variables, Slump)	E1	0.77	0.63	0.50	0.63	*	0.77	0.63	0.49	0.63	None
				E2	0.76	0.60	0.47	0.61	None					
			D1-2 (without Admixture variables, Slump)	E1	0.77	0.62	0.50	0.63	None	0.77	0.62	0.50	0.63	None
				E2	0.77	0.59	0.48	0.61	None					
			D1-3 (without Admixture variables, Slump)	E1	0.78	0.62	0.49	0.63	None	0.78	0.62	0.49	0.63	None
				E2	0.75	0.61	0.50	0.62	None	0.75	0.61	0.50	0.62	None

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D1-4 (without Admixture variables, Slump)	E1	0.76	0.64	0.53	0.64	None	0.76	0.64	0.53	0.64	None
				E2	0.75	0.61	0.50	0.62	None	reduces to same model as A1/B1/C1D1-3/E2/b				
			D2-1 (without Admixture variables, Slump)	E1	0.76	0.63	0.47	0.62	None	0.76	0.63	0.47	0.62	None
				E2	0.76	0.57	0.42	0.58	None					
			D2-2 (without Admixture variables, Slump)	E1	0.72	0.64	0.49	0.62	None	0.72	0.64	0.49	0.62	None
				E2	0.77	0.58	0.43	0.59	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.76	0.64	0.46	0.62	None	0.76	0.64	0.46	0.62	None
				E2	0.74	0.62	0.46	0.61	None					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.76	0.64	0.47	0.62	None	0.74	0.64	0.47	0.62	None
				E2	0.74	0.62	0.47	0.61	None					
A1 (without Cement Early Strength Gain Class)	B1	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.75	0.63	0.38	0.59	None					
				E2 without Age at first drying	0.74	0.60	0.36	0.57	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.74	0.64	0.40	0.59	None					
				E2 without Age at first drying	0.74	0.61	0.39	0.58	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.74	0.63	0.38	0.58	None					
				E2 without Age at first drying	0.74	0.61	0.38	0.58	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.74	0.63	0.37	0.58	None					
				E2 without Age at first drying	0.74	0.60	0.37	0.57	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.73	0.64	0.32	0.56	None					
				E2 without Age at first drying	0.74	0.60	0.32	0.55	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.72	0.65	0.35	0.57	None					
				E2 without Age at first drying	0.74	0.60	0.33	0.56	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1 without Age at first drying	0.72	0.64	0.33	0.56	None					
				E2 without Age at first drying	0.70	0.63	0.33	0.55	None					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1 without Age at first drying	0.71	0.64	0.33	0.56	None					
				E2 without Age at first drying	0.69	0.62	0.34	0.55	None					
A1 (without Cement Early Strength Gain Class)	B2-1 (without Stone Unconfined Compression and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.73	0.65	0.50	0.63	*	reduces to same model as A1/B2-1/C1/D1-1/f				
				E2	0.72	0.58	0.47	0.59	None					
			D1-2 (without Admixture variables, Slump)	E1	0.75	0.63	0.51	0.63	None	reduces to same model as A1/B2-1/C1/D1-2/f				
				E2	0.74	0.57	0.49	0.60	None					
			D1-3 (without Admixture variables, Slump)	E1	0.75	0.64	0.50	0.63	None	0.75	0.62	0.49	0.62	None
				E2	0.75	0.57	0.48	0.60	None					
			D1-4 (without Admixture variables, Slump)	E1	0.75	0.62	0.51	0.63	None	reduces to same model as A1/B2-1/C1/D1-4/f				
				E2	0.75	0.57	0.49	0.60	None					
			D2-1 (without Admixture variables, Slump)	E1	0.74	0.63	0.47	0.61	None					
			D2-2 (without Admixture variables, Slump)	E1	0.74	0.63	0.48	0.62	None	0.74	0.63	0.47	0.61	None
				E2	0.74	0.58	0.46	0.59	*					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.72	0.64	0.48	0.61	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.72	0.64	0.49	0.62	None	reduces to same model as A1/B2-1/C1/D3-2/E1/f				
				E2	0.72	0.59	0.46	0.59	None					
A1	B2-1	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.81	0.68	0.51	0.67	*	0.52	0.58	0.36	0.49	None
				E2 without Age at first drying	0.83	0.63	0.41	0.62	None	0.45	0.59	0.37	0.47	None

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.82	0.69	0.52	0.68	None	0.54	0.56	0.35	0.48	α
				E2 without Age at first drying	0.83	0.63	0.42	0.63	*	0.48	0.50	0.34	0.44	α
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.84	0.69	0.51	0.68	None	0.51	0.62	0.37	0.50	None
				E2 without Age at first drying	0.82	0.63	0.42	0.62	*	0.44	0.58	0.37	0.46	α
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.82	0.68	0.51	0.67	None	reduces to same model as A1/B2-1/C2/D1-3/E1/b				
				E2 without Age at first drying	0.82	0.62	0.41	0.62	*	reduces to same model as A1/B2-1/C2/D1-2/E1/b				
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.84	0.70	0.50	0.68	*	0.51	0.61	0.34	0.49	None
				E2 without Age at first drying	0.84	0.62	0.37	0.61	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.84	0.69	0.49	0.67	None	reduces to same model as A1/B2-1/C2/D2-1/E1/b				
				E2 without Age at first drying	0.84	0.62	0.37	0.61	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.84	0.67	0.49	0.67	*	0.71	0.66	0.48	0.62	None
				E2 without Age at first drying	0.84	0.61	0.38	0.61	*					
			D3-2	E1 without Age at first drying	0.85	0.68	0.49	0.67	*	reduces to same model as A1/B2-1/C2/D3-1/E1/b				
				E2 without Age at first drying	0.84	0.61	0.38	0.61	*					
A1 (without Cement Early Strength Gain Class)	B2-2 (without Stone LBD and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.72	0.61	0.47	0.60	*					
				E2	0.70	0.56	0.45	0.57	None					
			D1-2 (without Admixture variables, Slump)	E1	0.72	0.61	0.47	0.60	None					
			D1-3 (without Admixture variables, Slump)	E1	0.73	0.60	0.46	0.60	None					
			D1-4 (without Admixture variables, Slump)	E1	0.72	0.60	0.46	0.59	None					
			D2-1 (without Admixture variables, Slump)	E1	0.72	0.60	0.44	0.59	None					
			D2-2 (without Admixture variables, Slump)	E1	0.70	0.61	0.44	0.58	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.71	0.61	0.44	0.59	None					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.69	0.60	0.44	0.58	None					
A1 (without Cement Early Strength Gain Class)	B2-2 (without Stone LBD and Stone Flakiness Index)	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.49	0.59	0.29	0.46	*					
				E2 without Age at first drying	0.43	0.58	0.27	0.43	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.50	0.59	0.28	0.46	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.49	0.59	0.29	0.46	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.50	0.59	0.28	0.46	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.52	0.58	0.27	0.46	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.52	0.29	0.26	0.36	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.73	0.60	0.42	0.58	None					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.72	0.59	0.43	0.58	*					
A1 (without Cement Early Strength Gain	B2-3 (without Stone CDB, Stone UC and Stone	C1	D1-1 (without Admixture variables, Slump)	E1	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2										
			D1-2 (without Admixture variables, Slump)	E1										
			D1-3 (without Admixture variables, Slump)	E1										

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
Class)	Flakiness Index)		D1-4 (without Admixture variables, Slump)	E1										
			D2-1 (without Admixture variables, Slump)	E1										
			D2-2 (without Admixture variables, Slump)	E1										
			D3-1 (without Admixture variables, Slump)	E1										
			D3-2 (without Admixture variables, Slump)	E1										
A1 (without Cement Early Strength Gain Class)	B2-3 (without Stone CDB, Stone UC and Stone Flakiness Index)	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2 without Age at first drying										
			D1-2 (without Admixture variables)	E1 without Age at first drying										
			D1-3 (without Admixture variables)	E1 without Age at first drying										
			D1-4 (without Admixture variables)	E1 without Age at first drying										
			D2-1 (without Admixture variables)	E1 without Age at first drying										
			D2-2 (without Admixture variables)	E1 without Age at first drying										
			D3-1 (without Admixture variables)	E1 without Age at first drying										
			D3-2 (without Admixture variables)	E1 without Age at first drying										
A1 (without Cement Early Strength Gain Class)	B2-4 (without Stone CTE, Stone UC and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2										
			D1-2 (without Admixture variables, Slump)	E1										
			D1-3 (without Admixture variables, Slump)	E1										
			D1-4 (without Admixture variables, Slump)	E1										
			D2-1 (without Admixture variables, Slump)	E1										
			D2-2 (without Admixture variables, Slump)	E1										
			D3-1 (without Admixture variables, Slump)	E1										
			D3-2 (without Admixture variables, Slump)	E1										
A1 (without Cement Early Strength Gain Class)	B2-4 without CTE	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.82	0.64	0.45	0.64	*	0.46	0.59	0.33	0.46	None
				E2 without Age at first drying	0.71	0.63	0.43	0.59	*					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.79	0.63	0.46	0.63	*	0.41	0.54	0.33	0.43	α
				E2 without Age at first drying	0.60	0.61	0.33	0.51	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.78	0.63	0.47	0.63	*	0.43	0.56	0.31	0.43	None
				E2 without Age at first drying	0.60	0.61	0.33	0.51	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.79	0.63	0.46	0.63	*	reduces to same model as A1/B2-4/C2/D1-2/E1/b				
				E2 without Age at first drying	0.59	0.61	0.34	0.51	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.81	0.61	0.42	0.61	*					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.81	0.61	0.42	0.61	*					

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.81	0.60	0.44	0.62	*	0.65	0.55	0.42	0.54	None
				E2 without Age at first drying	0.72	0.61	0.36	0.56	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.81	0.60	0.44	0.62	*	reduces to same model as A1/B2-4/C2/D3-1/E1/b				
				E2 without Age at first drying	0.67	0.61	0.38	0.55	*					
A1 (without Cement Early Strength Gain Class)	B2-5 (without Stone LBD and Stone Flakiness Index)	C1	D1-1 (without Admixture variables)	E1	0.74	0.64	0.46	0.61	None					
				E2	0.72	0.60	0.45	0.59	None					
			D1-2 (without Admixture variables)	E1	0.73	0.61	0.46	0.60	None					
			D1-3 (without Admixture variables)	E1	0.74	0.61	0.45	0.60	None					
			D1-4 (without Admixture variables)	E1	0.74	0.60	0.45	0.60	None					
			D2-1 (without Admixture variables)	E1	0.73	0.60	0.41	0.58	None					
			D2-2 (without Admixture variables)	E1	0.73	0.62	0.43	0.59	None					
			D3-1 (without Admixture variables, Compressive Strength)	E1	0.72	0.62	0.44	0.59	*					
			D3-2 (without Admixture variables, Compressive Strength)	E1	0.70	0.62	0.44	0.59	None					
A1	B2-5 (without Stone LBD and Stone Flakiness Index)	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.74	0.66	0.51	0.64	None	0.48	0.60	0.36	0.48	None
				E2 without Age at first drying	0.64	0.64	0.46	0.58	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.73	0.66	0.53	0.64	None	0.47	0.60	0.37	0.48	None
				E2 without Age at first drying	0.68	0.64	0.47	0.60	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.73	0.66	0.52	0.64	None	0.45	0.60	0.37	0.47	None
				E2 without Age at first drying	0.66	0.64	0.42	0.57	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.72	0.66	0.52	0.63	None	0.47	0.60	0.34	0.47	None
				E2 without Age at first drying	0.68	0.64	0.40	0.57	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.74	0.65	0.48	0.62	None	0.45	0.59	0.34	0.46	None
				E2 without Age at first drying	0.64	0.63	0.39	0.55	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.75	0.64	0.48	0.62	None	reduces to same model as A1/B2-5/C2/D2-1/E1/b				
				E2 without Age at first drying	0.63	0.63	0.39	0.55	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.70	0.64	0.52	0.62	None	0.60	0.65	0.50	0.58	None
				E2 without Age at first drying	0.67	0.63	0.45	0.58	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.70	0.64	0.52	0.62	None	reduces to same model as A1/B2-5/C2/D3-1/E1/b				
				E2 without Age at first drying	0.62	0.63	0.41	0.55	None					
A1 (without Cement Early Strength Gain Class)	B2-6 (without Stone LBD, Stone UC and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2										
			D1-2 (without Admixture variables, Slump)	E1										
			D1-3 (without Admixture variables, Slump)	E1										
			D1-4 (without Admixture variables, Slump)	E1										
			D2-1 (without Admixture variables, Slump)	E1										
			D2-2 (without Admixture variables, Slump)	E1										
			D3-1 (without Admixture variables, Slump)	E1										

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D3-2 (without Admixture variables, Slump)	E1										
A1	B2-6 (without Stone LBD and Stone Flakiness Index)	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.76	0.63	0.45	0.61	None					
				E2 without Age at first drying	0.68	0.62	0.41	0.57	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.76	0.62	0.45	0.61	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.76	0.63	0.45	0.61	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.76	0.63	0.47	0.62	None	0.59	0.59	0.43	0.54	None
				E2 without Age at first drying	0.64	0.61	0.36	0.54	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.78	0.61	0.41	0.60	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.61	0.41	0.60	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.78	0.60	0.44	0.61	None					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.60	0.44	0.61	None					
A1 (without Cement Early Strength Gain Class)	B2-7 (without Stone CTE and Stone Flakiness Index)	C1	D1-1 (without Admixture variables)	E1	0.75	0.60	0.43	0.59	*					
				E2	0.74	0.56	0.42	0.57	None					
			D1-2 (without Admixture variables)	E1	0.75	0.60	0.44	0.60	None					
			D1-3 (without Admixture variables)	E1	0.75	0.60	0.44	0.60	None					
			D1-4 (without Admixture variables)	E1	0.76	0.59	0.44	0.60	None					
			D2-1 (without Admixture variables)	E1	0.74	0.60	0.40	0.58	None					
			D2-2 (without Admixture variables)	E1	0.74	0.60	0.41	0.58	None					
			D3-1 (without Admixture variables, Compressive Strength)	E1	0.74	0.60	0.42	0.59	*					
			D3-2 (without Admixture variables, Compressive Strength)	E1	0.73	0.61	0.43	0.59	None					
A1	B2-7 without CTE	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.82	0.63	0.45	0.63	*	0.50	0.62	0.36	0.49	None
				E2 without Age at first drying	0.77	0.62	0.40	0.60	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.81	0.63	0.47	0.64	None	0.53	0.61	0.32	0.49	None
				E2 without Age at first drying	0.79	0.62	0.42	0.61	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.82	0.63	0.45	0.63	None	0.49	0.62	0.37	0.49	None
				E2 without Age at first drying	0.77	0.62	0.41	0.60	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.81	0.63	0.46	0.63	None	reduces to same model as A1/B2-7/C2/D1-2/E1/b				
				E2 without Age at first drying	0.79	0.62	0.40	0.60	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.83	0.62	0.42	0.62	None	0.50	0.61	0.33	0.48	None
				E2 without Age at first drying	0.80	0.61	0.37	0.59	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.84	0.61	0.42	0.62	None	0.51	0.61	0.34	0.49	None
				E2 without Age at first drying	0.79	0.60	0.37	0.59	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.81	0.60	0.44	0.62	None	0.72	0.60	0.46	0.59	None
				E2 without Age at first drying	0.79	0.60	0.39	0.59	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.81	0.60	0.44	0.62	None	reduces to same model as A1/B2-7/C2/D3-1/E1/b				
				E2 without Age at first drying	0.78	0.60	0.39	0.59	None					
A2 (without Cement Early Strength	B1	C1	D1-1 (without Admixture variables)	E1	0.77	0.56	0.50	0.61	*					
				E2	0.78	0.53	0.47	0.59	None					
			D1-2 (without Admixture variables)	E1	0.77	0.56	0.50	0.61	*					
			D1-3 (without Admixture variables)	E1	0.76	0.56	0.50	0.61	None					
			D1-4 (without Admixture variables)	E1	0.78	0.57	0.46	0.60	None					

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	**B**	**C**	**D**	**E**	**α**	**ln(β)**	**ln(γ)**	**Average**		**α**	**ln(β)**	**ln(γ)**	**Average**	
Gain Class)			D2-1 (without Admixture variables)	E1	0.77	0.57	0.47	0.60	None					
			D2-2 (without Admixture variables)	E1	0.78	0.57	0.46	0.60	None					
			D3-1 (without Admixture variables, Compressive Strength)	E1	0.78	0.57	0.46	0.60	None					
			D3-2 (without Admixture variables, Compressive Strength)	E1	0.78	0.57	0.46	0.60	None					
A2 (without Cement Early Strength Gain Class)	B1	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.75	0.60	0.35	0.57	*					
				E2 without Age at first drying	0.75	0.55	0.35	0.55	*					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.74	0.59	0.37	0.57	*					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.75	0.60	0.35	0.57	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.75	0.60	0.35	0.57	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.74	0.61	0.34	0.56	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.74	0.61	0.35	0.57	None					
			D3-1 (without Admixture variables, Compressive Strength)	E1 without Age at first drying	0.73	0.60	0.37	0.57	None					
			D3-2 (without Admixture variables, Compressive Strength)	E1 without Age at first drying	0.72	0.61	0.35	0.56	None					
A2 (without Cement Early Strength Gain Class)	B2-1 (without Stone UC and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.76	0.56	0.50	0.61	*					
				E2	0.76	0.51	0.47	0.58	None					
			D1-2 (without Admixture variables, Slump)	E1	0.76	0.56	0.51	0.61	None					
			D1-3 (without Admixture variables, Slump)	E1	0.76	0.56	0.50	0.61	None					
			D1-4 (without Admixture variables, Slump)	E1	0.76	0.56	0.51	0.61	None					
			D2-1 (without Admixture variables, Slump)	E1	0.75	0.57	0.47	0.60	None					
			D2-2 (without Admixture variables, Slump)	E1	0.76	0.57	0.48	0.60	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.74	0.56	0.47	0.59	None					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.74	0.57	0.48	0.60	None					
A2	B2-1	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.83	0.63	0.52	0.66	*	0.54	0.57	0.37	0.49	None
				E2 without Age at first drying	0.82	0.57	0.43	0.61	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.83	0.63	0.53	0.66	None	0.54	0.56	0.37	0.49	α
				E2 without Age at first drying	0.83	0.57	0.44	0.61	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.83	0.63	0.52	0.66	None	reduces to same model as A2/B2-1/C2/D1-1/E1/b				
				E2 without Age at first drying	0.82	0.57	0.43	0.61	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.83	0.62	0.52	0.66	None	reduces to same model as A2/B2-1/C2/D1-2/E1/b				
				E2 without Age at first drying	0.82	0.57	0.43	0.61	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.82	0.64	0.47	0.64	None	0.54	0.56	0.34	0.48	None
				E2 without Age at first drying	0.82	0.57	0.37	0.59	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.83	0.65	0.47	0.65	None	reduces to same model as A2/B2-1/C2/D2-1/E1/b				
				E2 without Age at first drying	0.83	0.57	0.37	0.59	None					

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D3-1 (without Admixture variables, Compressive Strength)	E1 without Age at first drying	0.82	0.63	0.49	0.65	*	0.70	0.62	0.47	0.60	None
				E2 without Age at first drying	0.81	0.55	0.39	0.58	*					
			D3-2 (without Admixture variables, Compressive Strength)	E1 without Age at first drying	0.83	0.63	0.47	0.64	None	0.50	0.56	0.33	0.46	None
				E2 without Age at first drying	0.81	0.55	0.39	0.58	*					
A2 (without Cement Early Strength Gain Class)	B2-2 (without Stone LBD and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.74	0.53	0.46	0.58	*					
				E2	0.74	0.50	0.45	0.56	None					
			D1-2 (without Admixture variables, Slump)	E1	0.74	0.53	0.47	0.58	None					
			D1-3 (without Admixture variables, Slump)	E1	0.74	0.54	0.47	0.58	None					
			D1-4 (without Admixture variables, Slump)	E1	0.75	0.53	0.45	0.58	None					
			D2-1 (without Admixture variables, Slump)	E1	0.74	0.53	0.43	0.57	None					
			D2-2 (without Admixture variables, Slump)	E1	0.73	0.53	0.41	0.56	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.73	0.54	0.44	0.57	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.70	0.53	0.44	0.56	None					
A2	B2-2 (without Stone LBD and Stone Flakiness Index)	C2	D1-1 (without Admixture variables, Slump)	E1 without Age at first drying	0.72	0.56	0.49	0.59	None					
				E2 without Age at first drying	0.57	0.57	0.47	0.54	None					
			D1-2 (without Admixture variables, Slump)	E1 without Age at first drying	0.72	0.56	0.49	0.59	None					
			D1-3 (without Admixture variables, Slump)	E1 without Age at first drying	0.72	0.56	0.49	0.59	None					
			D1-4 (without Admixture variables, Slump)	E1 without Age at first drying	0.72	0.56	0.49	0.59	None					
			D2-1 (without Admixture variables, Slump)	E1 without Age at first drying	0.71	0.57	0.46	0.58	None					
			D2-2 (without Admixture variables, Slump)	E1 without Age at first drying	0.71	0.57	0.46	0.58	None					
			D3-1 (without Admixture variables, Slump)	E1 without Age at first drying	0.70	0.55	0.48	0.58	*					
			D3-2 (without Admixture variables, Slump)	E1 without Age at first drying	0.70	0.55	0.48	0.58	*					
A2 (without Cement Early Strength Gain Class)	B2-3 (without Stone CDB, Stone UC and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2										
			D1-2 (without Admixture variables, Slump)	E1										
			D1-3 (without Admixture variables, Slump)	E1										
			D1-4 (without Admixture variables, Slump)	E1										
			D2-1 (without Admixture variables, Slump)	E1										

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D2-2 (without Admixture variables, Slump)	E1										
			D3-1 (without Admixture variables, Slump)	E1										
			D3-2 (without Admixture variables, Slump)	E1										
A2 (without Cement Early Strength Gain Class)	B2-3 (without Stone CDB, Stone UC and Stone Flakiness Index)	C2	D1-1 (without Admixture variables, Slump)	E1 without Age at first drying	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2 without Age at first drying										
			D1-2 (without Admixture variables, Slump)	E1 without Age at first drying										
			D1-3 (without Admixture variables, Slump)	E1 without Age at first drying										
			D1-4 (without Admixture variables, Slump)	E1 without Age at first drying										
			D2-1 (without Admixture variables, Slump)	E1 without Age at first drying										
			D2-2 (without Admixture variables, Slump)	E1 without Age at first drying										
			D3-1 (without Admixture variables, Slump)	E1 without Age at first drying										
			D3-2 (without Admixture variables, Slump)	E1 without Age at first drying										
A2 (without Cement Early Strength Gain Class)	B2-4 (without Stone CTE, Stone UC and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2										
			D1-2 (without Admixture variables, Slump)	E1										
			D1-3 (without Admixture variables, Slump)	E1										
			D1-4 (without Admixture variables, Slump)	E1										
			D2-1 (without Admixture variables, Slump)	E1										
			D2-2 (without Admixture variables, Slump)	E1										
			D3-1 (without Admixture variables, Slump)	E1										
			D3-2 (without Admixture variables, Slump)	E1										
A2	B2-4 without CTE	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.77	0.57	0.46	0.60	*					
				E2 without Age at first drying	0.66	0.56	0.45	0.56	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.77	0.56	0.48	0.60	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.77	0.57	0.46	0.60	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.77	0.56	0.48	0.60	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.79	0.57	0.42	0.59	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.79	0.57	0.42	0.59	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.79	0.55	0.44	0.59	*					

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.54	0.44	0.59	None					
A2 (without Cement Early Strength Gain Class)	B2-5 (without Stone LBD and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.74	0.56	0.47	0.59	*					
				E2	0.75	0.52	0.43	0.57	None					
			D1-2 (without Admixture variables, Slump)	E1	0.74	0.55	0.47	0.59	None					
			D1-3 (without Admixture variables, Slump)	E1	0.75	0.55	0.46	0.59	None					
			D1-4 (without Admixture variables, Slump)	E1	0.75	0.56	0.44	0.58	None					
			D2-1 (without Admixture variables, Slump)	E1	0.74	0.55	0.43	0.57	None					
			D2-2 (without Admixture variables, Slump)	E1	0.73	0.56	0.44	0.58	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.74	0.55	0.44	0.58	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.74	0.55	0.45	0.58	None					
A2	B2-5 (without Stone LBD and Stone Flakiness Index)	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.71	0.59	0.51	0.60	*					
				E2 without Age at first drying	0.62	0.57	0.46	0.55	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.73	0.57	0.53	0.61	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.71	0.59	0.51	0.60	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.73	0.57	0.53	0.61	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.72	0.59	0.46	0.59	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.74	0.58	0.48	0.60	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.74	0.57	0.48	0.60	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.74	0.56	0.48	0.59	*					
A2 (without Cement Early Strength Gain Class)	B2-6 (without Stone LBD, Stone UC and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2										
			D1-2 (without Admixture variables, Slump)	E1										
			D1-3 (without Admixture variables, Slump)	E1										
			D1-4 (without Admixture variables, Slump)	E1										
			D2-1 (without Admixture variables, Slump)	E1										
			D2-2 (without Admixture variables, Slump)	E1										
			D3-1 (without Admixture variables, Slump)	E1										
			D3-2 (without Admixture variables, Slump)	E1										
A2	B2-6 (without Stone	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.77	0.57	0.46	0.60	None					
				E2 without Age at first drying	0.66	0.57	0.45	0.56	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.77	0.56	0.48	0.60	None					

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
	LBD)		D1-3 (without Admixture variables)	E1 without Age at first drying	0.77	0.57	0.46	0.60	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.77	0.56	0.48	0.60	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.79	0.57	0.42	0.59	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.79	0.57	0.42	0.59	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.79	0.55	0.44	0.59	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.54	0.44	0.59	*					
A2 (without Cement Early Strength Gain Class)	B2-7 (without Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.76	0.54	0.43	0.58	*					
				E2	0.75	0.51	0.41	0.56	None					
			D1-2 (without Admixture variables, Slump)	E1	0.76	0.54	0.45	0.58	None					
			D1-3 (without Admixture variables, Slump)	E1	0.77	0.54	0.44	0.58	None					
			D1-4 (without Admixture variables, Slump)	E1	0.77	0.53	0.43	0.58	None					
			D2-1 (without Admixture variables, Slump)	E1	0.76	0.54	0.43	0.58	None					
			D2-2 (without Admixture variables, Slump)	E1	0.76	0.54	0.43	0.58	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.76	0.54	0.42	0.57	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.75	0.55	0.43	0.58	None					
A2 (without Cement Early Strength Gain Class)	B2-7 (without Stone CTE and Stone FI)	C2	D1-1 (without Admixture variables, Slump)	E1 without Age at first drying	0.74	0.56	0.47	0.59	*					
				E2 without Age at first drying	0.75	0.52	0.43	0.57	None					
			D1-2 (without Admixture variables, Slump)	E1 without Age at first drying	0.74	0.55	0.47	0.59	None					
			D1-3 (without Admixture variables, Slump)	E1 without Age at first drying	0.75	0.55	0.46	0.59	None					
			D1-4 (without Admixture variables, Slump)	E1 without Age at first drying	0.74	0.55	0.47	0.59	None					
			D2-1 (without Admixture variables, Slump)	E1 without Age at first drying	0.74	0.55	0.43	0.57	None					
			D2-2 (without Admixture variables, Slump)	E1 without Age at first drying	0.73	0.56	0.44	0.58	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1 without Age at first drying	0.74	0.55	0.44	0.58	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1 without Age at first drying	0.74	0.55	0.45	0.58	None					
A3 (without Cement Early Strength Gain Class)	B1	C1	D1-1 (without Admixture variables, Slump)	E1	0.78	0.57	0.49	0.61	*					
				E2	0.78	0.53	0.46	0.59	None					
			D1-2 (without Admixture variables, Slump)	E1	0.78	0.56	0.49	0.61	None					
			D1-3 (without Admixture variables, Slump)	E1	0.78	0.56	0.48	0.61	None					
			D1-4 (without Admixture variables, Slump)	E1	0.77	0.57	0.47	0.60	None					

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D2-1 (without Admixture variables, Slump)	E1	0.78	0.56	0.47	0.60	None					
			D2-2 (without Admixture variables, Slump)	E1	0.77	0.57	0.47	0.60	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.77	0.58	0.46	0.60	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.78	0.58	0.46	0.61	None					
A3 (without Cement Early Strength Gain Class)	B1	C2	D1-1 (without Admixture variables, Slump)	E1 without Age at first drying	0.74	0.58	0.38	0.57	*					
				E2 without Age at first drying	0.72	0.55	0.38	0.55	None					
			D1-2 (without Admixture variables, Slump)	E1 without Age at first drying	0.74	0.60	0.39	0.58	None					
			D1-3 (without Admixture variables, Slump)	E1 without Age at first drying	0.74	0.61	0.39	0.58	None					
			D1-4 (without Admixture variables, Slump)	E1 without Age at first drying	0.74	0.60	0.39	0.58	None					
			D2-1 (without Admixture variables, Slump)	E1 without Age at first drying	0.73	0.60	0.38	0.57	None					
			D2-2 (without Admixture variables, Slump)	E1 without Age at first drying	0.74	0.61	0.38	0.58	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1 without Age at first drying	0.72	0.61	0.34	0.56	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1 without Age at first drying	0.72	0.61	0.34	0.56	None					
A3 (without Cement Early Strength Gain Class)	B2-1 (without Stone UC and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.76	0.58	0.50	0.61	*					
				E2	0.75	0.53	0.48	0.59	None					
			D1-2 (without Admixture variables, Slump)	E1	0.75	0.58	0.50	0.61	None					
			D1-3 (without Admixture variables, Slump)	E1	0.76	0.58	0.50	0.61	None					
			D1-4 (without Admixture variables, Slump)	E1	0.75	0.58	0.50	0.61	None					
			D2-1 (without Admixture variables, Slump)	E1	0.75	0.58	0.48	0.60	None					
			D2-2 (without Admixture variables, Slump)	E1	0.75	0.58	0.49	0.61	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.73	0.58	0.47	0.59	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.73	0.58	0.48	0.60	None					
A3	B2-1	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.83	0.65	0.52	0.67	None	0.53	0.58	0.37	0.49	None
				E2 without Age at first drying	0.82	0.58	0.42	0.61	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.83	0.65	0.52	0.67	None	0.54	0.57	0.37	0.49	None
				E2 without Age at first drying	0.83	0.59	0.43	0.62	None	0.47	0.55	0.37	0.46	None
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.83	0.65	0.51	0.66	None	reduces to same model as A3/B2-1/C2/D1-1/E1/b				
				E2 without Age at first drying	0.82	0.58	0.42	0.61	None					

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.83	0.64	0.52	0.66	None	0.54	0.53	0.33	0.47	α
				E2 without Age at first drying	0.82	0.58	0.42	0.61	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.84	0.66	0.52	0.67	None	0.54	0.57	0.35	0.49	None
				E2 without Age at first drying	0.83	0.58	0.41	0.61	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.84	0.65	0.51	0.67	None	reduces to same model as A3/B2-1/C2/D2-1/E1/b				
				E2 without Age at first drying	0.83	0.58	0.40	0.60	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.84	0.63	0.49	0.65	*	0.71	0.62	0.48	0.60	None
				E2 without Age at first drying	0.85	0.55	0.37	0.59	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.85	0.64	0.49	0.66	*	reduces to same model as A3/B2-1/C2/D3-1/E1/b				
				E2 without Age at first drying	0.85	0.56	0.37	0.59	*					
A3 (without Cement Early Strength Gain Class)	B2-2 (without Stone LBD and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.74	0.54	0.46	0.58	*					
				E2	0.74	0.51	0.44	0.56	None					
			D1-2 (without Admixture variables, Slump)	E1	0.73	0.54	0.46	0.58	None					
			D1-3 (without Admixture variables, Slump)	E1	0.74	0.53	0.46	0.58	None					
			D1-4 (without Admixture variables, Slump)	E1	0.73	0.53	0.45	0.57	None					
			D2-1 (without Admixture variables, Slump)	E1	0.72	0.54	0.45	0.57	None					
			D2-2 (without Admixture variables, Slump)	E1	0.72	0.53	0.44	0.56	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.72	0.55	0.44	0.57	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.69	0.54	0.44	0.56	None					
A3	B2-2 (without Stone LBD)	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.81	0.58	0.47	0.62	*	0.48	0.56	0.37	0.47	None
				E2 without Age at first drying	0.80	0.56	0.41	0.59	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.82	0.57	0.48	0.62	None	0.49	0.57	0.39	0.48	None
				E2 without Age at first drying	0.80	0.56	0.43	0.60	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.81	0.58	0.46	0.62	None	reduces to same model as A3/B2-2/C2/D1-1/E1/b				
				E2 without Age at first drying	0.80	0.56	0.41	0.59	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.83	0.56	0.47	0.62	None	0.52	0.56	0.38	0.49	None
				E2 without Age at first drying	0.80	0.55	0.42	0.59	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.83	0.57	0.46	0.62	None	0.51	0.50	0.34	0.45	None
				E2 without Age at first drying	0.81	0.55	0.40	0.59	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.83	0.56	0.45	0.61	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.83	0.55	0.42	0.60	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.82	0.54	0.44	0.60	*					
A3 (without Cement Early Strength Gain	B2-3 (without Stone CDB, Stone UC and Stone	C1	D1-1 (without Admixture variables, Slump)	E1	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2										
			D1-2 (without Admixture variables, Slump)	E1										
			D1-3 (without Admixture variables, Slump)	E1										

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
Class)	Flakiness Index)		D1-4 (without Admixture variables, Slump)	E1										
			D2-1 (without Admixture variables, Slump)	E1										
			D2-2 (without Admixture variables, Slump)	E1										
			D3-1 (without Admixture variables, Slump)	E1										
			D3-2 (without Admixture variables, Slump)	E1										
A3 (without Cement Early Strength Gain Class)	B2-3 (without Stone CDB, Stone UC and Stone Flakiness Index)	C2	D1-1 (without Admixture variables, Slump)	E1 without Age at first drying	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2 without Age at first drying										
			D1-2 (without Admixture variables, Slump)	E1 without Age at first drying										
			D1-3 (without Admixture variables, Slump)	E1 without Age at first drying										
			D1-4 (without Admixture variables, Slump)	E1 without Age at first drying										
			D2-1 (without Admixture variables, Slump)	E1 without Age at first drying										
			D2-2 (without Admixture variables, Slump)	E1 without Age at first drying										
			D3-1 (without Admixture variables, Slump)	E1 without Age at first drying										
			D3-2 (without Admixture variables, Slump)	E1 without Age at first drying										
A3 (without Cement Early Strength Gain Class)	B2-4 (without Stone CTE and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.75	0.54	0.44	0.58	*					
				E2	0.76	0.50	0.42	0.56	None					
			D1-2 (without Admixture variables, Slump)	E1	0.74	0.55	0.45	0.58	None					
			D1-3 (without Admixture variables, Slump)	E1	0.75	0.55	0.44	0.58	None					
			D1-4 (without Admixture variables, Slump)	E1	0.75	0.55	0.44	0.58	None					
			D2-1 (without Admixture variables, Slump)	E1	0.75	0.54	0.42	0.57	None					
			D2-2 (without Admixture variables, Slump)	E1	0.74	0.54	0.43	0.57	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.73	0.54	0.42	0.56	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.72	0.55	0.43	0.57	None					
A3	B2-4 without CTE	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.78	0.58	0.45	0.60	*					
				E2 without Age at first drying	0.68	0.57	0.45	0.57	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.78	0.58	0.45	0.60	None					

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.81	0.57	0.44	0.61	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.80	0.57	0.44	0.60	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.79	0.55	0.45	0.60	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.79	0.55	0.45	0.60	None					
A3 (without Cement Early Strength Gain Class)	B2-5 (without Stone LBD and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	0.74	0.57	0.46	0.59	None					
				E2	0.74	0.54	0.44	0.57	None					
			D1-2 (without Admixture variables, Slump)	E1	0.74	0.57	0.46	0.59	None					
			D1-3 (without Admixture variables, Slump)	E1	0.75	0.56	0.46	0.59	None					
			D1-4 (without Admixture variables, Slump)	E1	0.75	0.56	0.45	0.59	None					
			D2-1 (without Admixture variables, Slump)	E1	0.74	0.55	0.43	0.57	None					
			D2-2 (without Admixture variables, Slump)	E1	0.74	0.56	0.45	0.58	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.73	0.56	0.45	0.58	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.73	0.57	0.46	0.59	None					
A3	B2-5 (without Stone LBD)	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.82	0.57	0.45	0.61	None					
				E2 without Age at first drying	0.77	0.57	0.43	0.59	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.81	0.56	0.48	0.62	None	0.53	0.56	0.33	0.47	None
				E2 without Age at first drying	0.77	0.56	0.46	0.60	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.82	0.57	0.46	0.62	None	0.52	0.57	0.36	0.48	None
				E2 without Age at first drying	0.76	0.57	0.43	0.59	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.82	0.56	0.47	0.62	None	0.54	0.56	0.33	0.48	None
				E2 without Age at first drying	0.78	0.55	0.43	0.59	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.83	0.56	0.44	0.61	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.83	0.56	0.44	0.61	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.82	0.54	0.44	0.60	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.81	0.54	0.44	0.60	*					
A3 (without Cement Early Strength Gain Class)	B2-6 (without Stone LBD, Stone UC and Stone Flakiness Index)	C1	D1-1 (without Admixture variables, Slump)	E1	After elimination of additional covariates to remove problems with no variance in the design matrix, this combination of covariates comprised a subset of those of B2-1 and B2-2 and so was not tested separately.									
				E2										
			D1-2 (without Admixture variables, Slump)	E1										
			D1-3 (without Admixture variables, Slump)	E1										
			D1-4 (without Admixture variables, Slump)	E1										
			D2-1 (without Admixture variables, Slump)	E1										
			D2-2 (without Admixture variables, Slump)	E1										

Table I.1 Results of the multivariate multiple regression analyses

Covariate set (refer to Figure 6.3)					Adjusted R^2				Significant lack of fit?	Adjusted R^2				Significant lack of fit?
A	B	C	D	E	α	ln(β)	ln(γ)	Average		α	ln(β)	ln(γ)	Average	
			D3-1 (without Admixture variables, Slump)	E1										
			D3-2 (without Admixture variables, Slump)	E1										
A3	B2-6 (without Stone LBD)	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.78	0.58	0.45	0.60	*					
				E2 without Age at first drying	0.68	0.57	0.45	0.57	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.78	0.58	0.45	0.60	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.78	0.57	0.47	0.61	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.81	0.57	0.44	0.61	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.80	0.57	0.44	0.60	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.79	0.55	0.45	0.60	*					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.79	0.55	0.45	0.60	*					
A3 (without Cement Early Strength Gain Class)	B2-7 (without Stone CTE and Stone FI)	C1	D1-1 (without Admixture variables, Slump)	E1	0.74	0.57	0.46	0.59	*					
				E2	0.74	0.54	0.44	0.57	None					
			D1-2 (without Admixture variables, Slump)	E1	0.74	0.57	0.46	0.59	None					
			D1-3 (without Admixture variables, Slump)	E1	0.75	0.56	0.46	0.59	None					
			D1-4 (without Admixture variables, Slump)	E1	0.75	0.56	0.45	0.59	None					
			D2-1 (without Admixture variables, Slump)	E1	0.74	0.55	0.43	0.57	None					
			D2-2 (without Admixture variables, Slump)	E1	0.74	0.56	0.45	0.58	None					
			D3-1 (without Admixture variables, Slump, Compressive Strength)	E1	0.73	0.56	0.45	0.58	*					
			D3-2 (without Admixture variables, Slump, Compressive Strength)	E1	0.73	0.57	0.46	0.59	None					
A3	B2-7 without CTE	C2	D1-1 (without Admixture variables)	E1 without Age at first drying	0.82	0.57	0.45	0.61	*					
				E2 without Age at first drying	0.77	0.57	0.42	0.59	None					
			D1-2 (without Admixture variables)	E1 without Age at first drying	0.81	0.56	0.48	0.62	None	0.53	0.56	0.33	0.47	None
				E2 without Age at first drying	0.77	0.56	0.46	0.60	None					
			D1-3 (without Admixture variables)	E1 without Age at first drying	0.82	0.57	0.46	0.62	None	0.52	0.57	0.36	0.48	None
				E2 without Age at first drying	0.76	0.57	0.43	0.59	None					
			D1-4 (without Admixture variables)	E1 without Age at first drying	0.82	0.56	0.47	0.62	None	0.54	0.56	0.33	0.48	None
				E2 without Age at first drying	0.78	0.55	0.43	0.59	None					
			D2-1 (without Admixture variables)	E1 without Age at first drying	0.83	0.56	0.44	0.61	None					
			D2-2 (without Admixture variables)	E1 without Age at first drying	0.83	0.57	0.44	0.61	None					
			D3-1 (without Admixture variables)	E1 without Age at first drying	0.83	0.56	0.44	0.61	None					
			D3-2 (without Admixture variables)	E1 without Age at first drying	0.82	0.54	0.44	0.60	*					

Appendix J

Results of the diagnostic checks for the unweighted multivariate multiple regressions

The results of the following diagnostic checks for the unweighted multivariate multiple regressions are presented in Table J.1:

- Examination of the regression variate: Plots of studentised residuals versus predicted dependent variables, observed dependent variables and all covariates (those in the model and those not in the model).
- A check of the assumption of homoscedasticity: The plot of residuals versus the predicted dependent variable.
- Checks of the assumption of the normality of the error term distribution: Distribution and normal probability plot of residuals.
- Examination for outliers**:** Distribution of studentised residuals
- Influential and leverage points: Leverage, Mahalanobis D^2, studentised deleted residuals, Cook's distance and SDFFIT statistics.
- Multicollinearity: Variance Inflation Factor

Problematic results of diagnostic checks are indicated in red.

For each model, the diagnostic checks were repeated upon the removal of outliers and / or influential points. Model terms which subsequently were no longer significant are indicated at the end of the table.

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions

Diagnostic	Model 2*	Model 2* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 2* after deletion of 3 outliers (cases 169, 165 and 170) and leverage case (case 70)**	Model 3*	Model 3* after deletion of 3 outliers (cases 169, 165 and 170)
Average adjusted R^2 for the three dependent variables	0.621	0.666	0.660	0.623	0.670
Studentised residuals vs. predicted dependent variables	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK
Studentised residuals vs. observed dependent variables	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK
Studentised residuals vs. covariates in model	OK	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK	OK
Normal probability plot of residuals	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.	OK
Distribution of residuals	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.	OK
Observed vs. predicted dependent variables	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK
Leverage for p > 10 and n > 50 : critical values are > 2p/n p = number of coefficients + 1 / n = sample size	2p/n=0.259 15 cases exceed this value.	2p/n=0.263 14 cases exceed this value.	2p/n=0.265 13 cases exceed this value.	2p/n=0.259 16 cases exceed this value.	2p/n=0.263 15 cases exceed this value.
Mahalanobis D2	Case # 70 very large compared to others.	Case # 70 very large compared to others.	OK	Case # 70 very large compared to others.	Case # 70 very large compared to others.
Studentised deleted residuals (for n>50, significant if absolute value > 1.96 (0.95 confidence level))	7, 7 and 8 significant cases for α, ln(β) and ln(γ) respectively	8, 8 and 10 significant cases for α, ln(β) and ln(γ) respectively	8, 8 and 10 significant cases for α, ln(β) and ln(γ) respectively	3, 7 and 10 significant cases for α, ln(β) and ln(γ) respectively	5, 7 and 9 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance (significant if > 1 for medium to large n)	No significant cases. Statistic not calculated for case # 70.	No significant cases. Statistic not calculated for case # 70.	No significant cases.	No significant cases. Statistic not calculated for case # 70.	Case # 70 is significant for α and ln(β)
SDFFIT (significant if absolute value > 1.96)	No significant cases. Statistic not calculated for case # 70.	No significant cases. Statistic not calculated for case # 70.	Case # 136 is significant for ln(β)	No significant cases. Statistic not calculated for case # 70.	Case # 70 is significant for α and ln(β)
Multicollinearity	not checked	not checked	ln(VTSA) (VIF = 10.6)	not checked	not checked
Terms no longer significant	n/a	-	Sand River Windhoek	n/a	-

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions (continued)

Diagnostic	**Model 3* after deletion of 3 outliers (cases 169, 165 and 170) and leverage case (case 70)**	Model 4*	Model 4* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 4* after deletion of 3 outliers (cases 169, 165 and 170) and leverage case (case 70)**	Model 5*
Average adjusted R^2 for the three dependent variables	0.664	0.624	0.672	0.667	0.621
Studentised residuals vs. predicted dependent variables	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.
Studentised residuals vs. observed dependent variables	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.
Studentised residuals vs. covariates in model	OK	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK	OK
Normal probability plot of residuals	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.
Distribution of residuals	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.
Observed vs. predicted dependent variables	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK
Leverage	2p/n=0.268 13 cases exceed this value.	2p/n=0.247 15 cases exceed this value.	2p/n=0.251 14 cases exceed this value.	2p/n=0.253 13 cases exceed this value.	2p/n=0.247 14 cases exceed this value.
Mahalanobis D2	OK	Case # 70 very large compared to others.	Case # 70 very large compared to others.	OK	Case # 70 very large compared to others.
Studentised deleted residuals	5, 7 and 9 significant cases for α, ln(β) and ln(γ) respectively	5, 6 and 8 significant cases for α, ln(β) and ln(γ) respectively	9, 8 and 10 significant cases for α, ln(β) and ln(γ) respectively	9, 8 and 10 significant cases for α, ln(β) and ln(γ) respectively	7, 7 and 8 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	No significant cases.	No significant cases. Statistic not calculated for case # 70.	No significant cases. Statistic not calculated for case # 70.	No significant cases.	No significant cases. Statistic not calculated for case # 70.
SDFFIT	Case # 136 is significant for ln(β)	No significant cases. Statistic not calculated for case # 70.	No significant cases. Statistic not calculated for case # 70.	No significant cases.	No significant cases. Statistic not calculated for case # 70.
Multicollinearity	ln(VTSA) (VIF = 10.6)	not checked	not checked	ln(VTSA) (VIF = 10.7)	not checked
Terms no longer significant	Sand River Windhoek	n/a	-	Sand River Windhoek	n/a

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions (continued)

Diagnostic	Model 5* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 5* after deletion of 3 outliers (cases 169, 165 and 170) and leverage case (case 70)**	Model 6*	Model 6* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 6* after deletion of 3 outliers (cases 169, 165 and 170) and leverage case (case 70)**
Average adjusted R^2 for the three dependent variables	0.667	0.661	0.621	0.667	0.662
Studentised residuals vs. predicted dependent variables	OK	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK
Studentised residuals vs. observed dependent variables	OK	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK
Studentised residuals vs. covariates in model	OK	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK	OK
Normal probability plot of residuals	OK	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK
Distribution of residuals	OK	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK
Observed vs. predicted dependent variables	OK	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK
Leverage	2p/n=0.251 14 cases exceed this value.	2p/n=0.253 13 cases exceed this value.	2p/n=0.247 14 cases exceed this value.	2p/n=0.251 14 cases exceed this value.	2p/n=0.253 13 cases exceed this value.
Mahalanobis D2	Case # 70 very large compared to others.	OK	Case # 70 very large compared to others.	Case # 70 very large compared to others.	OK
Studentised deleted residuals	8, 9 and 11 significant cases for α, ln(β) and ln(γ) respectively	8, 9 and 11 significant cases for α, ln(β) and ln(γ) respectively	7, 7 and 9 significant cases for α, ln(β) and ln(γ) respectively	8, 8 and 12 significant cases for α, ln(β) and ln(γ) respectively	8, 8 and 12 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	No significant cases. Statistic not calculated for case # 70.	No significant cases.	Case # 70 is significant (for all three dependent variables)	No significant cases. Statistic not calculated for case # 70.	No significant cases.
SDFFIT	No significant cases. Statistic not calculated for case # 70.	No significant cases.	Case # 70 is significant (for all three dependent variables)	No significant cases. Statistic not calculated for case # 70.	No significant cases.
Multicollinearity	not checked	ln(VTSA) (VIF = 10.6)	not checked	not checked	ln(VTSA) (VIF = 10.7)
Terms no longer significant	-	Sand River Windhoek	n/a	-	Sand River Windhoek

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions (continued)

Diagnostic	Model 7*	Model 7* after deletion of 3 outliers (cases 169, 165 and 170)	Model 7* after deletion of 3 outliers (cases 169, 165 and 170) and leverage case (case 70)	**Model 7* after deletion of 3 outliers (cases 169, 165 and 170), leverage case (case 70) and sig SDFFITs (cases 79 and 136)**
Average adjusted R^2 for the three dependent variables	0.632	0.696	0.690	0.695
Studentised residuals vs. predicted dependent variables	α: 1 residual > 5 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	OK
Studentised residuals vs. observed dependent variables	α: 1 residual > 5 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	OK
Studentised residuals vs. covariates in model	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK
Normal probability plot of residuals	Outliers evident as indicated above. Otherwise fine.	OK	OK	OK
Distribution of residuals	Outliers evident as indicated above. Otherwise fine.	OK	OK	OK
Observed vs. predicted dependent variables	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK	OK
Leverage	2p/n=0.329 18 cases exceed this value.	2p/n=0.335 17 cases exceed this value.	2p/n=0.337 16 cases exceed this value.	2p/n=0.341 13 cases exceed this value.
Mahalanobis D2	Case # 70 very large compared to others.	Case # 70 very large compared to others.	Case # 79 large compared to others.	OK
Studentised deleted residuals	6, 7 and 11 significant cases for α, ln(β) and ln(γ) respectively	12, 6 and 10 significant cases for α, ln(β) and ln(γ) respectively	12, 6 and 10 significant cases for α, ln(β) and ln(γ) respectively	12, 5 and 10 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	Case # 70 is significant (for all three dependent variables)	No significant cases. Statistic not calculated for case # 70.	No significant cases.	No significant cases.
SDFFIT	Case # 70 is significant (for all three dependent variables), # 79 for α, # 169 for α, # 165 for ln(γ).	Statistic not calculated for case # 70. Case # 79 is significant for α, # 136 for ln(β).	Case # 79 is significant for α, # 136 for ln(β).	Case # 160 is significant for ln(γ).
Multicollinearity	not checked	not checked	not checked	Some Stone Type terms, some Sand Type terms, Lab, AgeDry, Temp and ln(VTSA) have VIF > 10
Terms no longer significant	n/a	-	Stone Quartz Alluvial	CEM II B-S GGVS

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions (continued)

Diagnostic	Model 8*	Model 8* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 8* after deletion of 3 outliers (cases 169, 165 and 170), leverage case (case 70) and sig SDFFITs (cases 79 and 136)**	Model 9*
Average adjusted R^2 for the three dependent variables	0.629	0.689	0.688	0.629
Studentised residuals vs. predicted dependent variables	α: 1 residual > 5 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	α: 1 residual > 5 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.
Studentised residuals vs. observed dependent variables	α: 1 residual > 5 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	α: 1 residual > 5 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.
Studentised residuals vs. covariates in model	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK
Normal probability plot of residuals	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.
Distribution of residuals	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.
Observed vs. predicted dependent variables	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK
Leverage	2p/n=0.329 18 cases exceed this value.	2p/n=0.335 17 cases exceed this value.	2p/n=0.341 13 cases exceed this value.	2p/n=0.318 15 cases exceed this value.
Mahalanobis D2	Case # 70 very large compared to others.	Case # 70 very large compared to others.	OK	Case # 70 very large compared to others.
Studentised deleted residuals	5, 7 and 10 significant cases for α, ln(β) and ln(γ) respectively	13, 6 and 11 significant cases for α, ln(β) and ln(γ) respectively	14, 4 and 11 significant cases for α, ln(β) and ln(γ) respectively	8, 6 and 10 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	No significant cases. Statistic not calculated for case # 70.	No significant cases. Statistic not calculated for case # 70.	No significant cases.	No significant cases. Statistic not calculated for case # 70.
SDFFIT	Statistic not calculated for case # 70. Cases # 79, 169 significant for α, # 165 for ln(γ).	Statistic not calculated for case # 70. Case # 79 is significant for α, # 136 for ln(β).	Case # 160, 54 is significant for ln(γ).	Statistic not calculated for case # 70. Cases # 169 significant for α, # 165 for ln(γ).
Multicollinearity	not checked	not checked	Some Stone Type terms, some Sand Type terms, Lab, AgeDry, Temp and ln(VTSA) have VIF > 10	not checked
Terms no longer significant	n/a	Stone/Sand Ratio	CEM II BS GGVS Stone Quartz Alluvial	n/a

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions (continued)

Diagnostic	Model 9* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 9* after deletion of 3 outliers (cases 169, 165 and 170), leverage case (case 70) and sig SDFFITs (cases 79)**	Model 10*	Model 10* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 10* after deletion of 3 outliers (cases 169, 165 and 170) and leverage case (case 70)**
Average adjusted R^2 for the three dependent variables	0.691	0.687	0.644	0.707	0.703
Studentised residuals vs. predicted dependent variables	OK	OK	α: 1 residual > 5 ln(β): 1 residual > 4 ln(γ): 1 residual > 5 Shape of plot fine.	OK	OK
Studentised residuals vs. observed dependent variables	OK	OK	α: 1 residual > 5 ln(β): 1 residual > 4 ln(γ): 1 residual > 5 Shape of plot fine.	OK	OK
Studentised residuals vs. covariates in model	OK	OK	OK	OK	
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	
Residuals vs. time (Test Year)	OK	OK	OK	OK	
Normal probability plot of residuals	OK	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK
Distribution of residuals	OK	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK
Observed vs. predicted dependent variables	OK	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK
Leverage	2p/n=0.323 14 cases exceed this value.	2p/n=0.325 14 cases exceed this value.	2p/n=0.318 19 cases exceed this value.	2p/n=0.323 18 cases exceed this value.	2p/n=0.325 17 cases exceed this value.
Mahalanobis D2	Case # 70 very large compared to others.	OK	Case # 70 very large compared to others.	Case # 70 very large compared to others.	OK
Studentised deleted residuals	9, 8 and 13 significant cases for α, ln(β) and ln(γ) respectively	8, 7 and 12 significant cases for α, ln(β) and ln(γ) respectively	7, 7 and 9 significant cases for α, ln(β) and ln(γ) respectively	11, 9 and 10 significant cases for α, ln(β) and ln(γ) respectively	11, 9 and 10 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	Case # 70 is significant (for all three dependent variables)	No significant cases.	No significant cases.	No significant cases. Statistic not calculated for case # 70.	No significant cases.
SDFFIT	Case # 70 is significant (for all three dependent variables) and # 79 for α	No significant cases.	Cases # 169 significant for α, # 165 for ln(γ).	No significant cases. Statistic not calculated for case # 70.	No significant cases.
Multicollinearity	not checked	Some Stone Type terms, some Sand Type terms, Lab, Temp and ln(VTSA) have VIF > 10	not checked	not checked	Some Stone Type terms, some Sand Type terms, Lab and ln(VTSA) have VIF > 10
Terms no longer significant	-	Stone Quartz Alluvial	n/a	-	Stone Quartz Alluvial

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions (continued)

Diagnostic	**Model 10* after deletion of 3 outliers (cases 169, 165 and 170) and leverage case (case 70)**	Model 11*	Model 11* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 11* after deletion of 3 outliers (cases 169, 165 and 170) and leverage cases (case 70 and 79)**	Model 12*
Average adjusted R^2 for the three dependent variables	0.703	0.621	0.682	0.677	0.617
Studentised residuals vs. predicted dependent variables	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 5 Shape of plot fine.	OK	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.
Studentised residuals vs. observed dependent variables	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 5 Shape of plot fine.	OK	OK	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.
Studentised residuals vs. covariates in model	OK	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK	OK
Normal probability plot of residuals	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.
Distribution of residuals	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.
Observed vs. predicted dependent variables	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK
Leverage	2p/n=0.325 17 cases exceed this value.	2p/n=0.306 19 cases exceed this value.	2p/n=0.311 19 cases exceed this value.	2p/n=0.313 17 cases exceed this value.	2p/n=0.235 14 cases exceed this value.
Mahalanobis D2	OK	Cases # 70 and 79 very large compared to others.	Cases # 70 and 79 very large compared to others.	OK	Case # 70 very large compared to others.
Studentised deleted residuals	11, 9 and 10 significant cases for α, ln(β) and ln(γ) respectively	5, 8 and 11 significant cases for α, ln(β) and ln(γ) respectively	11, 7 and 11 significant cases for α, ln(β) and ln(γ) respectively	11, 7 and 11 significant cases for α, ln(β) and ln(γ) respectively	5, 5 and 11 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	No significant cases.	No significant cases. Statistic not calculated for cases # 70 and 79.	No significant cases. Statistic not calculated for cases # 70 and 79.	No significant cases.	No significant cases. Statistic not calculated for case # 70.
SDFFIT	No significant cases.	Statistic not calculated for cases # 70 and 79. Cases # 169 significant for α, # 165 for ln(γ).	No significant cases. Statistic not calculated for cases # 70 and 79.	No significant cases.	No significant cases. Statistic not calculated for case # 70.
Multicollinearity	Some Stone Type terms, some Sand Type terms, Lab and ln(VTSA) have VIF > 10	not checked	not checked	Some Stone Type terms, some Sand Type terms, Lab have VIF > 10	not checked
Terms no longer significant	Stone Quartz Alluvial	n/a	-	Stone Quartz Alluvial Sand Ecca Grit	n/a

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions (continued)

Diagnostic	Model 12* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 12* after deletion of 3 outliers (cases 169, 165 and 170) and leverage case (case 70)**	Model 13*	Model 13* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 13* after deletion of 3 outliers (cases 169, 165 and 170), leverage case (case 70) and sig SDFFITs (cases 79, 136, 139)**
Average adjusted R^2 for the three dependent variables	0.664	0.658	0.617	0.687	0.701
Studentised residuals vs. predicted dependent variables	OK	OK	α: 1 residual > 5 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK
Studentised residuals vs. observed dependent variables	OK	OK	α: 1 residual > 5 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK
Studentised residuals vs. covariates in model	OK	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK	OK
Normal probability plot of residuals	OK	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK
Distribution of residuals	OK	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK
Observed vs. predicted dependent variables	OK	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK
Leverage	2p/n=0.240 14 cases exceed this value.	2p/n=0.241 13 cases exceed this value.	2p/n=0.435 14 cases exceed this value.	2p/n=0.443 12 cases exceed this value.	2p/n=0.446 16 cases exceed this value.
Mahalanobis D2	Case # 70 very large compared to others.	OK	Case # 70 very large compared to others.	Case # 70 very large compared to others.	OK
Studentised deleted residuals	7, 8 and 12 significant cases for α, ln(β) and ln(γ) respectively	7, 8 and 12 significant cases for α, ln(β) and ln(γ) respectively	8, 6 and 9 significant cases for α, ln(β) and ln(γ) respectively	10, 9 and 10 significant cases for α, ln(β) and ln(γ) respectively	9, 10 and 10 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	No significant cases. Statistic not calculated for case # 70.	No significant cases.	No significant cases. Statistic not calculated for case # 70.	Case # 70 is significant (for all three dependent variables)	No significant cases.
SDFFIT	No significant cases. Statistic not calculated for case # 70.	No significant cases.	Statistic not calculated for case # 70. Cases # 79, 169 significant for α, # 139 for ln(β), # 165 for ln(γ).	Cases # 70 and 79 are significant (for all three dependent variables) and # 136 and 139 for ln(β).	Cases # 54 significant for ln(γ).
Multicollinearity	not checked	ln(VTSA) (VIF = 11.4)	not checked	not checked	Some Cem Type terms, some Stone Type terms, some Sand Type terms, lnCC, PVF, Lab, Temp and ln(VTSA) have VIF > 10
Terms no longer significant	-	Stone Quartz Alluvial	n/a	-	Stone Quartz Alluvial

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions (continued)

Diagnostic	Model 14*	Model 14* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 14* after deletion of 3 outliers (cases 169, 165 and 170), leverage case (case 70) and sig SDFFITs (cases 79, 139)**	Model 15*
Average adjusted R2 for the three dependent variables	0.617	0.678	0.687	0.622
Studentised residuals vs. predicted dependent variables	α: 1 residual > 4 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	α: 1 residual > 5 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.
Studentised residuals vs. observed dependent variables	α: 1 residual > 4 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	α: 1 residual > 5 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.
Studentised residuals vs. covariates in model	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK
Normal probability plot of residuals	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.
Distribution of residuals	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.
Observed vs. predicted dependent variables	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK
Leverage	2p/n=0.329 21 cases exceed this value.	2p/n=0.335 18 cases exceed this value.	2p/n=0.337 16 cases exceed this value.	2p/n=0.341 20 cases exceed this value.
Mahalanobis D2	Case # 70 very large compared to others.	Case # 70 very large compared to others.	OK	Case # 70 very large compared to others.
Studentised deleted residuals	6, 10 and 8 significant cases for α, ln(β) and ln(γ) respectively	9, 11 and 10 significant cases for α, ln(β) and ln(γ) respectively	9, 12 and 10 significant cases for α, ln(β) and ln(γ) respectively	8, 7 and 8 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	Case # 70 is significant (for ln(β) and ln(γ))	No significant cases. Statistic not calculated for case # 70.	No significant cases.	No significant cases. Statistic not calculated for case # 70.
SDFFIT	Cases # 70 and # 139 significant for ln(β), # 70, 79, 165 for ln(γ).	Statistic not calculated for case # 70. Case # 79 significant for α, # 139 for ln(β), # 79 for ln(γ).	No significant cases.	Statistic not calculated for case # 70. Cases # 79, 169 significant for α, # 139 for ln(β), # 79 for ln(γ).
Multicollinearity	not checked	not checked	Some Stone Type terms, some Sand Type terms, Lab, Temp and ln(VTSA) have VIF > 10	not checked
Terms no longer significant	n/a	-	Stone Quartz Alluvial	n/a

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions (continued)

Diagnostic	**Model 15* after deletion of 3 outliers (cases 169, 165 and 170), leverage case (case 70) and sig SDFFITs (cases 79, 139)**	Model 16*	Model 16* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 16* after deletion of 3 outliers (cases 169, 165 and 170), leverage case (case 70) and sig SDFFITs (cases 79, 139)**	Model 17*
Average adjusted R2 for the three dependent variables	0.694	0.619	0.679	0.689	0.62
Studentised residuals vs. predicted dependent variables	OK	α: 1 residual > 4 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	α: 1 residual > 4 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.
Studentised residuals vs. observed dependent variables	OK	α: 1 residual > 4 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	α: 1 residual > 4 ln(β): 1 residual > 4 ln(γ): 1 residual > 4 Shape of plot fine.
Studentised residuals vs. covariates in model	OK	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK	OK
Normal probability plot of residuals	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.
Distribution of residuals	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK	Outliers evident as indicated above. Otherwise fine.
Observed vs. predicted dependent variables	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK
Leverage	2p/n=0.349 16 cases exceed this value.	2p/n=0.318 18 cases exceed this value.	2p/n=0.323 18 cases exceed this value.	2p/n=0.325 15 cases exceed this value.	2p/n=0.247 18 cases exceed this value.
Mahalanobis D2	OK	Case # 70 very large compared to others.	Case # 70 very large compared to others.	OK	Case # 70 very large compared to others.
Studentised deleted residuals	7, 8 and 8 significant cases for α, ln(β) and ln(γ) respectively	7, 7 and 9 significant cases for α, ln(β) and ln(γ) respectively	7, 10 and 9 significant cases for α, ln(β) and ln(γ) respectively	7, 10 and 12 significant cases for α, ln(β) and ln(γ) respectively	7, 7 and 9 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	No significant cases.	Case # 70 is significant (for all three dependent variables)	No significant cases. Statistic not calculated for case # 70.	No significant cases.	No significant cases. Statistic not calculated for case # 70.
SDFFIT	Cases # 54 significant for ln(γ).	Case # 70 significant (for all three dependent variables) and # 79 for α and ln(γ) and # 139 for ln(β).	Statistic not calculated for case # 70. Case # 79 significant for α, # 139 for ln(β), # 79 for ln(γ).	No significant cases.	No significant cases. Statistic not calculated for case # 70.
Multicollinearity	Some Stone Type terms, some Sand Type terms, Lab, AgeDry, Temp and ln(VTSA) have VIF > 10	not checked	not checked	Some Stone Type terms, some Sand Type terms, Lab, Temp and ln(VTSA) have VIF > 10	not checked
Terms no longer significant	Stone Quartz Alluvial CEM III A GGVS	n/a	-	Stone Quartz Alluvial CEM III A GGVS	n/a

Table J.1 Results of diagnostic checks for the unweighted multivariate multiple regressions (continued)

Diagnostic	Model 17* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 17* after deletion of 3 outliers (cases 169, 165 and 170), leverage case (case 70) and sig SDFFITs (cases 136)**	Model 18*	Model 18* after deletion of 3 outliers (cases 169, 165 and 170)	**Model 18* after deletion of 3 outliers (cases 169, 165 and 170), leverage case (case 70) and sig SDFFITs (cases 136)**
Average adjusted R2 for the three dependent variables	0.669	0.666	0.62	0.668	0.664
Studentised residuals vs. predicted dependent variables	OK	OK	α: 1 residual > 5 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK
Studentised residuals vs. observed dependent variables	OK	OK	α: 1 residual > 5 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK
Studentised residuals vs. covariates in model	OK	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK	OK
Normal probability plot of residuals	OK	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK
Distribution of residuals	OK	OK	Outliers evident as indicated above. Otherwise fine.	OK	OK
Observed vs. predicted dependent variables	OK	OK	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK
Leverage	2p/n=0.251 17 cases exceed this value.	2p/n=0.253 13 cases exceed this value.	2p/n=0.282 15 cases exceed this value.	2p/n=0.287 13 cases exceed this value.	2p/n=0.289 8 cases exceed this value.
Mahalanobis D2	Case # 70 very large compared to others.	OK	Case # 70 very large compared to others.	Case # 70 very large compared to others.	OK
Studentised deleted residuals	7, 10 and 11 significant cases for α, ln(β) and ln(γ) respectively	7, 9 and 12 significant cases for α, ln(β) and ln(γ) respectively	6, 8 and 9 significant cases for α, ln(β) and ln(γ) respectively	9, 8 and 12 significant cases for α, ln(β) and ln(γ) respectively	10, 8 and 12 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	No significant cases. Statistic not calculated for case # 70.	No significant cases.	No significant cases. Statistic not calculated for case # 70.	No significant cases. Statistic not calculated for case # 70.	No significant cases.
SDFFIT	Statistic not calculated for case # 70. Case # 136 significant for ln(β).	No significant cases.	No significant cases. Statistic not calculated for case # 70.	Statistic not calculated for case # 70. Case # 136 significant for ln(β).	No significant cases.
Multicollinearity	not checked	ln(VTSA) (VIF = 11.1)	not checked	not checked	ln(VTSA) (VIF = 11.6)
Terms no longer significant	-	Stone Quartz Alluvial CEM II BS GGVS	n/a	-	Stone Quartz Alluvial CEM II BS GGVS

Appendix K

Diagnostic plots for Model 17* (after removal of outliers and influential points)

For each of the three modelled parameters, α, ln(β) and ln(γ), the following diagnostic plots are presented for Model 17*:

Figure K.1 Standardised residuals vs. predicted values

Figure K.2 Predicted vs. observed values

Figure K.3 Standardised residuals vs. observed values (these plots have a slope of 1-R^2)

Figure K.4 Normal probability plot of studentised residuals

Figure K.5 Distribution of studentised residuals

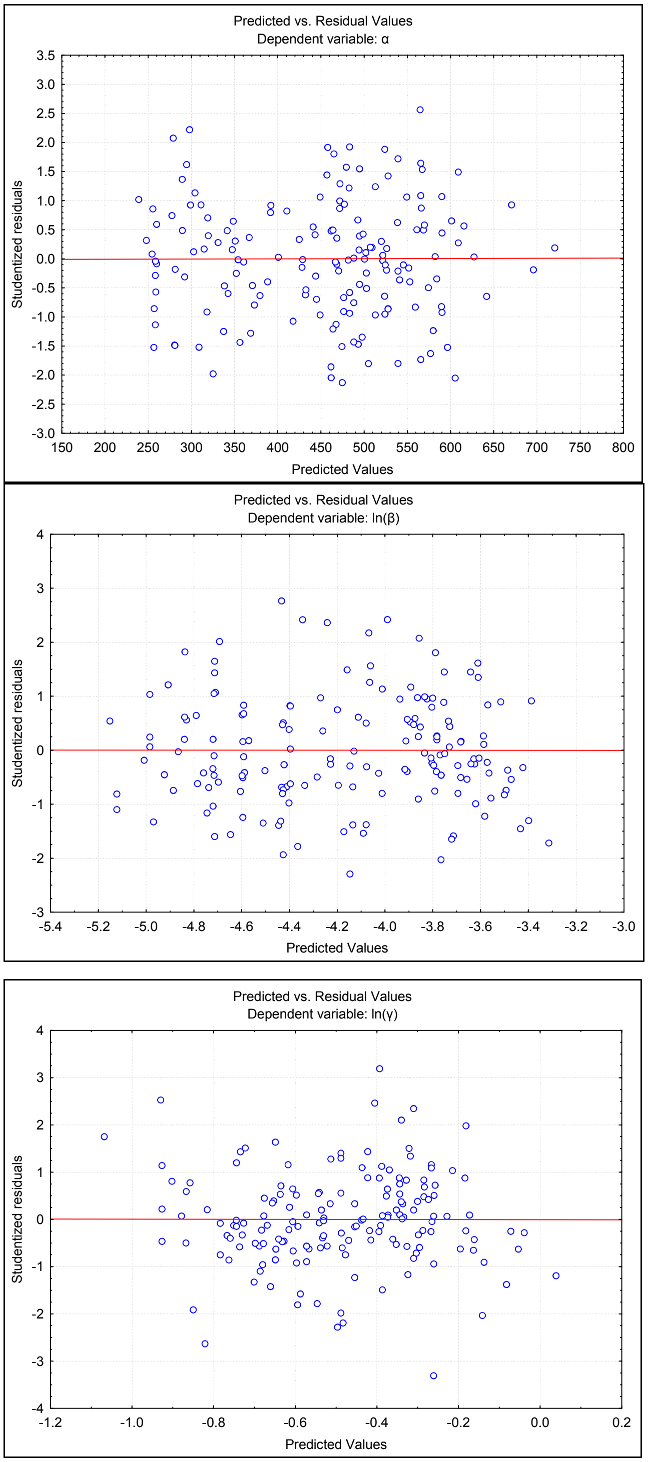


Figure K.1 Standardised residuals vs. predicted values

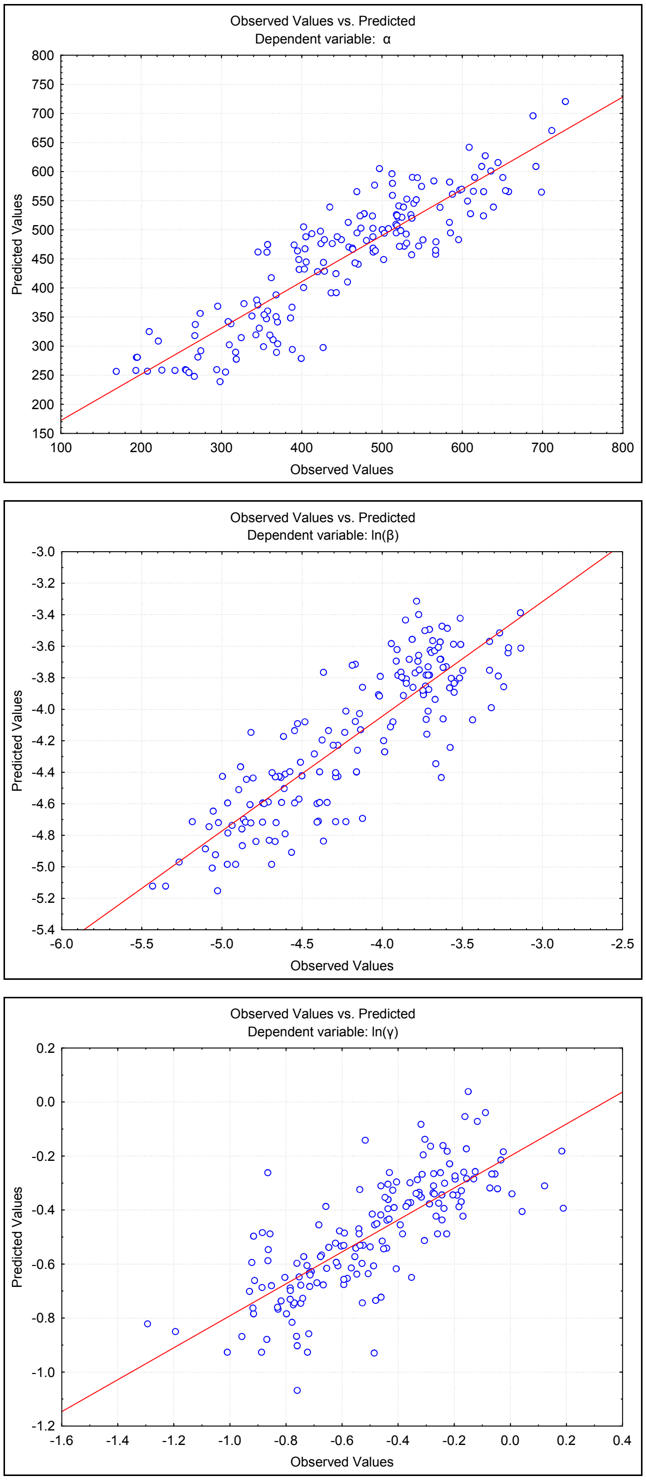


Figure K.2 Predicted vs. observed values

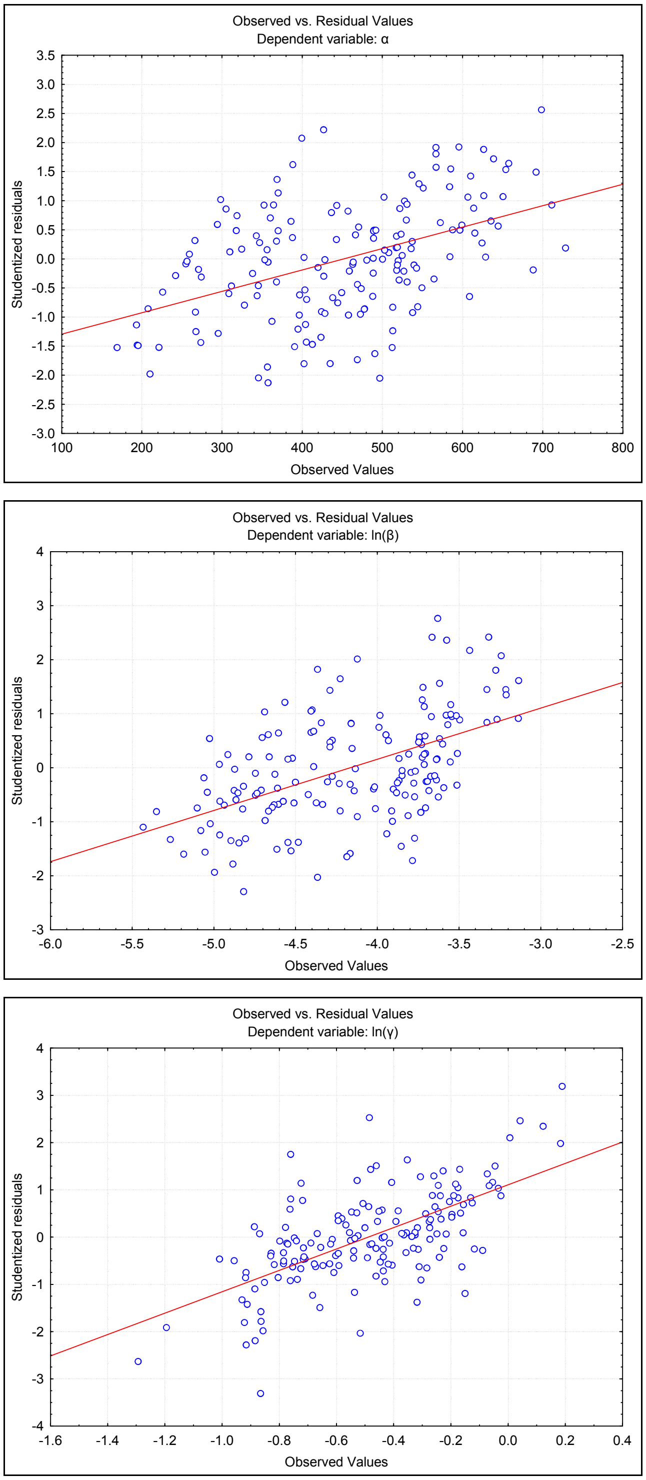


Figure K.3 Standardised residuals vs. observed values (these plots have a slope of 1-R^2)

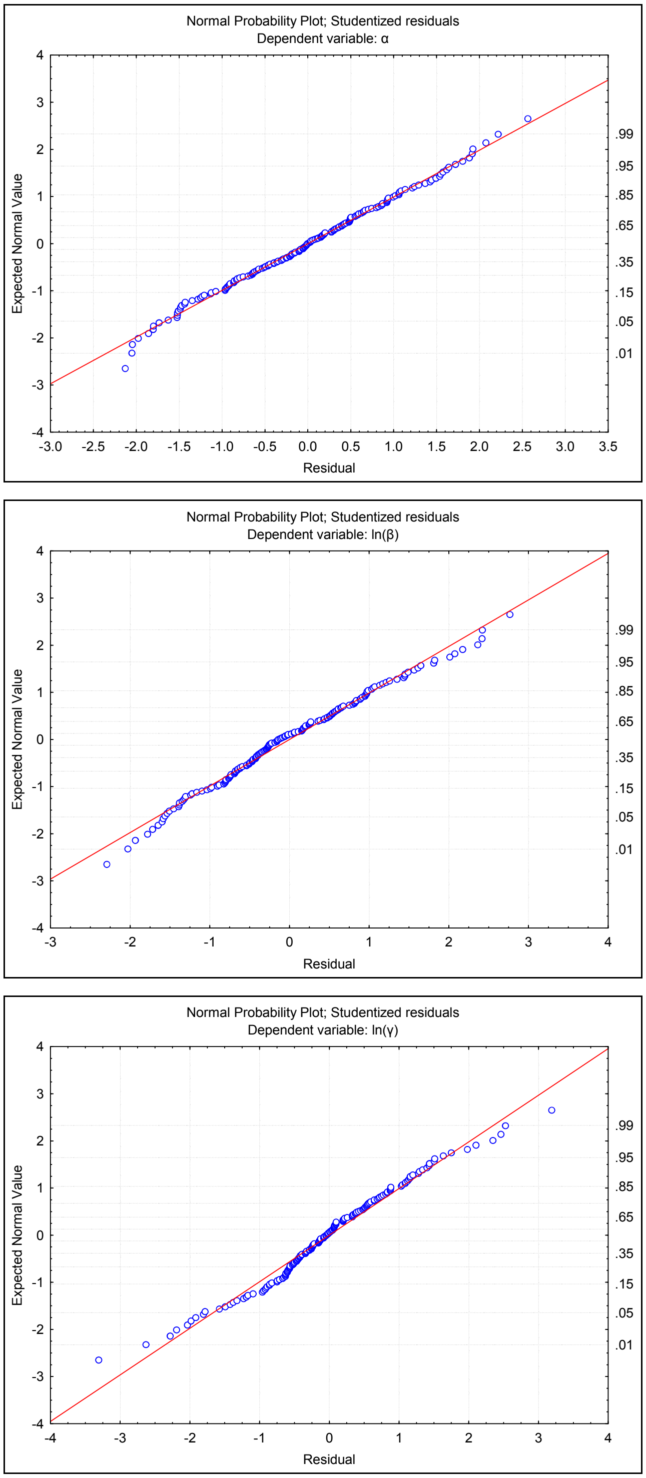


Figure K.4 Normal probability plot of studentised residuals

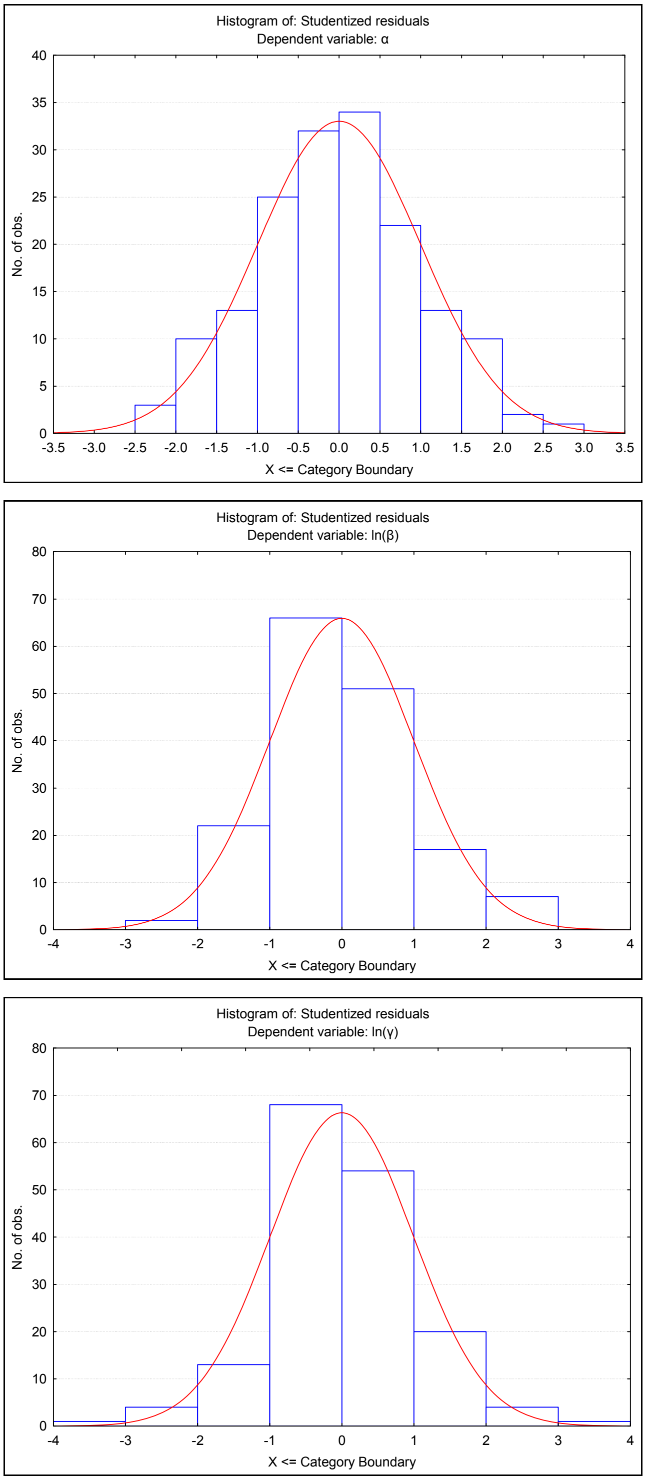


Figure K.5 Distribution of studentised residuals

Appendix L

SAS Program listing for the calculation of weights from the generalised variance for the fit of the growth curve model to a shrinkage profile

This Appendix contains the following sections:

- Input File
- SAS Program Listing
- Output

Notes regarding the output:

- The output entitled ‘Grid Search’ gives the results of the model fit at various starting values for the three model parameters. The best fit (lowest SS) is used as the starting point for the iterative phase of the model fitting algorithm.
- The next two sections, ‘Iterative Phase’ and ‘Estimation Summary’ relate to the execution of the Gauss-Newton Algorithm to estimate the parameters of the growth curve, and show that the criteria for convergence were met. No intercept was specified for this model since the growth curve model used does not have an intercept. The ‘missing’ data refers to additional values placed in the input file for model prediction.
- The next table shows that the model fit is highly significant. This is followed by the parameter estimates and their approximate 95% confidence intervals, as well as their correlation coefficients.
- This is followed by the covariance matrix of the three parameters, its determinant and its inverse, which is the weight output sought from the program.
- The program was run for each shrinkage profile individually and the weights were manually collected in the data file containing the covariate data and the growth curve parameters for each shrinkage profile, for the next step which was multivariate multiple regression employing these weights.

Input file

"Mix0275.xls"

Time	Shrinkage
0	0
10	68
50	194
100	255
200	345
302	373
401	395
492	391

Program Listing

```
PROC IMPORT OUT= CONCRETE.Mix0275
            DATAFILE= "C:\Documents and Settings\Petra Gaylard\My
                  Documents\Wits Masters\Research
                  Report\Analyses\x 9-10 Growth curves\All
                  profiles raw data\Mix0275.xls"
            DBMS=EXCEL REPLACE;
            SHEET="Sheet1$";
            GETNAMES=YES;
RUN;

goptions reset=all;
ods html;
ods graphics on;

data work;
   set concrete.Mix0275;
run;

proc nlin data=work best=10 hougaard smethod=golden maxiter=500
      outest=outinfo;
parameters a=100 to 800 by 50 b=.0010 to .0050 by .001
            g=.1 to 1 by 0.1;
      bounds a>0;
      bounds b>0;
      bounds g>0;
model Shrinkage = a*(1-exp(-b*time))**g;
output out=model_OLS predicted=pred residual=resid student=stu_resid
weight=weight;
run;

data outinfo2; set outinfo;
  if(_TYPE_="COVB");
  keep a b g;
  run;

proc iml;
use outinfo2;
read all into covmatrix;
print covmatrix;
determinant=det(covmatrix);
```

```
print determinant;
weight=1/sqrt(determinant);
print weight;
quit;

ods graphics off;
ods html close;
```

Output

```
The NLIN Procedure
                                Dependent Variable Shrinkage

                                       Grid Search
                                                              Sum of
                                 a           b           g     Squares

                             450.0     0.00400      0.5000      1347.7
                             450.0     0.00500      0.6000      1967.9
                             400.0     0.00500      0.5000      2186.3
                             450.0     0.00300      0.4000      2531.3
                             500.0     0.00200      0.4000      2533.8
                             450.0     0.00400      0.6000      3052.5
                             500.0     0.00300      0.5000      3488.7
                             450.0     0.00500      0.7000      3535.4
                             550.0     0.00200      0.5000      4111.7
                             600.0     0.00100      0.4000      4146.8
                                        Iterative Phase
                                                                     Sum of
                        Iter           a           b           g     Squares

                           0       450.0     0.00400      0.5000      1347.7
                           1       415.2     0.00554      0.5636       514.3
                           2       401.5     0.00704      0.6341       290.5
                           3       401.9     0.00761      0.6670       216.2
                           4       402.2     0.00757      0.6655       216.0
                           5       402.2     0.00757      0.6656       216.0
                           6       402.2     0.00757      0.6656       216.0
Execution Errors for OBS 1:
NOTE: Missing values were generated as a result of performing an operation on missing
values.
                NOTE: Convergence criterion met.
                                  Estimation Summary
                             Method                   Gauss-Newton
                             Iterations                          6
                             Subiterations                     105
                             Average Subiterations            17.5
                             R                           3.096E-7
                             PPC(b)                      4.893E-8
                             RPC(b)                      3.139E-6
                             Object                      2.03E-10
                             Objective                   216.0421
                             Observations Read                   8
                             Observations Used                   7
```

```
Observations Missing                1
NOTE: An intercept was not specified for this model.
```

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Model	3	674129	224710	4160.48	<.0001
Error	4	216.0	54.0105		
Uncorrected Total	7	674345			

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits		Skewness
a	402.2	7.4994	381.4	423.0	0.3748
b	0.00757	0.000998	0.00480	0.0103	0.1589
g	0.6656	0.0555	0.5114	0.8197	0.3094

Approximate Correlation Matrix

	a	b	g
a	1.0000000	-0.8209484	-0.5871992
b	-0.8209484	1.0000000	0.8882602
g	-0.5871992	0.8882602	1.0000000

COVMATRIX

56.241186	-0.006142	-0.244488
-0.006142	9.9525E-7	0.0000492
-0.244488	0.0000492	0.0030824

DETERMINANT

8.3894E-9

WEIGHT

10917.759

Appendix M

SAS Program listing for weighted multivariate multiple regression

This Appendix contains the following sections:

- SAS Program Listing
- Output

Notes regarding the output:

- The output is shown for the dependent variable α. The output is similar for the other two dependent variables.
- The first table shows that the overall model fit is highly significant.
- The next table shows the univariate significance tests for the covariates in the model.
- This is followed by the parameter estimates for all the covariates in the model.
- The last two sets of two tables each shows the assessment of the multivariate significance of two of the covariates in the model (similar output would be generated for each covariate). Of interest here is the multivariate significance of the covariate (as determined by Wilks' Lambda, Pillai's trace, the Hotelling-Lawley trace and Roy's Greatest Root), shown in the second table for each covariate.

Program Listing

```
PROC IMPORT OUT= CONCRETE.regression
            DATAFILE= "C:\Documents and Settings\Petra Gaylard\My
                  Documents\Wits Masters\Research
                  Report\Analyses\13 Multivariate multiple
                  regression\Data 2010 05 18 175 for SAS rem 3
                  outliers.xls"
            DBMS=EXCEL REPLACE;
            SHEET="Data";
            GETNAMES=YES;
RUN;

goptions reset=all;
ods html;
ods graphics on;

data work;
   set concrete.regression;
run;

proc glm outstat=model;
model a ln_b ln_g =  /*CemEarlyStrGainClass CEMIIAD CEMIIAL
      CEMIIAML*/ CEMIIASGGBS CEMIIASGGCS CEMIIASGGFS /*CEMIIAV*/
      CEMIIBMVL CEMIIBSGGBS CEMIIBSGGCS CEMIIBSGGFS /*CEMIIBV*/
      CEMIIIAGGBS CEMIIIAGGCS CEMIIIAGGFS CEMVA
      StoneWQ StoneShale StoneDolerite StoneGreywacke StoneDolomite
      StoneGranite StoneQuartzalluvial StonePretoriaQuartzite
      StoneTillite /*StoneQuartzite*/
      /*StoneRD  lnStoneWA StoneUC  StoneFI*/
      SandWQ /*SandShale*/ SandGranite SandDolerite SandAndesite
      SandKlipheuwelpit Sandnatural /*SandRiverWindhoek*/ SandCapeFlats
      SandRiverVaal
      SandPretoriaQuartzite /*SandTillite*/ SandQuartzite SandRiver
      SandEccaGrit
      /*SandRD SandFM*/
      lnCemContent
      /*FAContent GGBSContent GGCSContent GGFSContent LimestoneContent*/
      /*WaterContent SandContent StoneContent */ PVF
      /*CSFFract FAFract GGBSFract GGCSFract GGFSFract LimestoneFract*/
      /*WaterBinderRatio*/
      /*AggBinderRatio*/
      StoneSandRatio
      LabWits AgeDry ln2VolToSurfArea Temp /*Slump ComprStr*/;
manova H=_ALL_ / SUMMARY;
weight wtN;
** output out=model_OLS predicted=a_pred ln_b_pred ln_g_pred stdi=a_se
ln_b_se ln_g_se press=a_press ln_b_press ln_g_press;
run;

proc print data = model_OLS;
run;

ods graphics off;
ods html close;
```

Output

The GLM Procedure

Number of Observations Read	171
Number of Observations Used	167

Dependent Variable: a   a
Weight: wtN   wtN

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	39	6560328.609	168213.554	19.94	<.0001
Error	127	1071619.241	8437.947		
Corrected Total	166	7631947.850			

R-Square	Coeff Var	Root MSE	a Mean
0.859588	20.52386	91.85830	447.5683

Source	DF	Type I SS	Mean Square	F Value	Pr > F
CEMIIASGGBS	1	2452.773	2452.773	0.29	0.5907
CEMIIASGGCS	1	88461.860	88461.860	10.48	0.0015
CEMIIASGGFS	1	4635.742	4635.742	0.55	0.4599
CEMIIBMVL	1	21840.756	21840.756	2.59	0.1101
CEMIIBSGGBS	1	15748.994	15748.994	1.87	0.1743
CEMIIBSGGCS	1	114706.614	114706.614	13.59	0.0003
CEMIIBSGGFS	1	12130.625	12130.625	1.44	0.2328
CEMIIIAGGBS	1	122902.982	122902.982	14.57	0.0002
CEMIIIAGGCS	1	8508.726	8508.726	1.01	0.3172
CEMIIIAGGFS	1	1832.787	1832.787	0.22	0.6420
CEMVA	1	34967.380	34967.380	4.14	0.0439
StoneWQ	1	375501.072	375501.072	44.50	<.0001
StoneShale	1	589034.973	589034.973	69.81	<.0001
StoneDolerite	1	495186.974	495186.974	58.69	<.0001
StoneGreywacke	1	738452.100	738452.100	87.52	<.0001
StoneDolomite	1	541374.662	541374.662	64.16	<.0001
StoneGranite	1	22877.737	22877.737	2.71	0.1021
StoneQuartzalluvial	1	97.582	97.582	0.01	0.9145
StonePretoriaQuartzi	1	513661.627	513661.627	60.88	<.0001
StoneTillite	1	20309.021	20309.021	2.41	0.1233
SandWQ	1	150348.208	150348.208	17.82	<.0001
SandGranite	1	38639.786	38639.786	4.58	0.0343
SandDolerite	1	1118681.076	1118681.076	132.58	<.0001
SandAndesite	1	155036.413	155036.413	18.37	<.0001
SandKlipheuwelpit	1	422619.265	422619.265	50.09	<.0001
Sandnatural	1	599.362	599.362	0.07	0.7903
SandCapeFlats	1	8627.541	8627.541	1.02	0.3139
SandRiverVaal	1	201559.195	201559.195	23.89	<.0001
SandPretoriaQuartzit	1	19487.077	19487.077	2.31	0.1311
SandQuartzite	1	270717.107	270717.107	32.08	<.0001
SandRiver	1	22003.728	22003.728	2.61	0.1088
SandEccaGrit	1	53459.578	53459.578	6.34	0.0131
lnCemContent	1	4263.252	4263.252	0.51	0.4785
PVF	1	2723.801	2723.801	0.32	0.5709
StoneSandRatio	1	49880.257	49880.257	5.91	0.0164
LabWits	1	128282.357	128282.357	15.20	0.0002
AgeDry	1	7351.996	7351.996	0.87	0.3524
ln2VolToSurfArea	1	176488.189	176488.189	20.92	<.0001

Temp	1	4875.434	4875.434	0.58	0.4486

Source	DF	Type III SS	Mean Square	F Value	Pr > F
CEMIIASGGBS	1	1048.3324	1048.3324	0.12	0.7251
CEMIIASGGCS	1	3104.0489	3104.0489	0.37	0.5453
CEMIIASGGFS	1	19726.3265	19726.3265	2.34	0.1288
CEMIIBMVL	1	2470.9526	2470.9526	0.29	0.5894
CEMIIBSGGBS	1	10412.8302	10412.8302	1.23	0.2687
CEMIIBSGGCS	1	8021.2011	8021.2011	0.95	0.3314
CEMIIBSGGFS	1	7970.0854	7970.0854	0.94	0.3330
CEMIIIAGGBS	1	1211.9187	1211.9187	0.14	0.7053
CEMIIIAGGCS	1	65.9549	65.9549	0.01	0.9297
CEMIIIAGGFS	1	33966.4158	33966.4158	4.03	0.0469
CEMVA	1	7107.2477	7107.2477	0.84	0.3605
StoneWQ	1	15022.8721	15022.8721	1.78	0.1845
StoneShale	1	893.5677	893.5677	0.11	0.7454
StoneDolerite	1	72051.7280	72051.7280	8.54	0.0041
StoneGreywacke	1	145415.0879	145415.0879	17.23	<.0001
StoneDolomite	1	84218.8349	84218.8349	9.98	0.0020
StoneGranite	1	153726.6826	153726.6826	18.22	<.0001
StoneQuartzalluvial	1	383618.4387	383618.4387	45.46	<.0001
StonePretoriaQuartzi	1	34012.5301	34012.5301	4.03	0.0468
StoneTillite	1	32881.5085	32881.5085	3.90	0.0505
SandWQ	1	121184.6544	121184.6544	14.36	0.0002
SandGranite	1	26521.3268	26521.3268	3.14	0.0786
SandDolerite	1	138280.3418	138280.3418	16.39	<.0001
SandAndesite	1	55294.3510	55294.3510	6.55	0.0116
SandKlipheuwelpit	1	166711.9657	166711.9657	19.76	<.0001
Sandnatural	1	326038.5619	326038.5619	38.64	<.0001
SandCapeFlats	1	164968.5506	164968.5506	19.55	<.0001
SandRiverVaal	1	1504.1233	1504.1233	0.18	0.6736
SandPretoriaQuartzit	1	521.2188	521.2188	0.06	0.8041
SandQuartzite	1	14005.9606	14005.9606	1.66	0.2000
SandRiver	1	97207.3237	97207.3237	11.52	0.0009
SandEccaGrit	1	48525.9527	48525.9527	5.75	0.0179
lnCemContent	1	2393.2423	2393.2423	0.28	0.5953
PVF	1	3.1206	3.1206	0.00	0.9847
StoneSandRatio	1	52577.7536	52577.7536	6.23	0.0138
LabWits	1	113636.2304	113636.2304	13.47	0.0004
AgeDry	1	86621.2430	86621.2430	10.27	0.0017
ln2VolToSurfArea	1	36920.7391	36920.7391	4.38	0.0384
Temp	1	4875.4339	4875.4339	0.58	0.4486

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	1489.874462	872.822126	1.71	0.0903
CEMIIASGGBS	-9.317196	26.433467	-0.35	0.7251
CEMIIASGGCS	19.246674	31.732902	0.61	0.5453
CEMIIASGGFS	-48.795901	31.913808	-1.53	0.1288
CEMIIBMVL	-33.402984	61.726472	-0.54	0.5894
CEMIIBSGGBS	-28.033757	25.235697	-1.11	0.2687
CEMIIBSGGCS	33.838203	34.706114	0.97	0.3314
CEMIIBSGGFS	-33.901448	34.882303	-0.97	0.3330
CEMIIIAGGBS	-11.315599	29.857894	-0.38	0.7053
CEMIIIAGGCS	-3.159365	35.735078	-0.09	0.9297

CEMIIIAGGFS	-83.387877	41.562001	-2.01	0.0469
CEMVA	-46.690491	50.874026	-0.92	0.3605
StoneWQ	-296.892176	222.505453	-1.33	0.1845
StoneShale	-77.308230	237.563870	-0.33	0.7454
StoneDolerite	-668.475874	228.760969	-2.92	0.0041
StoneGreywacke	-4375.774310	1054.067857	-4.15	<.0001
StoneDolomite	-655.195201	207.388411	-3.16	0.0020
StoneGranite	-141.986915	33.265349	-4.27	<.0001
StoneQuartzalluvial	576.200022	85.455875	6.74	<.0001
StonePretoriaQuartzi	-577.539227	287.660618	-2.01	0.0468
StoneTillite	-398.482397	201.860649	-1.97	0.0505
SandWQ	-178.756514	47.168977	-3.79	0.0002
SandGranite	-395.716280	223.205394	-1.77	0.0786
SandDolerite	552.186104	136.402912	4.05	<.0001
SandAndesite	-604.092781	235.983509	-2.56	0.0116
SandKlipheuwelpit	4764.595280	1071.916660	4.44	<.0001
Sandnatural	359.479060	57.830574	6.22	<.0001
SandCapeFlats	4744.563196	1073.035394	4.42	<.0001
SandRiverVaal	-92.581349	219.280584	-0.42	0.6736
SandPretoriaQuartzit	-79.224344	318.762339	-0.25	0.8041
SandQuartzite	-328.998310	255.361610	-1.29	0.2000
SandRiver	3097.868165	912.707942	3.39	0.0009
SandEccaGrit	225.545583	94.051490	2.40	0.0179
lnCemContent	21.784228	40.904136	0.53	0.5953
PVF	6.226243	323.762655	0.02	0.9847
StoneSandRatio	-68.393555	27.398864	-2.50	0.0138
LabWits	-864.324433	235.524869	-3.67	0.0004
AgeDry	5.769387	1.800677	3.20	0.0017
ln2VolToSurfArea	-251.038758	120.011807	-2.09	0.0384
Temp	11.360114	14.944949	0.76	0.4486

Similarly for ln(b) and ln(g)

The GLM Procedure

Multivariate Analysis of Variance

Characteristic Roots and Vectors of: E Inverse * H, where

H = Type III SSCP Matrix for **CEMIIASGGBS**

E = Error SSCP Matrix

Characteristic Root	Percent	Characteristic Vector V'EV=1 a	ln_b	ln_g
0.01080342	100.00	0.00069721	0.14282572	0.10076627
0.00000000	0.00	-0.00000484	-0.15242032	0.33597827
0.00000000	0.00	0.00079607	-0.05976067	0.00000000

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall **CEMIIASGGBS** Effect

H = Type III SSCP Matrix for CEMIIASGGBS

E = Error SSCP Matrix

S=1 M=0.5 N=61.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.98931204	0.45	3	125	0.7176
Pillai's Trace	0.01068796	0.45	3	125	0.7176
Hotelling-Lawley Trace	0.01080342	0.45	3	125	0.7176
Roy's Greatest Root	0.01080342	0.45	3	125	0.7176

Characteristic Roots and Vectors of: E Inverse * H, where
H = Type III SSCP Matrix for **CEMIIASGGCS**
E = Error SSCP Matrix

Characteristic Root	Percent	Characteristic Vector V'EV=1 a	ln_b	ln_g
0.06684722	100.00	0.00035489	0.20718695	-0.22836821
0.00000000	0.00	0.00091211	-0.04409446	0.04860235
0.00000000	0.00	0.00040244	0.04828796	0.26176519

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall **CEMIIASGGCS** Effect
H = Type III SSCP Matrix for CEMIIASGGCS
E = Error SSCP Matrix
S=1 M=0.5 N=61.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.93734134	2.79	3	125	0.0436
Pillai's Trace	0.06265866	2.79	3	125	0.0436
Hotelling-Lawley Trace	0.06684722	2.79	3	125	0.0436
Roy's Greatest Root	0.06684722	2.79	3	125	0.0436

and so on for all the other terms.

Appendix N

Results of the diagnostic checks for the weighted multivariate multiple regressions

The results of the following diagnostic checks for the weighted multivariate multiple regressions are presented in Table N.1:

- Examination of the regression variate: Plots of studentised residuals versus predicted dependent variables, observed dependent variables and all covariates (those in the model and those not in the model).
- A check of the assumption of homoscedasticity: The plot of residuals versus the predicted dependent variable.
- Checks of the assumption of the normality of the error term distribution: Distribution and normal probability plot of residuals.
- Examination for outliers**:** Distribution of studentised residuals
- Influential and leverage points: Leverage, Mahalanobis D^2, studentised deleted residuals, Cook's distance and SDFFIT statistics.
- Multicollinearity: Variance Inflation Factor

Problematic results of diagnostic checks are indicated in red.

For each model, the diagnostic checks were repeated upon the removal of outliers and / or influential points. Model terms which subsequently were no longer significant are indicated at the end of the table.

Table N.1 Results of diagnostic checks for the weighted multivariate multiple regressions

Diagnostic	Model 19*W	**Model 19*W after deletion of 3 outliers (cases 169, 165 and 170)**	**Model 20*W after deletion of 3 outliers (cases 169, 165 and 170)**	**Model 21*W after deletion of 3 outliers (cases 169, 165 and 170)**	**Model 22*W after deletion of 3 outliers (cases 169, 165 and 170)**
Average adjusted R2 for the three dependent variables	0.676	0.691	0.697	0.698	0.732
Studentised residuals vs. predicted dependent variables	α: 1 residual > 6 ln(β): 1 residual > 4 ln(γ): 1 residual > 5 Shape of plot fine.	OK	OK	OK	OK
Studentised residuals vs. observed dependent variables	α: 1 residual > 4 ln(β): 1 residual near 4 ln(γ): 1 residual > 4 Shape of plot fine.	OK	OK	OK	OK
Studentised residuals vs. covariates in model	OK	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK	OK
Normal probability plot of residuals	Outliers evident as indicated above. Otherwise fine.	OK	OK	OK	OK
Distribution of residuals	Outliers evident as indicated above. Otherwise fine.	OK	OK	OK	OK
Observed vs. predicted dependent variables	α: value > 800 appears to be outlier ln(β): fine ln(γ): OK	OK	OK	OK	OK
Leverage	2p/n=0.096 (n=522) 14 cases exceed this value.	2p/n=0.096 (n=518) 14 cases exceed this value.	2p/n=0.096 (n=518) 20 cases exceed this value.	2p/n=0.096 (n=518) 21 cases exceed this value.	2p/n=0.127 (n=518) 25 cases exceed this value.
Mahalanobis D2	OK	OK	OK	OK	Case # 79 very large compared to others.
Studentised deleted residuals	11, 14 and 13 significant cases for α, ln(β) and ln(γ) respectively	13, 15 and 13 significant cases for α, ln(β) and ln(γ) respectively	14, 12 and 13 significant cases for α, ln(β) and ln(γ) respectively	13, 13 and 12 significant cases for α, ln(β) and ln(γ) respectively	16, 15 and 13 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	No significant cases.	No significant cases.	No significant cases.	No significant cases.	No significant cases.
SDFFIT	No significant cases.	No significant cases.	No significant cases.	No significant cases.	No significant cases.
Multicollinearity		ln(VTSA) (VIF = 14.7)	ln(VTSA) (VIF = 16.2) Sand Dolerite (VIF = 10.2)	ln(VTSA) (VIF = 16.5) Sand Dolerite (VIF = 10.6)	Some Stone Type terms, some Sand Type terms, Lab, Temp and ln(VTSA) have VIF > 10

Table N.1 Results of diagnostic checks for the weighted multivariate multiple regressions (continued)

Diagnostic	**Model 23*W after deletion of 3 outliers (cases 169, 165 and 170)**	Model 24*W after deletion of 3 outliers (cases 169, 165 and 170)	Model 24*W after deletion of 3 outliers (cases 169, 165 and 170) and 2 outliers (cases 20 and 25)	**Model 24*W after deletion of 3 outliers (cases 169, 165 and 170) and 2 outliers (cases 20 and 25) and leverage case # 79**	**Model 25*W after deletion of 3 outliers (cases 169, 165 and 170)**
Average adjusted R2 for the three dependent variables	0.731	0.724	0.738	0.738	0.708
Studentised residuals vs. predicted dependent variables	OK	α: OK ln(β): OK ln(γ): 1 residual > 4, 1 residual < -4 Shape of plot fine.	OK	OK	OK
Studentised residuals vs. observed dependent variables	OK	α: OK ln(β): OK ln(γ): 1 residual > 4, 1 residual < -4 Shape of plot fine.	OK	OK	OK
Studentised residuals vs. covariates in model	OK	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK	OK
Normal probability plot of residuals	OK	OK	OK	OK	OK
Distribution of residuals	OK	OK	OK	OK	OK
Observed vs. predicted dependent variables	OK	OK	OK	OK	OK
Leverage	2p/n=0.127 (n=518) 25 cases exceed this value.	2p/n=0.131 (n=518) 25 cases exceed this value.	2p/n=0.132 (n=514) 25 cases exceed this value.	2p/n=0.129 (n=511) 25 cases exceed this value.	2p/n=0.119 (n=518) 26 cases exceed this value.
Mahalanobis D2	Case # 79 very large compared to others.	Case # 79 very large compared to others.	Case # 79 very large compared to others.	OK	Case # 79 very large compared to others.
Studentised deleted residuals	16, 15 and 13 significant cases for α, ln(β) and ln(γ) respectively	12, 14 and 14 significant cases for α, ln(β) and ln(γ) respectively	13, 15 and 13 significant cases for α, ln(β) and ln(γ) respectively	10, 14 and 12 significant cases for α, ln(β) and ln(γ) respectively	15, 13 and 11 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	No significant cases.	No significant cases.	No significant cases.	No significant cases.	No significant cases.
SDFFIT	No significant cases.	No significant cases.	No significant cases.	No significant cases.	No significant cases.
Multicollinearity	Some Stone Type terms, some Sand Type terms, Lab, Temp and ln(VTSA) have VIF > 10	Some Stone Type terms, some Sand Type terms, Lab, Age Dry, Temp and ln(VTSA) have VIF > 10	Some Stone Type terms, some Sand Type terms, Lab, Age Dry, Temp and ln(VTSA) have VIF > 10	Some Stone Type terms, some Sand Type terms, Lab, Age Dry, Temp and ln(VTSA) have VIF > 10	Some Stone Type terms, some Sand Type terms, Lab, Temp and ln(VTSA) have VIF > 10

Table N.1 Results of diagnostic checks for the weighted multivariate multiple regressions (continued)

Diagnostic	**Model 26*W after deletion of 3 outliers (cases 169, 165 and 170)**	**Model 27*W after deletion of 3 outliers (cases 169, 165 and 170)**	**Model 28*W after deletion of 3 outliers (cases 169, 165 and 170)**	**Model 29*W after deletion of 3 outliers (cases 169, 165 and 170)**	**Model 30*W after deletion of 3 outliers (cases 169, 165 and 170)**
Average adjusted R2 for the three dependent variables	0.730	0.719	0.720	0.709	0.735
Studentised residuals vs. predicted dependent variables	OK	OK	OK	OK	OK
Studentised residuals vs. observed dependent variables	OK	OK	OK	OK	OK
Studentised residuals vs. covariates in model	OK	OK	OK	OK	OK
Studentised residuals vs. covariates not in model	OK	OK	OK	OK	OK
Residuals vs. time (Test Year)	OK	OK	OK	OK	OK
Normal probability plot of residuals	OK	OK	OK	OK	OK
Distribution of residuals	OK	OK	OK	OK	OK
Observed vs. predicted dependent variables	OK	OK	OK	OK	OK
Leverage	2p/n=0.143 (n=518) 26 cases exceed this value.	2p/n=0.127 (n=518) 29 cases exceed this value.	2p/n=0.127 (n=518) 28 cases exceed this value.	2p/n=0.119 (n=518) 17 cases exceed this value.	2p/n=0.116 (n=518) 35 cases exceed this value.
Mahalanobis D2	Case # 79 very large compared to others.	Case # 79 very large compared to others.	Case # 79 very large compared to others.	Case # 79 very large compared to others.	Case # 79 very large compared to others.
Studentised deleted residuals	17, 15 and 13 significant cases for α, ln(β) and ln(γ) respectively	10, 11 and 11 significant cases for α, ln(β) and ln(γ) respectively	10, 10 and 11 significant cases for α, ln(β) and ln(γ) respectively	16, 11 and 14 significant cases for α, ln(β) and ln(γ) respectively	13, 15 and 14 significant cases for α, ln(β) and ln(γ) respectively
Cook's distance	No significant cases.	No significant cases.	No significant cases.	No significant cases.	No significant cases.
SDFFIT	No significant cases.	No significant cases.	No significant cases.	No significant cases.	No significant cases.
Multicollinearity	Some Cement Type, some Stone Type terms, some Sand Type terms, Lab, Temp and ln(VTSA) have VIF > 10	Some Stone Type terms, some Sand Type terms, Lab, Age Dry, Temp and ln(VTSA) have VIF > 10	Some Stone Type terms, some Sand Type terms, Lab, Age Dry, Temp and ln(VTSA) have VIF > 10	Some Stone Type terms, some Sand Type terms, Lab, Temp and ln(VTSA) have VIF > 10	Some Stone Type terms, some Sand Type terms, Lab, Temp and ln(VTSA) have VIF > 10

Appendix O

SAS Program listing for the calculation of goodness-of-fit measures for the multivariate multiple regression models

This Appendix contains the following sections:

- SAS Program Listing
- Output

The program is used to calculate the MFPE, CAIC, CHQ and BIC goodness-of-fit indices for each regression model.

Program Listing

```
PROC IMPORT OUT= CONCRETE.model
            DATAFILE= "C:\Documents and Settings\Petra Gaylard\My
                  Documents\Wits Masters\Research
                  Report\Analyses\13 Multivariate multiple
                  regression\Model21A_SAS.xls"
            DBMS=EXCEL REPLACE;
            SHEET="Data";
            GETNAMES=YES;
RUN;

goptions reset=all;
ods html;
ods graphics on;

PROC IML;
 use concrete.model;
 read all var {y1 y2 y3} into Y;          /*Dependent variables*/
 read all var {x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15
      x16 x17 x18 x19 x20 x21 x22 x23 x24} into Xm;
            /*Independent variables*/

  n=NROW(Xm);      /* Number of cases*/
  k=NCOL(Xm)+1;    /* Total number of parameters incl intercept*/
  q=NCOL(Y);       /* Total number of dependent variables*/
  print n, k, q;

  Xp=J(N,1,1)||Xm;
  Hp=Xp*(INV(Xp`*Xp))*Xp`;
  SSE=Y`*(I(n)-Hp)*Y;
  MSE=det(SSE/(n-k));
  MFPE = det(MSE)*((n+k)/n)**q;
  CAIC = log(det(MSE)) + ((2*k*q)+(q*(q+1)))/(n-k-q-1);
  CHQ=log(det(MSE))+(2*q*k*LOG(LOG(n)))/(n-k-q-1);
  BIC=n*(log(det(SSE/n)))+((k*q)+(0.5*q*(q+1)))*LOG(n);
  print MFPE, CAIC, CHQ, BIC;

ods graphics off;
ods html close;
```

Output

```
     N
    167

     K
    25

     Q
     3

   MFPE
 8.3105661

   CAIC
 2.8729361

    CHQ
 3.4737649

    BIC
 617.0488
```

Appendix P

Values of $(Z'Z)^{-1}$ and $\hat{\Sigma}$ required for the calculation of the standard error of predicted values of α, ln(β) and ln(γ)

$(Z'Z)^{-1}$

69.882	-0.820	-1.078	-0.877	0.101	0.585	-0.020	0.820	1.221	8.810	0.380	-0.522	-0.277	-6.666	-1.484	-0.029	-0.002	-0.386	-8.250	-0.894
-0.820	0.121	0.072	0.002	0.015	0.016	-0.028	-0.045	-0.003	-0.064	-0.012	0.001	0.030	0.004	0.124	-0.001	0.000	0.018	0.039	0.000
-1.078	0.072	0.654	0.083	0.014	-0.003	-0.001	-0.006	-0.095	-0.079	-0.020	0.011	0.035	0.020	0.120	0.000	0.000	0.018	0.064	0.001
-0.877	0.002	0.083	0.209	0.032	0.023	-0.004	-0.026	-0.105	-0.098	0.008	0.000	-0.011	-0.143	0.003	0.001	0.000	0.003	0.096	0.015
0.101	0.015	0.014	0.032	0.292	0.095	0.030	0.025	0.023	-0.008	0.029	0.016	0.025	-0.372	0.038	-0.001	0.000	0.005	-0.001	-0.009
0.585	0.016	-0.003	0.023	0.095	0.613	0.021	0.017	0.034	0.020	0.037	0.006	0.017	-0.407	0.017	-0.002	0.000	-0.001	-0.032	-0.012
-0.020	-0.028	-0.001	-0.004	0.030	0.021	0.146	0.059	0.037	-0.008	0.018	0.049	0.048	0.013	0.007	0.000	0.000	0.003	0.023	-0.010
0.820	-0.045	-0.006	-0.026	0.025	0.017	0.059	0.562	0.064	0.125	0.014	0.040	0.051	-0.017	0.006	0.000	0.000	0.001	-0.112	-0.026
1.221	-0.003	-0.095	-0.105	0.023	0.034	0.037	0.064	0.150	0.152	0.025	0.028	0.040	-0.048	-0.005	-0.001	0.000	-0.002	-0.140	-0.026
8.810	-0.064	-0.079	-0.098	-0.008	0.020	-0.008	0.125	0.152	1.726	0.059	-0.097	-0.051	-0.827	-0.128	-0.001	0.000	-0.033	-1.328	-0.103
0.380	-0.012	-0.020	0.008	0.029	0.037	0.018	0.014	0.025	0.059	0.272	0.014	0.009	-0.063	-0.040	0.000	0.000	-0.009	-0.031	0.000
-0.522	0.001	0.011	0.000	0.016	0.006	0.049	0.040	0.028	-0.097	0.014	0.108	0.067	0.112	0.010	0.001	0.000	0.006	0.106	-0.009
-0.277	0.030	0.035	-0.011	0.025	0.017	0.048	0.051	0.040	-0.051	0.009	0.067	0.114	0.078	0.063	0.000	0.000	0.008	0.056	-0.018
-6.666	0.004	0.020	-0.143	-0.372	-0.407	0.013	-0.017	-0.048	-0.827	-0.063	0.112	0.078	2.861	-0.029	0.004	0.000	-0.006	0.881	0.098
-1.484	0.124	0.120	0.003	0.038	0.017	0.007	0.006	-0.005	-0.128	-0.040	0.010	0.063	-0.029	0.229	-0.001	0.000	0.034	0.087	-0.005
-0.029	-0.001	0.000	0.001	-0.001	-0.002	0.000	0.000	-0.001	-0.001	0.000	0.001	0.000	0.004	-0.001	0.000	0.000	0.000	0.002	0.000
-0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-0.386	0.018	0.018	0.003	0.005	-0.001	0.003	0.001	-0.002	-0.033	-0.009	0.006	0.008	-0.006	0.034	0.000	0.000	0.010	0.023	0.001
-8.250	0.039	0.064	0.096	-0.001	-0.032	0.023	-0.112	-0.140	-1.328	-0.031	0.106	0.056	0.881	0.087	0.002	0.000	0.023	1.239	0.104
-0.894	0.000	0.001	0.015	-0.009	-0.012	-0.010	-0.026	-0.026	-0.103	0.000	-0.009	-0.018	0.098	-0.005	0.000	0.000	0.001	0.104	0.020

$\hat{\Sigma}$

3553.9819	-4.498731	-4.968883
-4.498731	0.0922294	0.0351575
-4.968883	0.0351575	0.0360255