EVALUATION OF MUNICIPAL WATER DEMAND AND RELATED PARAMETERS

HENDRINA JOHANNA VAN ZYL

A Dissertation submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, in fulfilment of the requirements of the degree of Master of Science in Engineering

Johannesburg 2006

ABSTRACT

South Africa is a water scarce country that constantly strives to apply its available water resources in the most efficient and equitable manner. Different users including industry, domestic, agriculture and the environment, vie for the available resources and have to be awarded an equitable and adequate share to ensure sustainability in the long term. Proper water demand management in South Africa is thus of critical importance. Researchers, municipal engineers and engineering consultants require good and reliable data and design guidelines on water demand patterns and factors influencing water demand to achieve this aim.

The most commonly used South African guideline for municipal water demand was first published in 1983 (CSIR, 1983), and is currently included in the document titled "Guidelines for human settlement planning and design" (CSIR, 2003). The guideline provides upper and lower limits for domestic consumer demands as a function of stand area. This study analyses measured water consumption data of more than a million stands (domestic and non-domestic) contained in the municipal treasury data for a cross-selection of South African towns and cities. The main objectives of the study were to evaluate the existing South African domestic demand guidelines and to evaluate the parameters that influence domestic and non-domestic water demand.

The main results of the study indicate that the domestic demands of most of the users fall well within the upper and lower limit envelope curves of the South African design guideline although for larger stands the design guideline was found to be conservative. Stepwise multiple variable regressions were applied to domestic and non-domestic consumption data to determine the most significant variables influencing water demand. In a large majority of cases, either the stand area or stand value had the greatest significance. There is strong evidence that both domestic and non-domestic demand increase with increasing stand area and stand value. In the case of domestic demand, stand value was used as an indicator for household income level. The study found that inland domestic water demand is significantly and consistently higher than coastal demand. This trend was not observed for non-domestic demand. However, the development level (city vs. small town) was found to have a marked impact on non-domestic demand. The study also found that the frequency distributions of annual average daily demand (AADD) for specific non-domestic user categories seem to have a lognormal distribution. Preliminary revisions to the existing domestic demand estimation guideline is proposed. Preliminary estimation curves for seven non-domestic user categories are presented as an indication of what demand estimation guidelines for non-domestic users could look like.

DECLARATION

I hereby declare that this dissertation submitted for a Masters degree in Civil Engineering at the University of the Witwatersrand, apart from the help recognised, is my own work and has not formerly been submitted to another university for a degree.

____

H. J. VAN ZYL

ACKNOWLEDGEMENTS

My gratitude is due to many individuals and institutions who have helped directly and indirectly in the production of this dissertation:

To my Maker, Who gave me the ability and opportunity to discover more of His creation also in this study.

I am indebted to Dr Adesola Ilemobade for his guidance, advice and enthusiasm.

I would like to thank Kobus, my husband, for his support, guidance and valuable advice throughout my studies.

This project would not have been possible without financial support by the Water Research Commission. I would like to extend a word of appreciation for this opportunity.

I would like to express my gratitude to GLS Consulting in Stellenbosch for providing data, SWIFT® software and assistance. I would especially like to thank Leon Geustyn, Heinz Jacobs and Kerry Fair.

I am also indebted to my employer, VelaVKE for allowing me the freedom to pursue my own personal goals and I would like to specially mention David Gertzen in this regard.

This project was possible due to the co-operation of many institutions. I therefore wish to extend my gratitude to the following:

- Rand Water for the appointment to undertake the study.
- Municipalities that made water consumption records available for analysis.
- The South African Weather service for providing climatic data.
- The South African Demarcation Board for making socio-economic statistics available.
- Statistical Consulting services of the University of Johannesburg for assistance with the statistical analyses, Prof Francois Steffens in particular.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

AADD	Annual Average Daily Demand (kl/day)
CVM	Contingent Valuation Method
DL	Development Level
GL	Geographic Location
MAP	Mean Annual Precipitation (mm)
MAE	Mean Annual Evaporation (mm)
SAWS	South African Weather Service
WRC	Water Research Commission (South Africa)

1 INTRODUCTION

1.1 Background

The underlying motive of this study was the conservation of a limited natural resource, water, that is indispensable for human life. Proper planning and future water demand management is essential in an economically developing and water scarce country such as South Africa.

A key input in water demand management and planning for municipal services is the estimation of present and prediction of future water demand. Water demand estimates are used to calculate peak water demand and sewer flows and thus determine municipal water and sewer infrastructure requirements, which in turn decide water authorities' budgets and capital investment needs.

The literature review that was undertaken as part of this study indicated that annual average water consumption is a function of a large number of variables, including type of supply, land use, climate, stand area, population density, and the socio-economic profile of the supply area. The document "Guidelines for Human Settlement Planning and Design" (CSIR, 2003) is commonly used in South Africa to estimate municipal water demand. In this guideline, water demand is linked to the type of supply and whether the supply area is a developing or developed community. The guideline provides upper and lower limits for annual average daily demand in residential stands based on stand area. The designer has to take other factors, such as climate and income into account when selecting an appropriate demand for a given area. How this is to be done is not suggested.

The expansion of urban areas, the continuing development taking place in South Africa and the constant need for potable water services have created a requirement for more accurate water demand estimates. Inaccurate estimates lead to a deficiency in basic design information that could lead to inadequate service or inequitable water distribution.

1.2 Objectives

The following aims were set for the study:

i. A comprehensive literature review of previous work done in South Africa on domestic and non-domestic water consumption patterns and the dependency of these consumption patterns on various parameters.
ii. Analysis of treasury data of domestic and non-domestic water consumption for a cross selection of South African towns (rural) and cities (urban) to determine the extent of dependency on the following and other possible parameters:
    a. Stand Area
    b. Household income
    c. Household size
    d. Geographic location
    e. Climate / annual rainfall / evaporation
    f. User type (i.e. land use)
iii. Evaluate current South African guidelines for domestic and non-domestic water demand given the outcome of the above analyses.

1.3 Scope of the Study

The main difference in the methodology of this study and those of many previous studies is that in this study, municipal water meter readings were used to determine water demand information. This made it possible to study a very large number of consumers. A large number of records made it possible to conduct meticulous statistical analyses, to investigate the distribution of the data in greater detail and to have representative samples of specific data characteristics.

The overriding problem with studying water demand is that quality data is not readily available. Even a large logging exercise can realistically only reach a small proportion of users. Furthermore, a logging survey is expensive and therefore, further limits the length of the monitoring period and the number of users monitored. This study considered municipal water meter readings to be an ideal source of water demand information, since the readings are taken on a regular basis by virtually all municipalities in South Africa. Possible arguments against using meter readings for studying water demand are:

- Consumer meters do not accurately register the amount of water used;
- It is difficult to access and extract water demand information from municipal treasury systems;

- Meter readings are not always taken monthly and are estimated for some months;
- Meter readings are not always accurate as meters clock over or meter replacements take place;
- Customer information such as address, income level or user type contained in treasury systems is not always accurate.

It is true that the accuracy of a consumer meter declines over the years. However, it is in the interest of the utility (municipality) and the consumer that the meter register as accurately as possible the amount of water used and therefore, meter maintenance programmes should be in place in most utilities. Meters seldom are designed to under-register as they age in order to benefit the customer rather that the water utility (Garlipp, 1979). It was assumed for this study that the accuracy of the consumer meters studied is adequate.

Until recently, the wealth of water demand information in municipal treasury systems was difficult to access. Actual meter readings were often hidden in complicated database setups or could not be directly or easily linked to user information. In most cases it was not possible to analyse data programmatically i.e. using a computer and software. However, the past decade has seen significant software developments that now enable engineers to abstract and analyse demand information from treasury databases for selected municipalities that have employed these software tools (Jacobs et al., 2004). One such a software tool is Swift®. This software allows a user to interrogate and access municipal treasury databases to obtain demographic data, stand characteristics (area and value) and recorded water consumption for individual consumer connections. A number of municipalities have implemented Swift® including Tshwane Metropolitan Municipality, Ekurhuleni, Johannesburg Water and most of the municipalities of the Western Cape. The existing databases cover years of consumption data for hundreds of thousands of users. This study therefore, with the collaboration of GLS Consulting Engineers and the various municipalities, extracted recorded monthly meter readings with the associated demographic and stand characteristics for more than 2.5 million stands for at least a period of 2 years from 48 different municipalities countrywide.

To address problems like meter clock overs or replacements, this study made use of the data cleaning functions contained in Swift®. In addition to the Swift® data cleaning procedures this study also followed a rigorous data cleaning process to ensure the integrity of the final data used for analysis. Even given the thorough data cleaning and filtering procedures that were applied, it is expected that some inaccurate data will still be contained in the dataset. This is one of the limitations of this study. However, a very large number of records (more than a million) was analysed and therefore, although data inaccuracy will inevitably lead to some degree of variation in the final results, significant correlations and trends were still expected.

Municipalities used in the study, were selected based on their economic importance and distribution to represent different climatic and economic regions of South Africa and of course, on the availability of the data. It was decided to undertake the water demand part of the study in five main tasks:

Task 1: Identify and confirm the towns and cities in terms of the available data and willingness to be involved in the study. Collect the available data.

Task 2: Extract the relevant data from the available treasury databases. Verify and clean the data. Obtain specific characteristics of each dataset in order to confirm that a representative sample of users will be studied with regards to economic, climatic and user type characteristics.

Task 3: Data analyses to determine relationships between the average daily demand and stand area, stand value, household income level, household size and other potential parameters.

Task 4: Evaluation of the current South African guidelines (CSIR, 2003) commonly used to estimate domestic and non-domestic water demand, given the outcome of the study analyses.

Task 5: Documentation of the results.

The study has a number of limitations, including the following:

- Water consumption is an inherently variable process and any measured data will thus include a measure of variability and uncertainty.
- Alternative water sources were not considered in this study. The treasury data did not identify stands with alternative water sources. The most common alternative sources are groundwater from boreholes, rainwater collected from roofs and on-site re-use of grey water. Usually water from alternative sources in residential developments is used for garden irrigation. This will definitely influence water demand patterns in the affected stands (most likely larger residential stands).

- The study investigated demand patterns of non-domestic water demand. However, this analysis relied greatly on the accuracy of the user type codes assigned to the non-domestic users in the treasury data.
- The climatic parameters that were included relate to the measurement years of the treasury data for the specific datasets. The weather parameters during the time of demand measurement were not compared to the long term average to check whether the measured water demand was subjected to significant influences by abnormal weather patterns.

1.4 Layout of the Document

This dissertation comprises six chapters and three appendices. Chapter 2 summarises a literature review of previous work done on domestic and non-domestic water consumption patterns and also discusses the existing water demand estimation guidelines (CSIR, 2003).

Chapter 3 constitutes a detailed discussion of the data used in this study. The data collection, cleaning and verification processes are discussed and the data is described in detail. Chapter 4 describes the methodology followed with regards to the data analysis.

Chapter 5 provides an in-depth summary of the all the results of this study and Chapter 6 concludes the findings of the study. In Chapter 6 recommendations for future work are also made.

Appendix A gives a comprehensive summary of the characteristics of the data that was used in this study. Appendix B summarises climatic data that was used in the study with regards to mean annual precipitation and mean annual evaporation measurements. Appendix C summarises all the regression results that were obtained with the data analysis. Appendix C is provided electronically on CD-Rom as the size (325 printed pages) makes it impractical to provide a hard copy.

2 LITERATURE REVIEW

2.1 South African Water Demand Guidelines

2.1.1 Domestic Water Demand

The estimation of peak water demands often consisted of estimating the population, multiplying by an average daily per capita use and then applying peak-to-average ratios based on entire cities (Howe & Linaweaver, 1967). It has been recognised that domestic water demand estimates should be preferably based on actual water consumptions per township as recorded by the municipal treasury (City of Johannesburg 1989, Howe & Linaweaver 1967). However, information on actual water consumptions is not always readily available and as a consequence guidelines for domestic demand estimation are still mostly based on stand area (Jacobs et al., 2004; CSIR, 2003)

The first guideline that was compiled in South Africa with the aim to provide information with regards to the provision of engineering services in residential townships was the so-called "Blue Book" (CSIR 1983), taking its name from the ring binder in which it was issued. It was published by the Department of Community Development and it was based on the experience of various municipal, design and planning engineers and town planners and it had the input of several technical committees. One of the sections of the "Blue Book" is dedicated to water supply. It contains information regarding design criteria, materials, construction and provides guidelines for demand estimation for water reticulation design and storage facilities. The "Blue Book" is mainly only applicable to urban residential areas with access to water-borne sanitation.

In the late 1980's the Department of Development Aid with support of the South African Housing Advice Council developed a guideline for the provision of engineering services for developing communities with a focus on low cost services, the so-called "Green Book" (1986). In 1994, the CSIR published a revised guideline that addressed and combined the guidelines of the "Blue Book" and the "Green Book" with the title "Guidelines for the Provision of Engineering Services and Amenities in Residential Township Development" the so-called "Red Book" (CSIR, 1994). The "Red Book" has been revised since its publication in 1994. The first "Red Book" was considered to have a number of shortcomings which restricted its usefulness in the drive to produce sustainable and vibrant human settlements as opposed to mere human settlements (CSIR, 2003). In terms of its mandate, the CSIR Division of Building and Construction Technology has undertaken to maintain the “Red Book” as a continually updated “living document” (CSIR, 2003). A revision of the ”Red Book” was published in 2000, with another revision in August 2003. The revisions in August 2003 applied to Chapter 9-Water Supply and Chapter 10-Sanitation.

The average water demand estimation guidelines of the most recent publication of the “Red Book” (CSIR, 2003) have remained unchanged since the first publication of the original guideline in the “Blue Book” (CSIR, 1983). However, the most recent publication of the guideline distinguishes between water demand in developing and developed areas. The following definitions are given (CSIR, 2003):

“Developing areas are considered to be those areas where the level of services to be installed may be subject to future upgrading to a higher level.”

“Developed areas are considered to be those areas where the services installed are already at their highest level and therefore will not require future upgrading.”

Table 2.1 and Table 2.2 summarise the “Red Book” guideline for domestic water demand in developing areas.

Table 2-1: Domestic Water Demand for Developing Areas (CSIR, 2003 – Table 9.10)

Type of Water Supply	Typical Consumption (litre/ca/d)	Range (litre/ca/d)
Communal Water Point		
• Well or standpipe at considerable distance (>1000 m)	7	5-10
• Well or standpipe at medium distance (250 - 1 000 m)	12	10-15
• Well nearby (<250 m)	20	15-25

Table 2-2: Domestic Water Demand in Developing Areas Equipped with Standpipes, Yard Connections and House Connections (CSIR, 2003 – Table 9.11)

Type of Water Supply	Type of Consumption (litre/ca/d)	Range (litre/ca/d)
Standpipe (200m walking)	25	10-50
Yard Connection With dry sanitation With LOFLOs With full-flush sanitation	55	50-100 30-60 45-75 60-100
House connection (developed areas)		60-475

Type of Water Supply	Type of Consumption (litre/ca/d)	Range (litre/ca/d)
Development level:		
Moderate	80	48-98
Moderate to high	130	80-145
High	250	130-280
Very high	450	260-480

For developed areas, the average daily water demand for domestic users (single residential stands) is based on stand area as represented in Figure 2.1 below (CSIR, 2003). Two envelope curves are given, an upper and a lower limit. Although, the "Red Book" mentions certain factors influencing domestic water demand (i.e. climate, income level, cost of water), it does not incorporate them into the estimates of the lower and upper demand limits. The user is therefore left without assistance on how to incorporate the factors into the demand estimates.

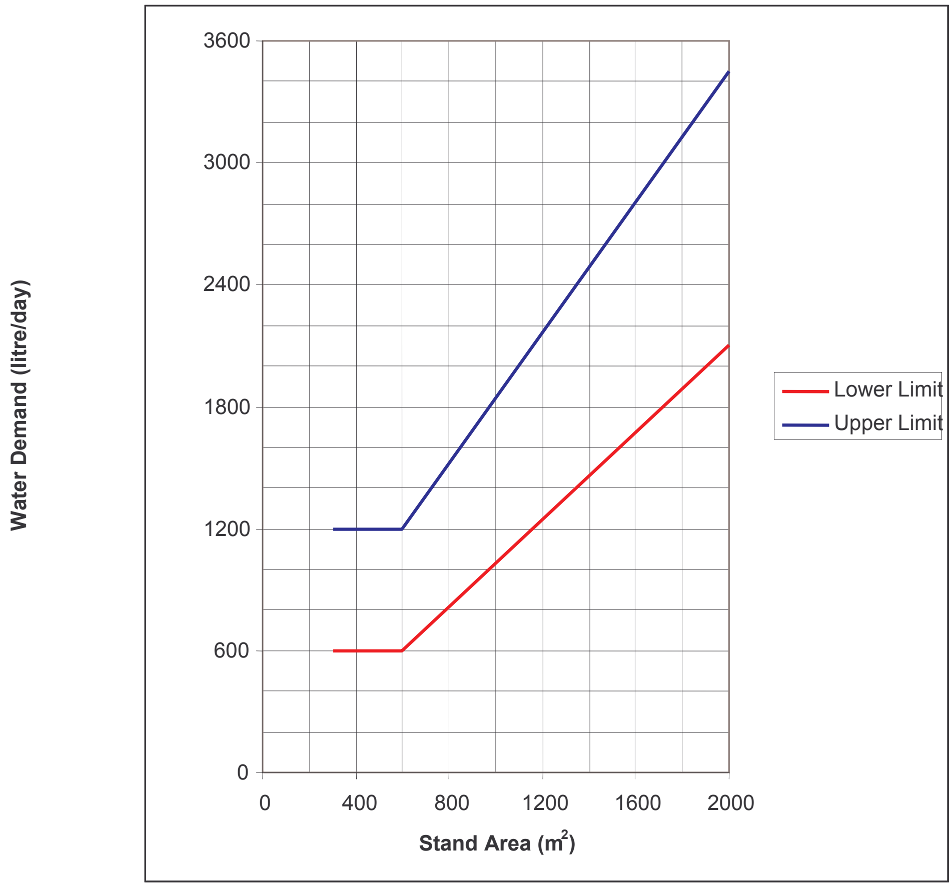


Figure 2-1: Annual Average Daily Water Demand for Domestic Use in Developed Areas (CSIR, 2003 – Figure 9.9)

New guideline curves for residential water demand estimation were proposed in 2004 by Jacobs et al. The updated guidelines are based on stand area as the only influencing variable, but take into account four different geographic regions and also provide separate estimation curves for township and suburban areas. Jacobs et al. (2004) investigated nearly 600 000 domestic users country-wide using municipal treasury data and found the "Red Book" guideline to be too conservative. Despite the fact that the work by Jacobs et al. provides a much improved stand-area based guideline for estimating residential demand, it still has the limitation of being stand-area based only (i.e. a single variable model).

2.1.2 Non-Domestic Water Demand

It is generally recommended that non-domestic water demands should be based on field measurements as it is extremely difficult to estimate non-domestic demand (CSIR, 2003). The City of Johannesburg also recommends in its water supply guidelines that non-domestic demands should be determined where possible from the City Treasurer's records on actual water consumption (City of Johannesburg, 1989).

The "Red Book" guideline with regards to non-domestic water demand is summarised in Tables 2.3 for developing areas and Table 2.4 for developed areas.

Table 2-3: Non-Domestic Water Demand in Developing Areas (CSIR, 2003 – Table 9.12)

Non-Domestic Users	Water Demand
Schools: Day Boarding	 15-20 90-140 litres/pupil/day
Hospitals	220-300 litres/bed/day
Clinics	5 litres/bed/day – out patients 40-60 litres /bed/day – in patients
Bus stations	15 litres/user/day for those persons outside the community
Community Halls / Restaurants	65-90 litres/seat/day

Table 2-4: Non-Domestic Water Demand in Developed Areas (CSIR, 2003 – Extract of Table 9.14)

Category	Type of Development	Unit	Annual Average Water Demand (litres/day) unless otherwise stated
4	Offices and Shops	100 m^2 of gross floor area[a]	400
5	Government and municipal	100 m^2 of gross floor area	400
6	Clinic	100 m^2 of gross floor area	500

Category	Type of Development	Unit	Annual Average Water Demand (litres/day) unless otherwise stated
7	Church	Erf	2000
8	Hostels	Occupant	150
9	Developed Parks	Hectare of erf area	=<2 ha: 15 kl[b,c] >2 ha and =<10 ha : 12.5 kl >10 ha: 10 kl
10	Day School / Crèche	Hectare of erf area	As per developed parks[d]
11	Boarding School	Hectare of erf area plus boarders	As per developed parks plus 150 litre/boarder
12	Sports ground	Hectare of erf area	As per developed parks

a: Gross floor area obtained using applicable floor space ratio from town planning scheme

b: Demand for developed parks to be considered as drawn over six hours on any particular day in order to obtain the peak demand

c: Where the designer anticipates the development of parks and sports grounds to be of a high standard, e.g. 25mm of water applied per week, the annual average water demand should be taken as follows: =<2 ha: 50 klitre; >2 ha and =<10 ha: 40 klitre ; >10 ha : 30 klitre

2.2 South African Studies of Water Demand

The following sections provide a summary of the literature review specifically with regards to previous work done in South Africa in the field of municipal water demand estimation. The highlights and limitations of the previous studies are summarised to bring to light the remaining unresolved problems that this study intended to address.

2.2.1 Garlipp (1979)

The research conducted by Garlipp studied domestic demand in various South African cities and also the possible factors influencing domestic demand. Water consumption was studied for cities as a whole (Pretoria, Bloemfontein, Cape Town, Port Elizabeth and Durban) and for individual consumers and sectors. Data was sourced from meter readings and water meter books (individual customers). Sample sizes were approximately 20 % of the residential sectors of the cities studied.

The study provides a breakdown of internal and external domestic water consumption in three South African cities namely Durban, Johannesburg (Witwatersrand) and Cape Town (Cape

Peninsula). The data for this analysis was obtained by sending out questionnaires to the engineering population of South Africa. The author found that 73 % of domestic consumption in the Witwatersrand was for outdoor use compared to the Cape Peninsula where only 40 % of domestic water consumed was used outdoors. In Durban, 45 % of domestic water consumed was used outdoors. Garlipp also found that the average daily domestic consumption in the Witwatersrand (2240 litres/stand/day) was significantly more than in the Cape Peninsula (914 litres/stand/day).

Garlipp also studied climatic variables and related water consumption patterns. This was investigated for entire cities (i.e. base unit is a city). It was found that after prolonged rainfall, water consumption decreased. Temperature had a positive correlation with water consumption i.e. with an increase in temperature, water consumption increased. An interesting finding is that domestic water was largely consumed internally at lower temperatures and externally at higher temperatures. This study evaluated seasonal variation in water demand for entire cities (i.e. base unit is a city) in Southern Africa and found that at least one month each year was found to have a monthly consumption that was less than 80% of the average annual monthly consumption. The author mentioned a study on Southern African cities that indicated that water consumption could be the result of differential tariff structures, restriction of water flows or restricting consumption for certain purposes such as garden watering in drought periods. The author pointed out that differential water tariffs in Windhoek saved approximately 20% water over a 6 month period. Garlipp stated that metering in conjunction with regular reading and an effective tariff structure and diligent collection system could reduce water consumption significantly.

Garlipp found the most significant parameter that influences domestic water consumption to be household size. This South African study indicated that water consumption per capita increased with stand area and income but decreased with an increase in household size. Household size did not affect external domestic water use. This study found that income influenced domestic water consumption positively and followed an S- curve. Stand area was also found to have a positive correlation with domestic water use. It significantly influenced external domestic use. The type of stand coverage (i.e. grass, paving, shrubbery etc.) was one of the main factors determining external domestic use. Garlipp also considered stand area as a good proxy for income. Boreholes were also found to significantly affect external domestic use. It was found that less water was consumed externally on stands with access to boreholes.

Highlights and Limitations of the Study

i. The study by Garlipp is the first of its kind in South Africa and was conducted before the publication of the “Blue Book”. It provided a valuable base for further research in water demand estimation in South Africa.

ii. The study investigated the effect of a number of factors on domestic water demand patterns in urban areas countrywide in five different cities. The study found household size was the most significant parameter that influenced domestic water demand. Other factors that positively affected domestic demand were income, stand area (only external use) and prolonged high temperatures. Access to borehole water and rainfall had a negative correlation with domestic water demand.

iii. The study measured water demand for cities as a whole and thus evaluated the effect of climatic factors on the basis of an entire city. Socio-economic data such as income, stand area and household size were collected by means of surveys. The response on these surveys was poor and various questionnaires had to be sent out. Some of the surveys were only conducted among the engineering fraternity of South Africa, which may be viewed as a biased sample.

iv. The study distinguished between ethnicity in most of its results, the research being conducted in a previous political era. This makes it difficult to compare with research being conducted in current day South Africa.

2.2.2 Stephenson and Turner (1996)

The study by Stephenson and Turner provides an important understanding of urban water demand patterns. The research focused on the Gauteng area and investigated one high income residential area (242 stands), seven medium income residential areas (7 119 stands), two low income residential areas (2 370 stands) and three suburbs where the land use was predominantly of a commercial and industrial nature. The definition used in the study for low, middle and high income is summarised as:

Very Low Income:	<R 10 000 per annum per household
Low Income:	>=R 10 000 and < R 25 000 per annum per household
Middle Income:	>=R 25 000 and < R 50 000 per annum per household
High Income:	>= R 50 000 per annum per household

The study areas were isolated with regards to their water supply. The water supply into the areas was measured by means of water meters and data loggers. Statistical data on characteristics such as population and income was obtained from the Central Statistical Services.

Domestic Demand - Effect of Stand Area

The relationship between the average stand area and the average AADD (Annual Average Daily Demand) in each study area is presented in the article. It should be observed that the average AADD value for each residential area included water losses and possibly some non-residential use (may be greater than residential use). Stephenson and Turner compared the findings of their study with the "Red Book" guideline (CSIR, 1994 and CSIR, 2003), for lower and higher domestic demand limits. Figure 2.2 represents this comparison.

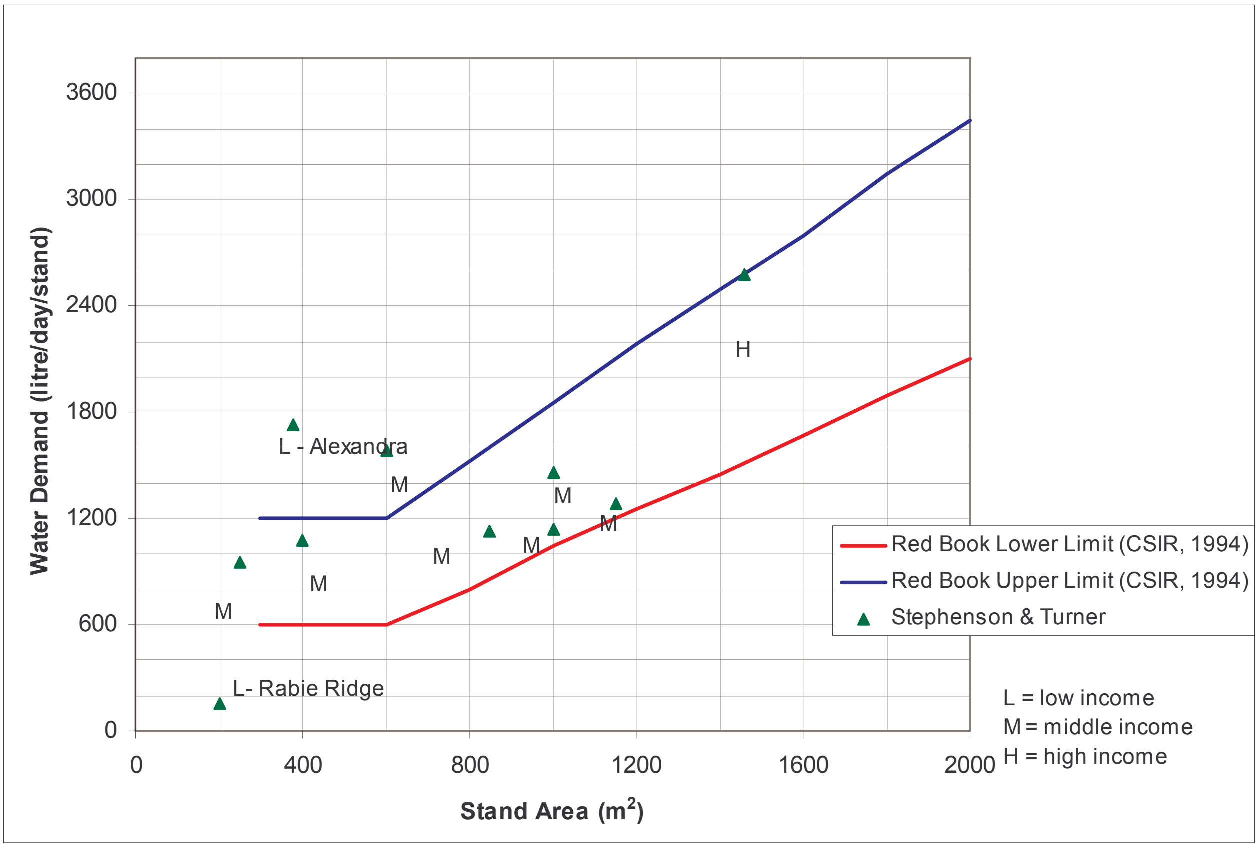


Figure 2-2: Evaluation of Existing Guidelines for Domestic Water Demand in the Gauteng Area (Stephenson and Turner, 1996)

Figure 2.2 indicates that the average per stand water demand of the majority of the study areas fell within the design guideline envelope recommended by the "Red Book" (CSIR, 1994 and 2003).

The areas whose average per stand consumptions did not fall within the guideline envelope were low income areas. The one area, Alexandra, had domestic water demand much higher than predicted by the guideline. Stephenson and Turner noted that the study area in Alexandra was unusually densely populated which resulted in very high demand per stand. The other low income area investigated by Stephenson and Turner, i.e. Rabie Ridge had significantly lower water demand than anticipated by the lower limit of the guideline curve (CSIR, 1994 and 2003). Stephenson and Turner indicated that the reason might be that Rabie Ridge had no house connections or waterborne sanitation (at the time of the study.)

This study concluded that generally it could be said that stand area had a direct influence on water demand. However, the type of housing had an effect and also the level of service (water and sanitation) had a significant impact, as was evident in the Rabie Ridge study area.

Domestic Demand – Effect of Income

The study reported that it was commonly acknowledged that domestic water demand is directly proportional to income per stand and income per person. The data analysis verified this statement. However, the study indicated that where this relationship was significant for per person consumptions, it was not as evident in the per stand consumptions. Figure 2.3 below shows the work of Stephenson and Turner for income versus per stand water demand in the Gauteng area.

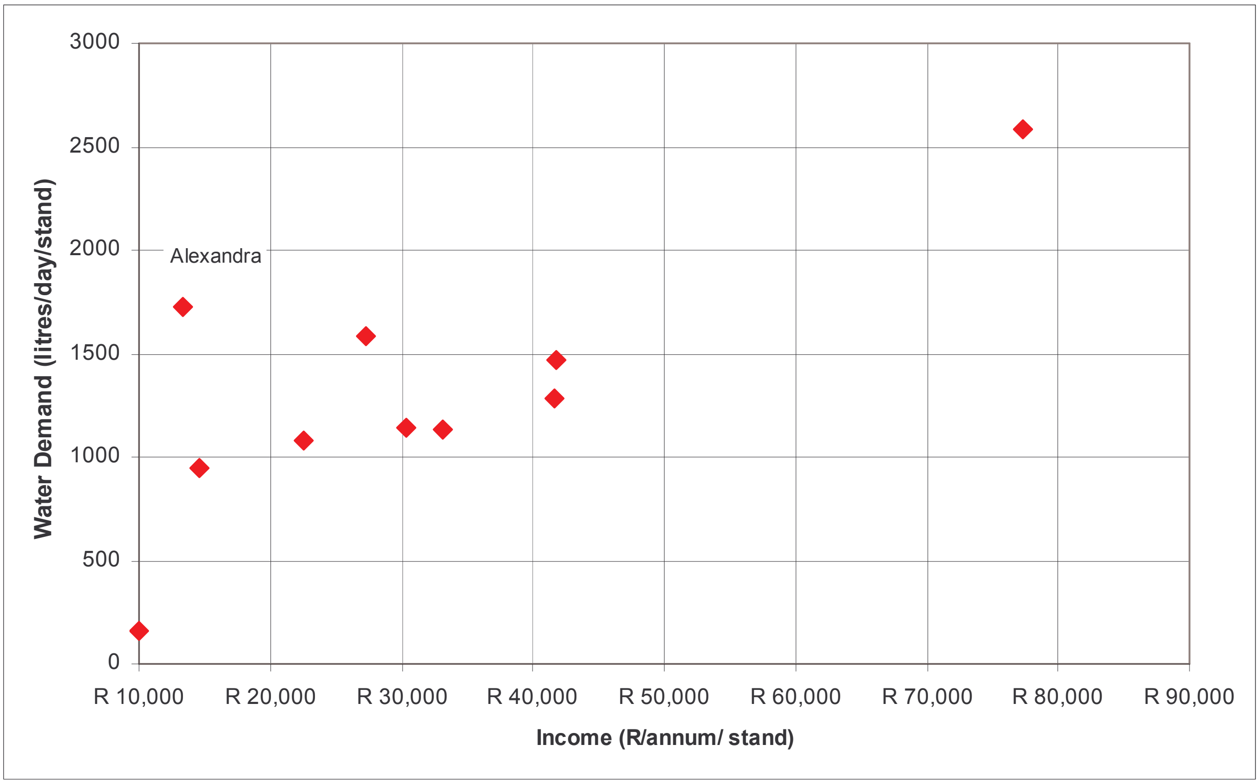


Figure 2-3: Effect of Income – (Stephenson and Turner, 1996)

The study area in Alexandra exhibited a high per stand water demand for a low income level (Figure 2.3). Compared to the other data points, this could be seen as an outlier. Stephenson and Turner gave the reason for Alexandra's exceptional high water demand as the fact that the study area in Alexandra was very densely populated.

<u>*Highlights and Limitations of the Study*</u>

i. The study investigated a substantial number of users (9 731 domestic stands) in the Gauteng area for all income levels and provides a valuable base for further research.

ii. Stephenson and Turner confirmed that domestic water demand can be related to stand area as recommended by the "Red Book" (CSIR 1994 and 2003) and its related guidelines.

iii. The study confirmed that factors such as income, population density, supply type, housing type can substantially influence water demand and thus result in deviations from the "Red Book" guidelines.

iv. The study investigated both formal residential developments and less formal residential developments such as Rabie Ridge (i.e. suburban versus township stands).

v. A possible limitation is the use of average stand area for all the stands in a study area or zone, as this could lead to the misrepresentation of stand area.

vi. The AADD presented by Stephenson and Turner for each study area can be expected to be higher than the actual domestic water demand because it included water losses and possibly even fire water demand and some non-domestic water use.

2.2.3 Van Vuuren and Van Beek (1997)

In 1997, Van Vuuren and Van Beek undertook a study in the Pretoria supply area for the Water Research Commission with the collaboration of the Municipality of Pretoria to review existing guidelines for urban domestic and industrial water demand, based on measured water consumptions. The study investigated domestic water consumption data for a period from March 1982 to October 1994 for 53 reservoir supply areas. The analysis distinguished between high, middle and low income users. Non-domestic water consumption was also examined in 16 reservoir supply areas with an acceptable proportion of industrial users, including Rosslyn industrial area. The results of the study provide valuable insights on possible factors influencing domestic and non-domestic water demand.

Domestic Demand

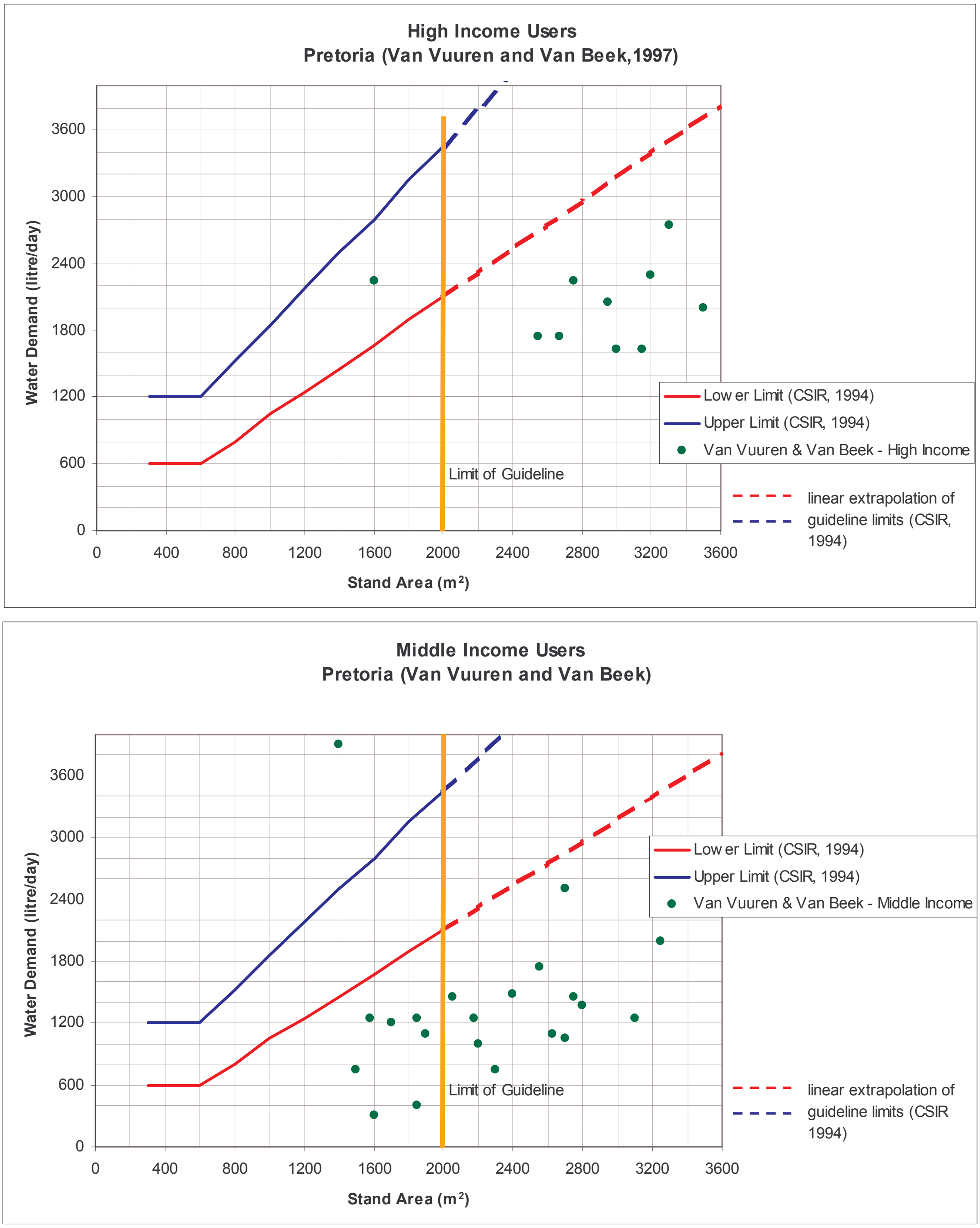
The study indicated a strong correlation between domestic demand and the income level of the users. High income users consumed significantly more water than middle and low income users. The study found that climate (rainfall and temperature) had a significant influence on water demand patterns. However, the income status of a household influenced specifically the outdoor water demand, which was closely linked to climate. It was shown that the influence of climate on domestic water demand, in low income areas, was negligible since outdoor water use was much less in these areas. An interesting finding of their work is that consumers, without exception, decreased their water consumption with the implementation of water restrictions. However, the investigation indicated that high income users took longer, than middle and low income users to respond to the implementation of the restrictions but recovered quicker to their pre-restriction water consumption level when the restrictions were lifted.

Non-Domestic Demand

The results indicated a significant correlation between the total area of the industrial development and water consumption. The study investigated the influence of climate (rainfall and temperature) on industrial water demand patterns and found no significant correlation.

Evaluation of Existing Guidelines

Van Vuuren and Van Beek evaluated the existing "Red Book" guidelines (CSIR, 1994) for domestic demand estimation with the results of their data analysis for Pretoria. They compare the lower and higher limits of the guideline for the three income level groups (low, middle and high) defined in their study. This comparison is represented graphically below (Van Vuuren and Van Beek ,1997) in Figure 2.4.



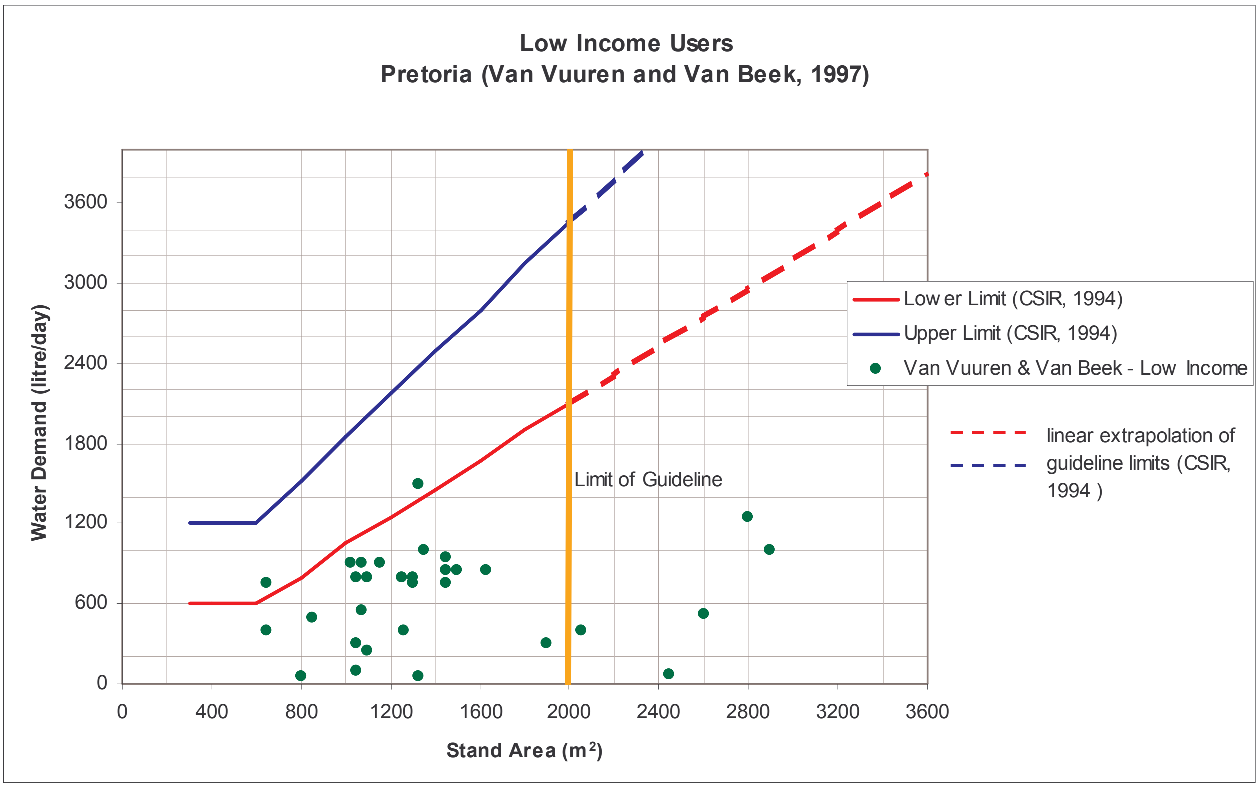


Figure 2-4: Evaluation of Existing Guidelines for Domestic Water Demand for the Pretoria Supply Area (Van Vuuren and Van Beek, 1997)

The study by Van Vuuren and Van Beek found that the domestic water demand in the Pretoria supply area was less than what the guideline (CSIR, 1994) recommends for all the income level groups defined in their study.

<u>*Highlights and Limitations of the Study*</u>

i. The study investigated the water demand of both domestic and non-domestic users for the entire Pretoria supply area, based on measured consumptions for a period of approximately twelve years (March 1982 to October 1994).

ii. The study confirmed a strong correlation between domestic demand and income per stand.

iii. The work presented interesting findings that relate climate to water demand and income per stand levels. High income users consumed more water than middle and low income users. Water demand patterns of high income users were more climate sensitive than middle and low income users.

iv. All users responded to water restrictions, although high income users took longer to respond. High income users however, recovered quicker when the restrictions were lifted.

v. The study evaluated the existing design guidelines (CSIR, 1994 and 2003) and found that the daily domestic water demand in the Pretoria supply area was less than the guideline envelope for all the income categories defined in the study.

vi. With the evaluation of the existing guidelines, the majority group with regards to income in a supply area was determined and the average AADD and average stand are for the group was determined.

vii. Non-domestic water demand was found not to be sensitive to climate (temperature and rainfall) for the non-domestic users investigated in this study.

viii. The study had limitations with regards to the data accuracy of some parameters, such as the meter reading intervals and land use characteristics.

ix. Income levels were defined and users categorised by means of property tax information.

x. The study was limited to formal residential developments with per stand water connections and water-borne sanitation. Informal residential settlements are excluded from the study.

xi. It is of note that stand areas for the low income category considered with this work were significantly greater (ranging from 600 to 2 900 m^2) than those stand areas considered by Stephenson and Turner (1996) for the low income category, which ranged from 200 – 400 m^2.

2.2.4 Veck and Bill (2000)

Veck and Bill undertook a study for the Water Research Commission of South Africa in 2000 to investigate the estimation of the price elasticity of residential water demand for different income groups. They approached the problem of estimating the price elasticity of residential water demand using a Contingent Valuation Method or model (CVM). Contingent evaluation models are models where the input data is obtained by means of a survey to acquire perceived information from individuals (users). In social-psychological terms it is a measure of behavioural intention. In the case of this study, willingness to pay for water at increased tariffs was used as an indicator of these behavioural intentions. A CVM is valuable in areas where data on the parameters that influence water demand is limited. The results of this approach are not as reliable or accurate as actual measured responses.

The researchers conducted the study in Alberton and Thokoza. In Alberton, 111 users were interviewed and 50 users in Thokoza. The study comprised a two-stage survey. The objective of the first stage of the survey was to obtain water demand profiles of the different users groups being investigated. The second stage involved the acquisition of Contingent Values (i.e. answers to the questions such as how much water the users would buy if the price increased). The study indicated that the price of water was an important consideration so far as domestic consumption was concerned. The authors found that the residential water demand in the study area had a negative price elasticity. In other words, water demand decreased with increases in price. The study found the price elasticity of indoor domestic water demand to be more or less the same for all income groups (low, middle and high). However, the price elasticity for outdoor domestic water demand was significantly higher (more negative) for high and middle income users than for low income users. The price elasticity for low income users for indoor and outdoor use was very similar.

Price elasticity changes with time and it is therefore possible to distinguish between short-term and long term elasticity. Price increases will have an immediate effect on water demand patterns (i.e. short term elasticity) but will not immediately change house and plumbing fixtures. In the longer term, increased water tariffs result in water saving plumbing fixtures and thus higher elasticity values. The findings of this study gave short term price elasticities obtained through the Contingent Valuation Method. An econometric model was also used to gauge long-term price elasticity.

The findings of the study compared well with similar work overseas, using CVM and short-term macro-econometric methods to determine price elasticity. Table 2.5 summarises the short term price elasticities for the various income groups in Alberton/Thokoza that resulted from employing the Contingent Valuation Method by Veck and Bill (2000):

Table 2-5: Water Price Elasticity (Veck and Bill, 2000)

Group	No of Respondents	Indoor	Outdoor	Total
High Income	52	-0.14	-0.47	-0.19
Middle Income	59	-0.12	-0.46	-0.17
Low Income	50	-0.14	-0.19	-0.14

Figure 2.5 represents the impact of a price increase on water consumption, using these short term elasticities that Veck and Bill (2000) found.

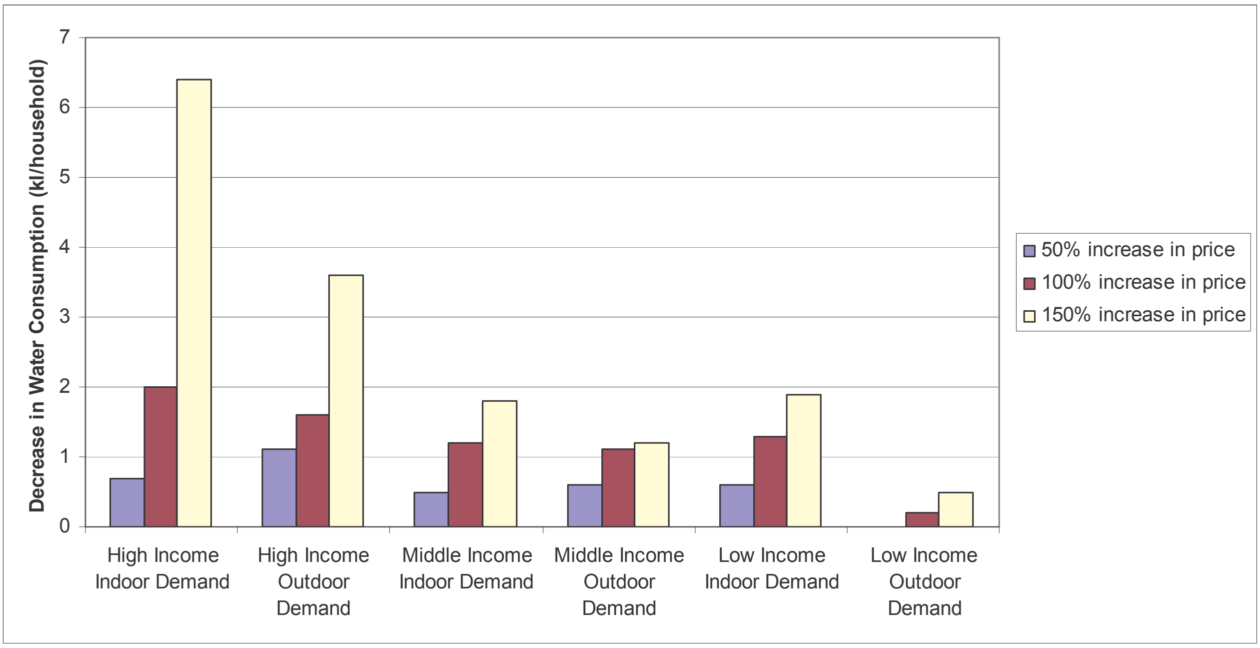


Figure 2-5: Effect of an Increase in Water Price on Water Demand (Veck and Bill, 2000)

Veck and Bill found, with a macro-econometric model, that the long term price elasticity in the study area for all income groups was *-0.73*. Compared to the average short term elasticity obtained with the study, it gave an indication of how much more price increases could affect water demand in the long term.

Veck and Bill also investigated the water demand profile of the users in the study area, not through direct measurement of water consumption but through the Contingent Valuation Method. The results can thus be termed *perceived* water demand of the users. The results of their study in this regard are presented below in Table 2.6.

Table 2-6: Perceived Water Usage for Various Income Groups (Veck and Bill, 2000)

Group	Perceived Water Demand (kL/month)	Perceived Water Demand (litres/day)
Alberton Middle Income	23	756
Alberton High Income	28	921
Thokoza Low Income	20	658

Highlights and Limitations of the Study

i. The study by Veck and Bill is one of the first of its kind in South Africa, investigating the impact of price on water demand using a Contingent Valuation Method.

ii. The results of the study compared well with similar studies undertaken overseas.

iii. The study indicated that the price elasticity of water demand was negative, which implied that if water prices increased, demand would decrease. An interesting finding was that the outdoor demand was more price elastic than indoor demand for high and middle income groups. This finding corresponded with the findings of a significant study in this field undertaken in the USA (Howe and Linaweaver, 1967).
iv. The study was limited to 150 respondents who partook in the contingent valuation survey.
v. The contingent valuation method, although it has its merits in areas where limited data on demand variables is available, has its limitations. The major limitation being, that it is based on perceived information obtained from users and is thus less accurate than actual measured responses to price increases.

2.2.5 Van Zyl et al. (2003)

Van Zyl et al (2003) investigated the elasticity of water price, pressure, income and stand area for residential water demand in the residential areas of Alberton, Boksburg, Centurion and Midrand in Gauteng Province, comprising more than 110 000 users. The study used end-use modelling. In end-use modelling, points of water consumption are modelled and thus the characteristics of end-uses and the behaviour of users can be incorporated in a very detailed model. The work was commissioned by Rand Water as a pilot study to investigate the strengths and weaknesses of end-use modelling as a water demand predictor.

The study grouped the end-uses into indoor consumption, outdoor consumption and leakage. The variables examined were household income, water price, stand area and available pressure.

Data for the study was obtained through Rand Water consumer surveys in the study areas, treasury data obtained through the SWIFT® software package and published local and international research. The authors of this paper also distinguished between suburb and township stands, based on property valuations included in the treasury data. The study analysed the data to identify ranges of elasticity values for the modelling parameters (income, stand area, pressure and price). The effect of these elasticity values on water demand was then evaluated through a sensitivity analysis. The findings of the sensitivity analysis are discussed below:

Domestic Demand – The Effect of Stand Area

Treasury data for more than 110 000 users was analysed to obtain stand area elasticities for water demand. Only data for users consuming more than 0.1 kl/day and less than 30 kl/day and with stands between 200 m^2 and 2 000 m^2 was used. The data was grouped into four stand value categories and the elasticity for each category was determined.

The study assumed that indoor consumption is not affected by stand area i.e. the elasticity of stand area for indoor water demand is zero. Thus, the net effect of a change in stand area was a function of the outdoor demand elasticities. The study indicated that stand area elasticity for *outdoor* water demand was high for both suburbs and townships, namely 1.6 for suburbs and 1.28 for townships. The results of the sensitivity analysis showed that a *50 %* reduction in stand area would result in a decrease in water demand of *28-40 %* in suburbs and *12 %* in townships. However, Van Zyl et al. noted that reduction in stand area might imply densification (such as sub-division for townhouse developments) which might result in increased water demand of the total supply area. The authors noted that water demand increase due to densification in township areas were usually not much tempered by the accompanying decrease in stand area, because outdoor water demand was a small fraction of the total water demand in township areas.

Domestic Demand – The Effect of Income

The study indicated that income had a significant effect on water demand. The study analysed detailed Rand Water consumer survey data generated in 2001. The resulting income elasticity values of *0.28* for suburban areas and *0.21* for townships were generated.

The results of the sensitivity analysis showed that a 20 % increase in real household income would result in water demand increasing between *4-7 %* in suburban areas and between *2-8* % in townships. Van Zyl et al. emphasized the point that although income clearly had an effect on consumption, it might not affect the overall consumption of a given area or suburb by much due to movement of people, with large changes in income, in or out of the suburb.

Domestic Demand – The Effect of Water Price

The study differentiated between townships and suburban areas and indoor and outdoor domestic demand. Both long term and short term elasticity were considered with the sensitivity analysis. For suburbs, the study assumed 50 % outdoor consumption and 20 % for townships. The elasticity values used in the sensitivity analysis were obtained from published local and international research. The results indicated that the price of water was an important determinant of water demand. A summary of the sensitivity analysis is represented in tabular format below:

Table 2-7: Effect of Water Price on Domestic Demand (Van Zyl et al., 2003)

		Townships		**Suburbs**	
		Decrease in Demand	**Elasticity**	**Decrease in Demand**	**Elasticity**
Short Term Effect	50% Increase in Price	2 -25 %	-0.0 to -0.8	7-15%	-0.05 to -0.5
Long Term Effect	50% Increase in Price	3 -44%	-0.0 to -1.6	13 -27%	-0.1 to -1.0

The authors discussed the results and explained the larger variation in townships by two factors, namely that:

- generally lower income levels in townships ensured a greater negative reaction with water price increases
- water consumption in townships was already low, close to basic human need, and even if the water price increased, people were not able to reduce their consumption by much.

The researchers observed that the findings for townships were also complicated by the issue of non-payment for services and the free basic water policy implemented by government.

Domestic Demand – The Effect of Pressure

Van Zyl et al demonstrated that pressure had a small but significant effect on domestic water demand. The researchers indicated that pressure mainly affected leakage in a water distribution system, although pressure could be expected to have an effect on non-leakage consumption as well. Water waste is usually linked to pressure in a system. The study used pressure elasticity values of between *0.15* and *0.25* for domestic demand which were based on the estimated effect on actual consumption and specifically excluded losses in the network.

The study found that a 50 % reduction in available pressure would result in a decrease of between *10 -16* % in water demand in suburban areas and a decrease of between *7-13 %* in townships (excluding reduction in leakage).

Highlights and Limitations of the Study

i. The work presented by Van Zyl et al provides essential pointers to factors that have an impact on domestic demand patterns. The merit of the work lies also in the fact that typical South African conditions are investigated with regards to suburban and township developments.

ii. The study is limited to a pilot study and the aim was not to develop a comprehensive model of domestic demand. Potential parameters influencing water demand such as climate, geographic location, level of service and age of infrastructure were not considered in the analysis.

iii. Although the study focused on a restricted number of variables in Gauteng only, the results can serve as general markers to parameters that influence water demand.

iv. The study indicated that price had the most significant impact on domestic water demand patterns.

v. Household income, stand area and pressure had positive demand elasticities, i.e. an increase in income, stand area or pressure resulted in an increased water demand.

vi. The study indicated the strengths and weaknesses of end-use modelling and reported on its potential application in South Africa

vii. The sensitivity analysis was limited to one factor at a time and although useful in highlighting the impact of the individual parameters, it cannot be used to estimate the combined effects of different user types and parameters.

viii. The end-use model was limited to two types of domestic users namely suburban and townships. Additional user categories such as parks, schools, business commercial sites etc. had not been considered.

2.2.6 Jacobs et al. (2004)

The work by Jacobs et al presents a valuable discussion with regards to estimating residential water demand in southern Africa using a single-coefficient model that relates water demand with stand area. The study analysed the measured water consumption of 582 997 single residential stands in various towns and cities in southern Africa by using treasury databases. The following municipalities were investigated:

- **Ekurhuleni Metropolitan Municipality**: Alberton, Thokoza, Benoni, Daveyton, Etwatwa, Boksburg, Vosloorus, Brakpan, Germiston, Kathlehong,, Kempton Park, Tembisa, Springs, Kwa-Thema
- **Cape Town**: Blaauwberg, Helderberg, Tygerberg (West, Central and South)
- **Tshwane (Gauteng)**: Akasia, Atteridgeville, Centurion, Mamelodi, Pretoria, Soshanguve, Randfontein
- **Johannesburg**: Midrand
- **George**
- **Windhoek** (Namibia)

The study analysed consumption data of at least 12 months varying from December 1999 to July 2003 for only single residential stands, with one water meter per stand (and vice versa) and using less than 20 kl per day. Only stands with areas between 50 m^2 and 2 050 m^2 were used in the study. For the Ekurhuleni dataset the stands are categorised as either "suburban" or "township". This classification was made by viewing the treasury records with SG (Surveyor General) cadastral information in GIS (Geographic Information System) format and then classifying the stands as suburban or township based on knowledge of the areas.

Domestic Demand – Effect of Stand area

The researchers of this paper proved that a strong relationship exists between domestic AADD (Annual Average Daily Demand) and stand area. Although the work acknowledged that numerous factors influence water demand, stand area is used as a single explanatory variable in the model. The authors noted that a model based on stand area alone has limited application and should be used only when better estimation methods are not available. The study presented three unique models for three Southern African geographic regions that could be used to estimate the AADD for single residential stands, namely:

- Coastal winter rainfall region
- Inland summer rainfall region (for suburban and township developments separately)
- Coastal all-year rainfall region

For each model, a guideline curve and an upper and a lower envelope curve are presented. Figures 2.6, 2.7, 2.8 and 2.9 represent the three models and their guideline curves and also compares the curves with the existing "Red Book" guidelines. This investigation found that the "Red Book" guidelines were too conservative, leading to possible over-design and potential unnecessary expenditure by water authorities.

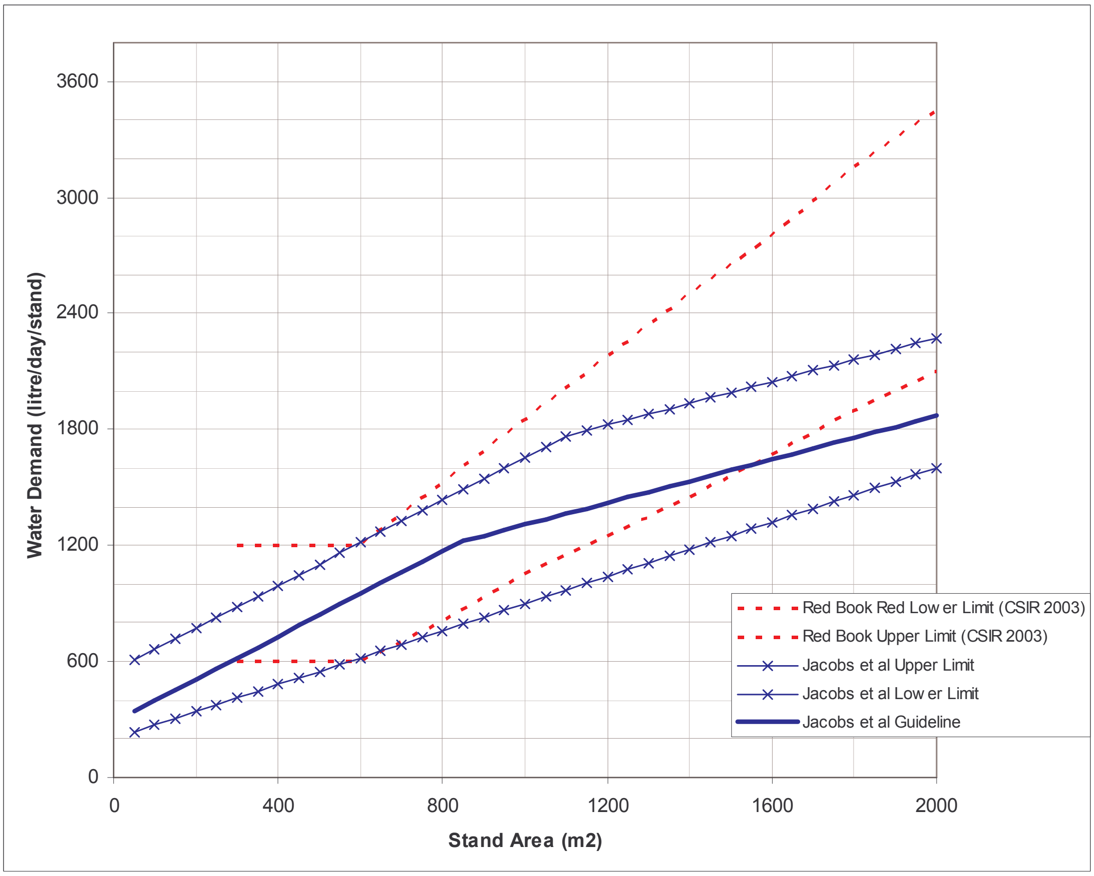


Figure 2-6: Southern African Coastal Winter Rainfall Region Model (Suburban and Townships) (Jacobs et al., 2004)

Jacobs et al compared their findings with other work that evaluated residential domestic demand. The results of Jacobs et al compared very well with the findings of Stephenson and Turner (1996), Van Zyl et al. (2003) and a similar study in Windhoek (Water Transfer Consultants 2001, WCE et al 2003).

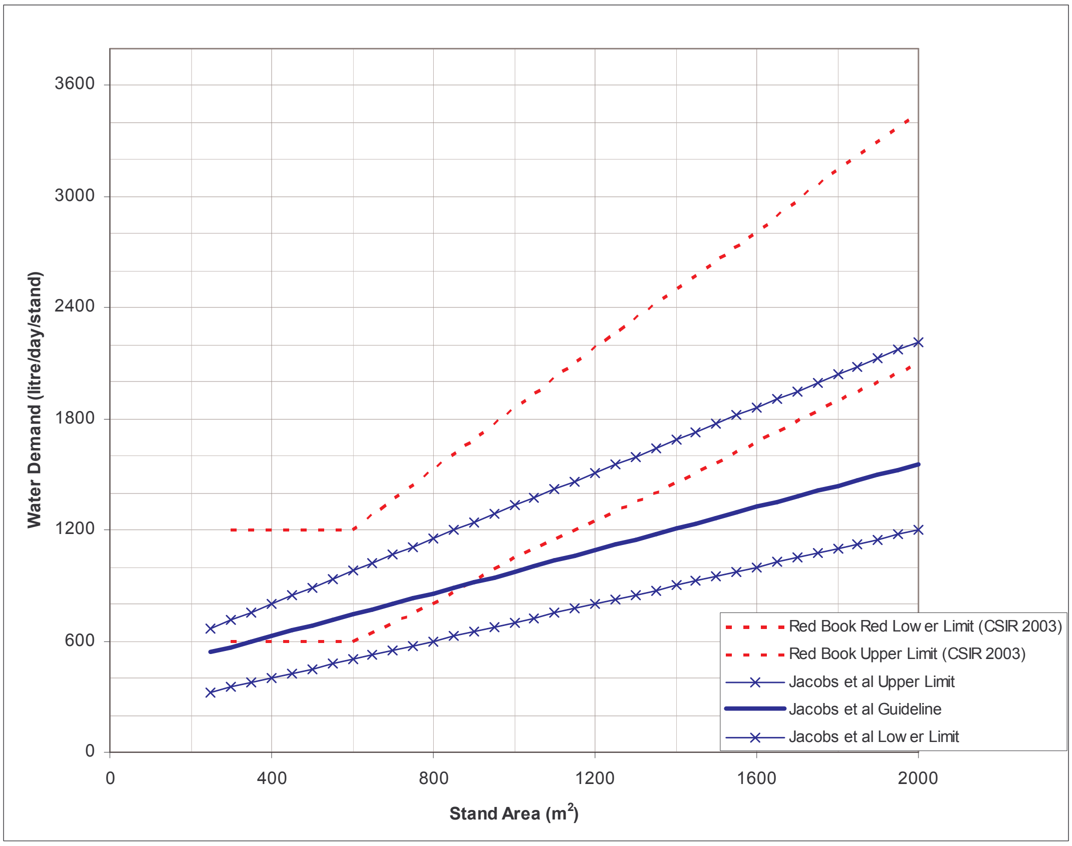


Figure 2-7: Southern Africa Inland Summer Rainfall Region Model- Suburban (Jacobs et al., 2004)

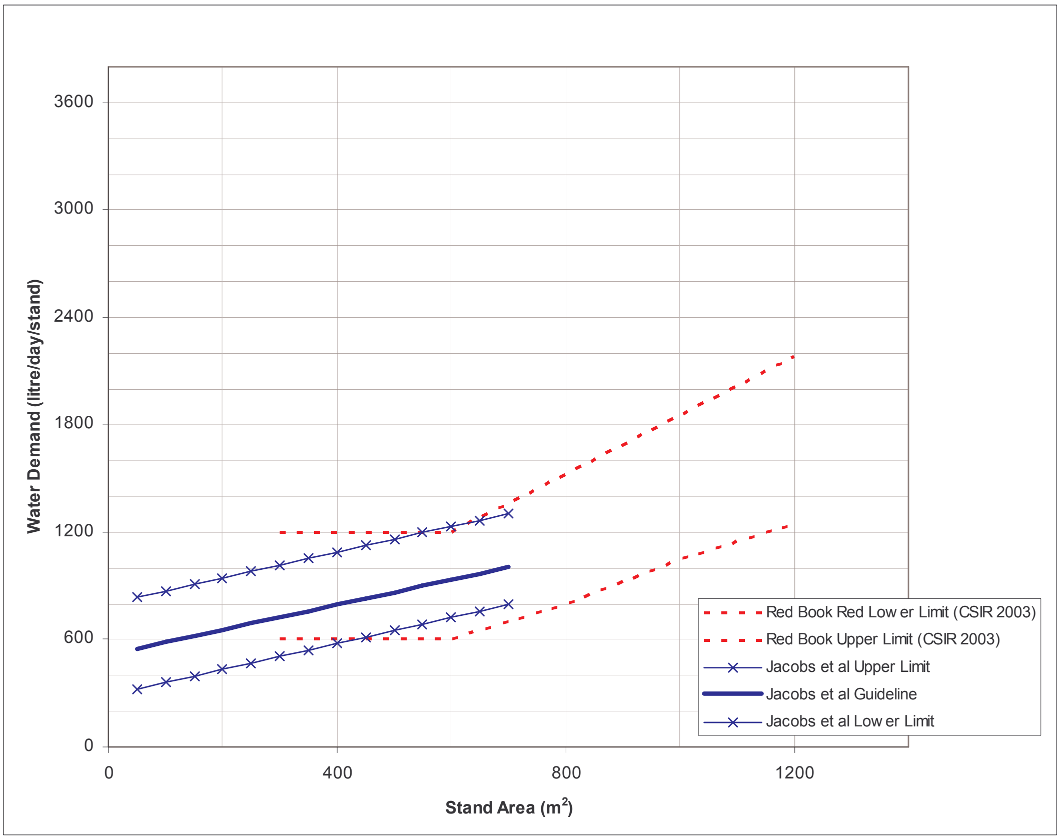


Figure 2-8: Southern Africa Inland Summer Rainfall Region Model – Townships (Jacobs et al., 2004)

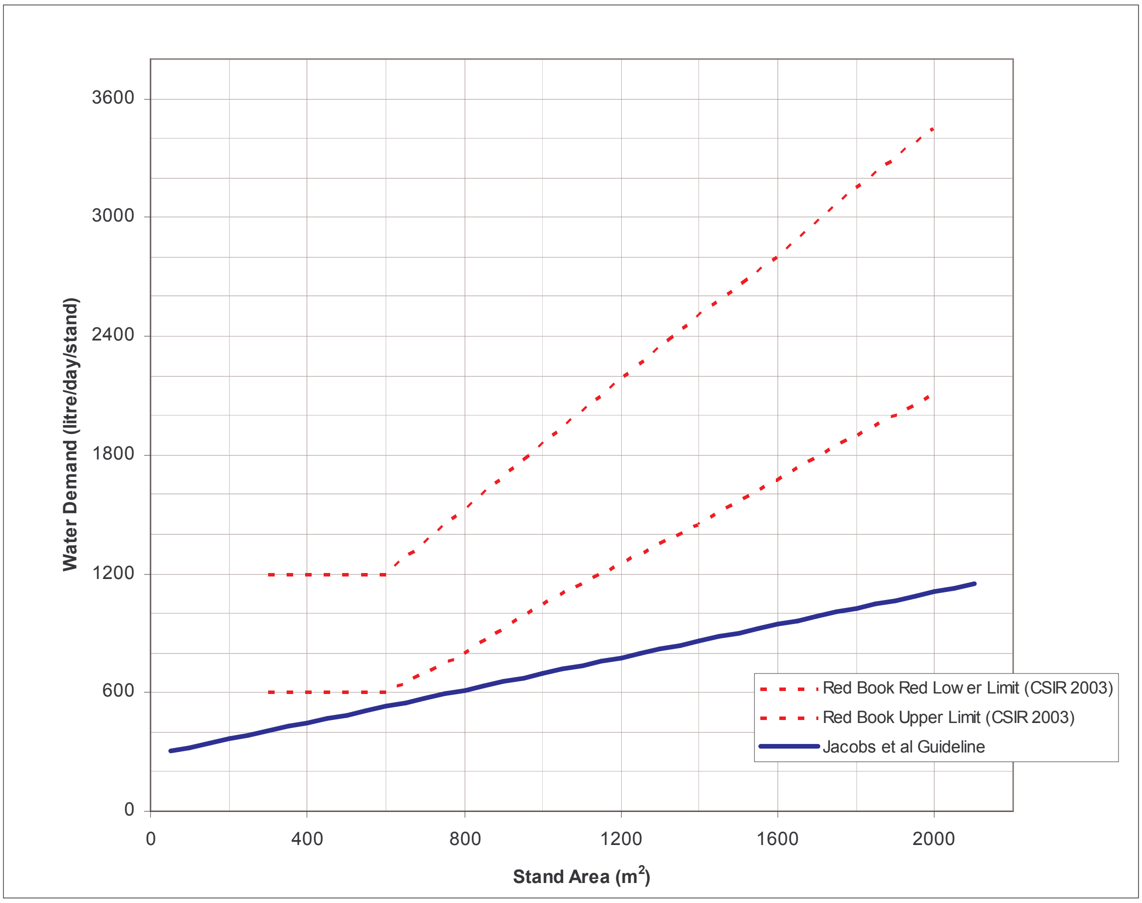


Figure 2-9: Southern Africa Coastal Annual Rainfall Region Model (Suburban & Townships) Restricted to George (Jacobs et al., 2004)

<u>*Highlights and Limitations of the Study*</u>

i. The study investigated the measured water consumption of *582 997* single residential stands in various towns and cities in South Africa using treasury databases. The large number of records made it possible to conduct statistical analyses and to investigate the distribution of data for stand area intervals of 100 m^2. This literature review indicates that this study is the first of its kind in South Africa, investigating such a large number of users for various regions in southern Africa.

ii. The investigation confirmed that a strong correlation exists between the annual average water demand for single residential stands and the stand area. Three single variable models are presented for three different geographic (climatic) regions in Southern Africa relating annual average residential demand to stand area.

iii. The study found the existing design guidelines (CSIR 2003) to be conservative.

iv. The research findings compare well with the results of previous studies undertaken in South Africa.

v. The single variable model does not address other parameters possibly influencing water demand apart from stand area. The designer has to take these parameters into consideration by selecting the appropriate demand within the given envelopes of the model. The study attempted to recognise the weather (climate) variable by generating models for three different climatic regions. However, the sample size for especially the coastal all-year rainfall region is limited to one town (George).

vi. The influence of socio-economic level is partly accounted for by presenting models for townships and suburban developments separately. However this distinction was only made for the Ekurhuleni dataset (i.e. the Southern Africa Inland Summer Rainfall region).

vii. The influence of weather was only investigated for a limited number of years of and in some cases one year.

2.2.7 Husselmann (2004)

In the study by Husselmann a subset of the database used in this study was evaluated to estimate the independent effects of stand area and stand value on average water demand. Data for 769 393 users were extracted for the following towns and cities:

Alberton, Atteridgeville, Benoni, Boksburg, Brakpan, Centurion, Edenvale, Kempton Park, Mamelodi, Northern Pretoria, Pretoria and Springs.

The data was filtered to exclude erroneous and unrealistic data points, as well as users falling outside the limits set for the study. The study was limited to residential users with stand areas between 200 and 1 400 m^2, and stand values between R 50 000 and R 245 000. Users with an average daily consumption of less than 0,01k*l*/day or more than 10k*l*/day were also excluded. After filtering, 194 816 users were left to base their study on.

Stand value (the sum of the municipal valuations of the value of the stand itself and improvements) was used as a surrogate for income. While no direct relationship between income and stand value could be determined, it was felt that stand value gave a good indication of income. It is also an internationally accepted practice to use stand value as a surrogate for income (Dandy et al 1997). Municipal stand valuations have not been updated to reflect the current market values, but are assigned on a reasonably consistent basis. It was thus assumed that the municipal valuations provided a good representation of stand values relative to each other.

The data set was categorised into six stand area and six stand value categories, each representing a roughly equal number of data points. The category boundaries used are given in Table 2.8. To reduce scatter, the data in each category was further grouped into 5 % intervals, thus producing 20 data points in each category.

Table 2-8: Stand Area and Stand Value Categories used in the study by Husselmann (2004)

Category no.	**Stand Area (m^2)**	**Stand value (R)**
1	200 – 260	50 000 - 65 000
2	260 – 300	65 000 - 85 000
3	300 – 380	85 000 - 130 000
4	380 – 760	130 000 - 170 000
5	760 - 1 000	170 000 - 200 000
6	1 000 - 1 400	200 000 - 245 000

Domestic Demand – the effect of Income and Stand Area

To investigate the effects of the stand area on AADD, the AADD was plotted against stand area for each stand value category. Since the stand values in each category are similar, the graphs represent the effect of stand area independent of stand value. A linear trend line was fitted to the data. A typical graph is shown in Figure 2.10 for the R 65 000 to R 85 000 stand value category.

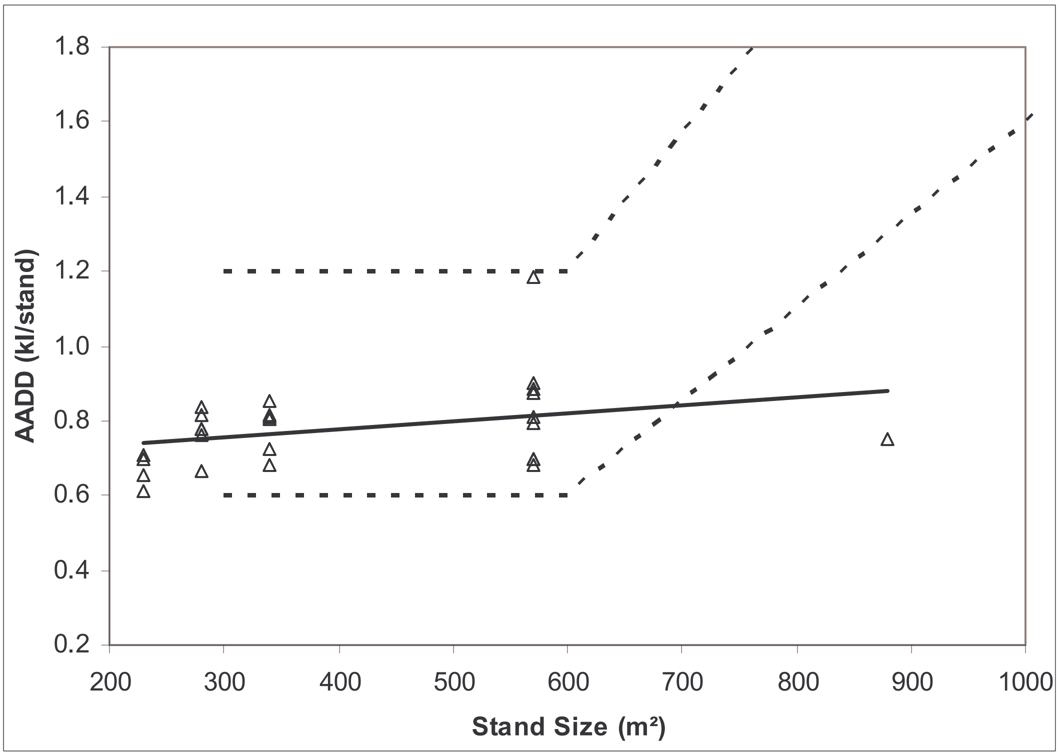


Figure 2-10: AADD vs. Stand Area for the R 65 000 to R85 000 Stand Value Category." Red Book" (CSIR,2003) is also shown. (Husselmann, 2004)

The trend lines for each of the six stand value categories were then plotted on the same axes as shown in Figure 2.11. The Red Book minimum and maximum guideline values are also shown. Figure 2.11 shows a clear relationship between AADD and stand area for all the stand value categories, with the AADD increasing with increasing stand area. The higher the stand value, the greater the AADD for the same stand area. This relationship can also be shown by plotting AADD against stand value for the different stand area categories. However, stand value is a subjective measure, likely to vary between municipal areas and with time. The authors thus felt that stand area was the better indicator of the demand envelope, with stand value (or income) playing a part in determining the relative position of a given supply area between the upper and lower bounds.

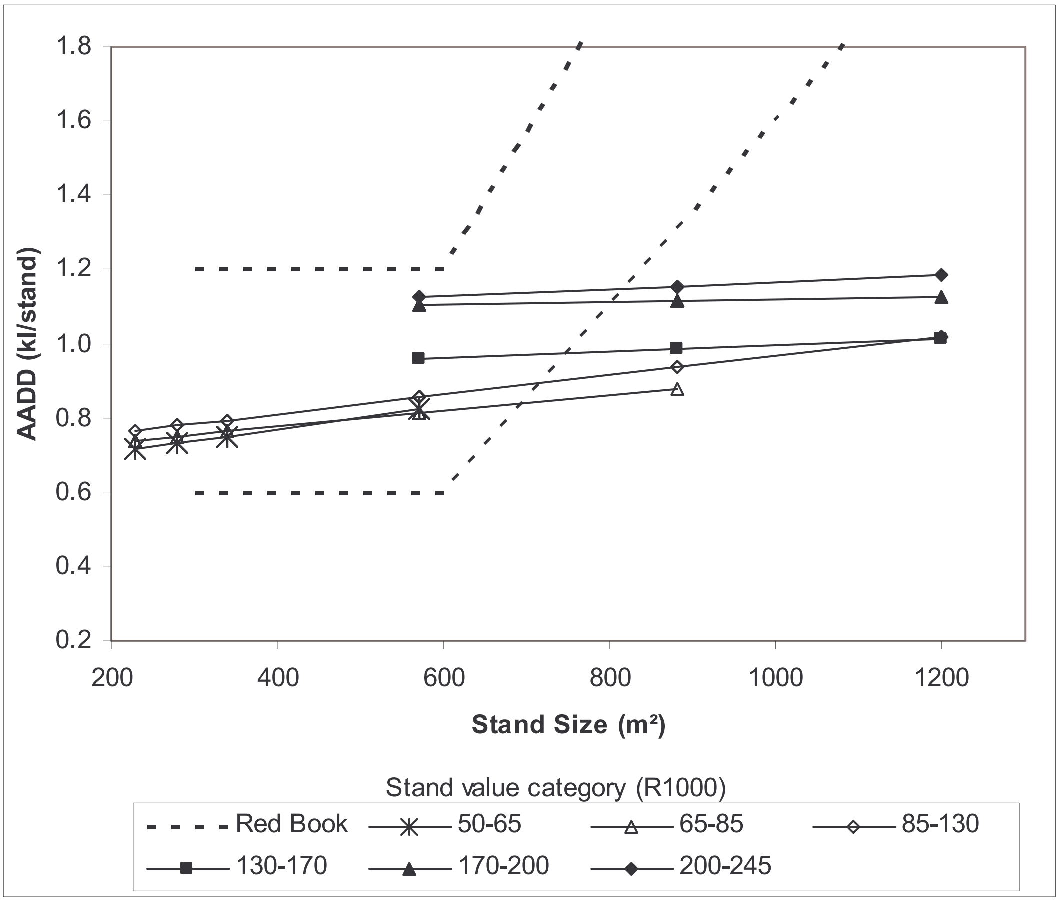


Figure 2-11: AADD as a Function of Stand Area for Different Stand Value Categories (Husselmann, 2004)

It can also be noted from Figure 2.11 that the higher the stand value, the smaller is the slope of the AADD vs. stand area trend line. The researchers therefore argued that this is directly opposed to the implicit assumption in the Red Book graph, which increases the slope for larger stand areas.

Finally, the authors indicated that while the AADD values for smaller stands fall within the "Red Book" guideline values, the "Red Book" possibly overestimated the AADD for larger stands. In this study, for stands larger than 800 m^2, the trend lines fell below even the minimum envelope line proposed by the Red Book.

The authors propose a new design envelope for AADD. To provide a better estimate of the AADD Husselmann plotted all the datapoints on which the lines in Figure 2.11 were based on a graph showing the Red Book Guideline envelopes (Figure 2.12). Husselman's new proposed envelope is also shown in Figure 2.12. This curve differs substantially from the Red Book graph in terms of the slopes and values of the envelope lines. The inflection points of the envelope lines remain at a stand area of 600 m^2, but their slopes decrease rather than increase. The proposed minimum

AADD limit is lower than that of the Red Book for all stand areas. The proposed maximum AADD limit exceeds that of the Red Book for stand areas between 300 and 700 m^2, but is lower for the rest. Significantly, its maximum line is lower than even the minimum Red Book line for stand areas exceeding 1 050 m^2. This represents a substantial reduction in design AADD for larger stands and will have a marked impact on the cost of providing water supply infrastructure for these areas.

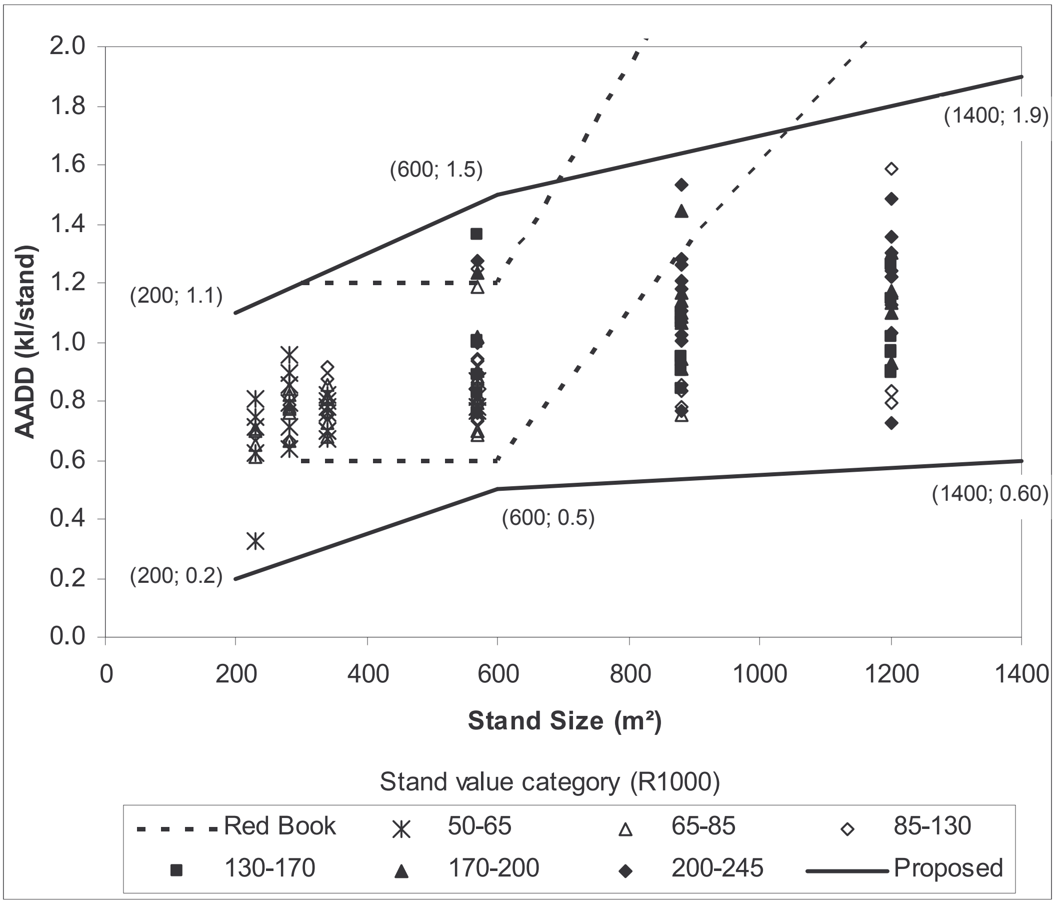


Figure 2-12: Proposed New Design Envelope for AADD showing data points and the Red Book Envelopes (Husselmann, 2004)

<u>*Highlights and Limitations of the Study*</u>

i. The study investigated measured consumption data for 769 393 residential stands in Gauteng.
ii. The data was divided into six stand area and six stand value categories to investigate the independent effects of stand value and size on water demand.
iii. The study found that there is a definite trend of increasing water demand with increases in both stand area and stand value (and thus income).
iv. It was concluded that stand area is a good measure for defining an AADD range. The design engineer should use information such as income, climate, level of living and other relevant

information to determine an appropriate AADD within the minimum and maximum range.

v. It was shown that increasing stand value increased the AADD (and thus moved it closer to the upper AADD range), and also reduced the slope of the curve.

vi. A comparison of the results with the proposed Red Book (CSIR 2003) guidelines for water demand showed that the Red Book guideline could underestimate the AADD in the 300 to 700 m^2 stand area range, but overestimate the AADD for stand areas larger than 700 m^2. A new envelope curve for estimating AADD is proposed based on the data analysed.

2.3 Summary of Major Unresolved Problems

i. The review of previous work done in South Africa indicates that the following are significant parameters that affect domestic water demand:

- Stand area
- Household income
- Water Price
- Available Pressure
- Type of development (suburban vs. township)
- Climate

These findings provided a base for the further work undertaken in this study.

ii. Some work has been done on the influence of climate. The study by Van Vuuren and Van Beek (1997) presented interesting findings regarding the combined effect of climate and income but was limited to the Pretoria supply area (one climatic region) and did not consider typical township developments. Jacobs et al (2004) considered the influence of climate on domestic water demand for three climatic regions but only with regards to stand area in a single variable model. Garlipp conducted a meticulous study on the effect of climate on domestic water demand, but considered cities as a whole (i.e. the water demand for a city was evaluated against climate). This study investigates the effect of climate for individual water consumers for various user categories in various types of developments (city vs. small towns) in various climatic regions in South Africa.

iii. Most of the previous work reviewed considered parameters influencing water demand individually. This literature review indicates that research of non-domestic demand patterns is very limited. This study investigates non-domestic water demand and the possible parameters influencing non-domestic water demand for a cross selection of South African towns and cities for various non-domestic user categories.

iv. Most of the studies considered with this review, were undertaken in the Gauteng area. Only the work by Jacobs et al (2004) considered different geographic regions in Southern Africa and the study by Garlipp (1979) considered other cities and regions in South Africa. However the study by Jacobs et al (2004) considered a single variable namely stand area. Although Garlipp's (1979) work is very valuable in this regard it was undertaken nearly 30 years ago and a lot has changed in the socio-economic and political characteristics of the country. This study investigates water users across South Africa in recent years (1998 – 2005) and considered a number of possible parameters influencing water demand.

v. Apart from the study of Jacobs et al (2004) that investigated nearly 600 000 domestic users country wide, the study by Van Zyl et al (2003) that investigated 110 000 domestic users and the study by Husselmann (2004) with nearly 800 000 users, the other studies investigated a limited number of users. A large number of records make it possible to conduct meticulous statistical analyses and to investigate the distribution of data in greater detail and to possibly have more representative samples of specific data characteristics. This study has extracted consumption data for more than 2.5 million users.

vi. The literature review indicates that the existing design guidelines (The "Red Book") (CSIR 2003) may be conservative (Jacobs et al,, 2004; Husselmann, 2004; Van Vuuren and Van Beek, 1997).

3 THE DATA

3.1 Introduction

Good quality data is always hard to come by and this is also the case with water consumption records. Even a large logging exercise, apart from being expensive, cannot realistically aim to reach a statistically reliable sample of households. This study considered municipal meter readings as an ideal source for water consumption information since the readings are taken on a regular basis by virtually all municipalities in the country. However, this information is generally difficult to access and are typically discarded (and thus lost) after a short period of time.

In recent years, GLS Consulting Engineers developed a software product called SWIFT®. This product allows the user to access municipal treasury databases to obtain demographic and water consumption information for large numbers of users (domestic and non-domestic). SWIFT® has been implemented by many local authorities throughout South Africa, covering different economic, socio-economic, climatic and other regions.

This study is based on water consumption data extracted from various SWIFT® databases developed for different municipalities throughout the country. The data reflects municipal water meter readings used for customer billing and thus also included errors present in these databases. Verification steps were taken to minimise the number of errors present in the analysis. The data cleaning process is discussed in detail in the following section.

Data on climatic and socio-economic parameters that possibly influence water consumption was sourced from the South African Weather Service and the South African Demarcation Board. The process of linking this data to the SWIFT® data sets is also described in this chapter.

3.2 Water Consumption Data

3.2.1 Data Collection

Municipal treasury data for a number of municipalities was collected in January to May 2005 from various SWIFT® databases operated by GLS Consulting Engineers. An on-site survey was required to establish which municipal treasury databases were available for extraction. Furthermore a complete understanding of the SWIFT® software was necessary in order to understand what Swift offered in terms of data. The data storing and primary data cleaning mechanisms had to be clearly understood to ensure that the final data extracted was what was needed for the study and that it was as accurate as possible.

The data extracted from the SWIFT® databases covers a time period of eight years (1998 – 2005) and consequently the data was available in various versions of SWIFT®. The automation process could not be altogether automated since the data was sourced from various municipalities that used different coding systems that had to be standardised in SWIFT®. Each database had to be thoroughly checked after extraction to ensure all the relevant data was included. This study made use of Microsoft Access ® to analyse and store the data. The water consumption and related data thus had to be exported from the Swift® databases in a readable format (PARADOX ®). Some SWIFT® data was very easily extracted but other information such as the municipal land use codes and the assignment of these codes to the standard SWIFT® land use codes was more complicated and required interpretation and checking after each extraction.

After extraction and importing the data into Microsoft Access® format a preliminary assessment of the data was done to ascertain whether there was any missing data for specific towns or municipalities. The reasons for the missing data were investigated. For example, is the treasury data simply not populated in the databases of is it available in another format or another database? If the missing data was not available an assessment was made on the impact this might have on the affected datasets and the subsequent analyses. If the impact was likely to be significant the affected datasets were discarded. For example if the land use codes were not populated, this would have a significant impact on further analysis and the data would thus be discarded.

The treasury data extracted with SWIFT® contained the following information per water meter that was used in the research:

- Unique identifier
- Municipality
- Town
- Stand area

- Stand Value (R)
- Value of Improvement on the stand (R)
- Land use code
- Suburb
- Zoning codes
- Water meter information (meter registration number, installation date, measurement unit)
- Flag whether vacant or not
- Flag whether large user or not
- Number of units (in the case of group housing)
- Water meter reading dates
- Monthly water meter readings
- Number of readings
- Calculated AADD (Annual Average Daily Demand) (kl/day)
- Adjustment codes (where data have been manipulated by SWIFT® procedures)

3.2.2 Description of the Data

Forty eight (48) municipal treasury databases were collected and extracted for further analysis in this study. This includes 4 metros (Johannesburg, Tshwane, Ekurhuleni and Cape Town) and 151 cities and towns. Table 3.1 lists the municipalities studied, the period of water consumption monitoring, and the total number of stands (uncleaned) for each municipality.

Table 3-1: Summary of Municipal Treasury Data Used.

Municipality	Towns and Cities	Period of Readings	# of Stands
Beaufort West	Beaufort West	Dec 02 - Feb 04	6 941
BergRiver	Veldrif, Porterville, Piketberg, Laaiplek, Aurora, Eendekuil, Dwarskersbos, Redelingshuys	Feb 03 - Mar 04	10 012
Blaauwberg	Blaauwberg, Milnerton, Mamre, Atlantis, Melkbos	Jun 00 - Jul 02	49 298
Breede River	Robertson, Bonnievale, Ashton, Montagu, McGregor	Oct 02 - Oct 04	20 079
Breede Valley	Worcester	Nov 99 - Jan 04	15 075
Buffalo City	East London	Jan 03 - Feb 05	119 748
Cape Agulhas	Bredasdorp, Waenhuiskrans, Struisbaai, Suiderstrand, Napier, Klipdale, Lagulhas	Mar 99 - Feb 04	8 576
Cederberg	Clanwilliam, Citrusdal, Lamberts Bay, Graafwater, Elands Bay, Leipoldtville	Dec 02 - Dec 03	6 881
Drakenstein	Wellington, Paarl, Gouda, Bainskloof, Hermon, Mbekweni	Dec01 - Dec 03	37 076
Ekurhuleni	Alberton	Aug 02 - Sep 04	40 065

Municipality	Towns and Cities	Period of Readings	# of Stands
Ekurhuleni	Germiston	Aug 02 - Sep 04	117 846
Ekurhuleni	Kempton Park	Aug 02 - Sep 04	87 779
Ekurhuleni	Lethabong	Aug 02 - Sep 04	25 150
Ekurhuleni	Nigel	Sep 02 - Oct 04	27 420
Ekurhuleni	Springs	Sep 02 - Oct 04	45 461
Ekurhuleni	Benoni	Sep 02 - Oct 04	100 921
Ekurhuleni	Boksburg	Aug 02 - Sep 04	72 334
Ekurhuleni	Brakpan	Sep 02 - Oct 04	63 812
George	George, Heroldsbay, Wildernis, Twee Rivieren	Jul 99 - Jul 03	33 229
Helderberg	Gordons Bay, Strand, Somerset West, Maccasar	Apr 98 - Apr 03	37 830
Johannesburg	Lethabong	Feb 02 - Mar 04	6 398
Johannesburg	Midrand	Feb 02 - Mar 04	39 699
Johannesburg	Northern Johannesburg (Sandton, Randburg,Bryanston, Sunninghill, Morningside)	Jan 02 - Feb 04	123 262
Johannesburg	Southern Johannesburg 1	Feb 02 - Mar 04	96 235
Johannesburg	Southern Johannesburg (Soweto)	Feb 02 - Mar 04	160 076
Johannesburg	Eastern Johannesburg	Feb 02 -Mar 04	118 315
Johannesburg	Western Johannesburg	Feb 02-Mar 04	105 494
Langeberg	Albertina, Gouritsmond, Heidelberg, Jongensfontein, Riversdale, Slangrivier, Stilbaai, Theronsville, Witsand	Dec 02 -Jan 04	16 008
Matzikama	Klawer, Lutzville, Vanrhynsdorp, Vredendal	Jan 02 - Feb 04	8 317
Mossel Bay	Mossel Bay, Dana Bay, Klein Brak, Hartenbos	Oct 01 - Oct 03	32 157
Oostenberg	Brackenfell, Blue Downs, Kuilsrivier	Nov 99 - Dec 01	61 300
Oudtshoorn	Oudtshoorn	Nov 01 - Nov 03	14 335
Overstrand	Gansbaai, Hermanus, Kleinmond, Stanford	Feb 00 -Feb 04	30 307
Plettenberg Bay	Plettenberg Bay	Dec 99 - Nov 04	9 155
Randfontein	Randfontein, Mohlakeng	Apr 01 - May 03	17 979
Saldanha Bay	Langebaan,Paternoster, Jacobsbaai, Vredenburg, St Helena Bay, Hopefield, Ongegund, Louwville	Apr 99 - Mar 04	23 481
Sedibeng	Sharpville, Sebokeng, Boipatong, Evaton, Stretford, Vereeniging, Meyerton, Kanana, Hartebeestfontein, Klipwater, Walkerville	June 02 - Jul 04	170 126
Stellenbosch	Stellenbosch	Oct 99 - Oct 01	14 335
Swartland	Abbotsdale, Chatsworth, Grottobaai, Darling, Malmesbury, Morreesburg, Koringberg, Riebeeck Wes, Riebeeck Kstl, Yzerfontein	Oct 01 - Oct 03	22 763
Theewatersklf	Caledon, Grabouw, Genadedal, Greyton, Riversonderend, Villiersdorp	Sep 03 -Oct 04	16 305
Tshwane	Atteridgeville	Feb 2002 - Feb 2005	17 753
Tshwane	Centurion	Dec 2002 -Jan 2005	35 652
Tshwane	Mamelodi	Feb 2002 - Feb 2005	38 493
Tshwane	NPMSS - Akasia, Rosslyn, Shoshanghuve	Feb 2003 -Mar 2005	84 900
Tshwane	Odi (Garankua)	Dec 2002 - Mar 2005	45 793
Tshwane	Pretoria	Feb 2002 -Feb 2005	146 691
Tshwane	Temba (Hammanskraal)	Jan 2004 - Feb 2005	12 108
Tygerberg	Belville, CapeTown, Delft, Durbanville, Goodwood, Mfuleni, Khayelitsha, Parow	Aug 2001 - Aug 03	196 036
TOTAL			**2 589 006**

Figure 3.1 portrays the location of each municipality that was studied and the major water regions of South Africa. The water regions were proposed by Basson et al. (1997) and published in a report by the Department of Water Affairs and Forestry. The regions were created by grouping together catchments of broadly similar hydrometeorological characteristics.

Figure 3-1: Geographic Locations of the Municipalities Studied

It is clear from Figure 3.1 that the municipalities investigated were fairly representative of the South Western, Southern Coastal and Northern water regions of the country. Only one municipality in the Eastern Coastal region and one municipality in the Central region were available for this study. Two water regions namely the Eastern Inland and the Karoo regions had no representation in the results of this study. However, since municipal treasury data was obtained using SWIFT®, this study was restricted to municipalities that have implemented Swift. This is one of the limitations of this study.

Figure 3.2 overleaf presents the monitoring period for each dataset. The different colours of the time bars in Figure 3.2 denote the water region for the municipality.

Table 3.2 below summarises the dataset characteristics per water region studied. Appendix A contains a comprehensive characterisation of each dataset that was used in the study.

Table 3-2: Summary of Dataset Characteristics per Water Region

Water Region (DWAF) (Basson et al., 1997)	**Municipalities**	**Number of Data Sets**	**Total Number of Stands**	**Total Number of Vacant Stands**	**Total Number of Domestic Stands**	**Total Number of Stands with Unknown Land use**
Central	Sedibeng	1	170 126	129 357	144 135	8 081
Eastern Coastal	Buffalo City (East London)	1	119 748	47 877	102 665	11 795
Northern	Ekurhuleni, Johannesburg Water, Randfontein, Tshwane	24	1 629 636	697 706	1 377 457	155 784
South Western	BergRiver, Blaauwberg, Breede River, Breede Valley, Cape Agulhas, Cederberg, Drakenstein, Helderberg, Matzikama, Oostenberg, Overstrand, Saldanha Bay, Stellenbosch, Swartland, Theewaterskloof, Tygerberg	16	557 671	157 165	457 613	38 888
Southern Coastal	Beaufort West, George, Langeberg, Mossel Bay, Oudtshoorn, Plettenberg Bay	6	111 825	33 472	68 685	15 895
TOTAL		**48**	**2 589 006**	**1 065 577**	**2 150 555**	**230 443**

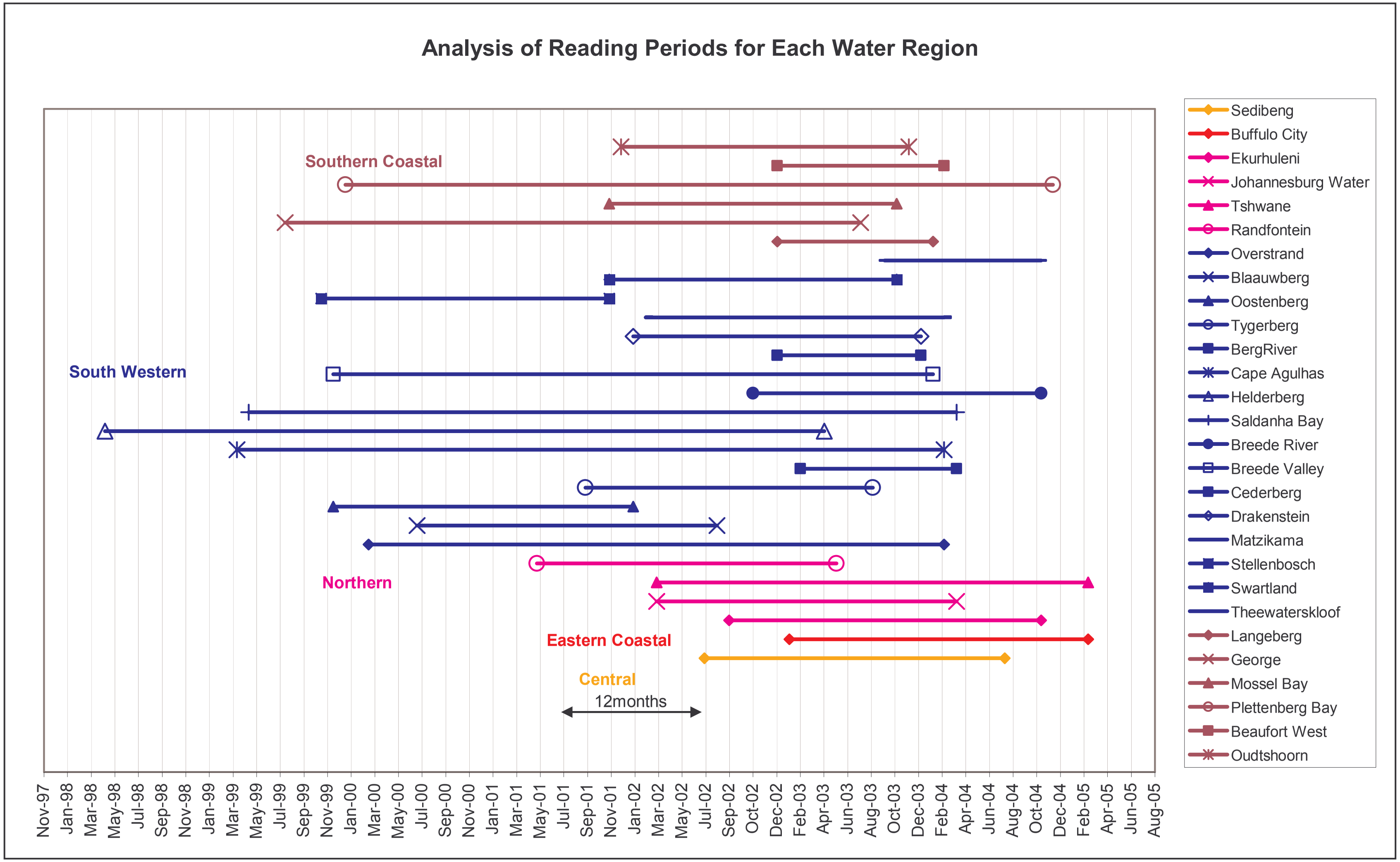


Figure 3-2: Monitoring Period of Each Dataset Used in the Study

The largest number of consumers investigated in this study was located in the Northern water region. This group of consumers were all located in Gauteng and include the three metros of Johannesburg, Tshwane and Ekurhuleni. The next largest group is situated in the South Western region and includes the metro of Cape Town. A large number of stands (41% of the total number) were assigned as 'vacant' stands in the treasury systems. Not all of these 'vacant' stands had zero water consumption; quite a number had relevant and regular readings. It may be possible that although these stands had at one time in the past been vacant, they had since been occupied, but the treasury systems had not been updated. However, it was considered prudent to exclude all stands that were flagged as vacant from further analysis even if some stands had relevant water consumption records. This process is described in the following section (3.2.3).

This investigation relied heavily on the accuracy of the land use codes assigned by the treasury systems. The land use codes were used to categorise the consumers in user type categories prior to analysis. Where no land use code has been assigned (8.9 % of the total number of stands) the data was of no use in this study and was discarded.

3.2.3 Data Verification

The data used in this study was obtained from municipal meter readings entered into treasury databases for billing purposes and then extracted using SWIFT®. It was necessary to verify the data to identify and eliminate potential errors in an objective manner. Various possible errors could be included in the database, namely:

- Wrong water meter readings;
- Errors in entering the meter readings into the treasury database;
- Errors in transferring the data to and from SWIFT®;
- Under- or over-registering of water meters;
- Meters not read regularly;
- Meter readings not taken on the same day of the month;
- Missing or erroneous user data, such as stand area, stand value or suburb;
- Incorrect land use assignment.

Primary Data Cleaning

The data collection process created a database for each municipality in Microsoft Access® format (48 databases). Before all the databases were consolidated into a single database, each municipal database was subjected to a primary data cleaning procedure. The aim of the primary data cleaning procedure was to eliminate records with obvious errors in the meter readings.

The primary data cleaning made use of the SWIFT® adjustment codes. SWIFT® assigns adjustment codes to every treasury record during the process of calculation of water demand and the consolidation of records. The adjustment codes describe the procedure of manipulation that was carried out on the record by SWIFT® in order to calculate an AADD value for the water meter. The primary data cleaning process used these adjustment codes as a guideline to disregard or omit records from further analysis. The table below lists the SWIFT® adjustment codes and provides a description of the meaning of the code (i.e. what action was carried out by SWIFT®). The table indicates whether records with specific adjustment codes were discarded in the primary data cleaning procedure or not.

Table 3-3: Primary Data Cleaning Procedure

Code No.	Code	Description	Discard Record?
1	None	No readings were found for the erf.	Yes
2	Single	A single reading was found for the erf.	Yes
3	Dates	Inconsistent dates. The system could not determine the order in which the readings should be placed due to date inconsistencies. Both the AADD and TWD are set to 0.	Yes
4	Clock	Clock overs. This is where the meters have clocked over. The system will manage this situation, given the display size of the meter.	No
5	New Meter	This is where the new meters have been installed. The system will handle the new meter from the date it receives the first new reading.	No
6	Spike	Spike interpolation. The system has detected a spike in the readings and will use interpolation to filter it out.	No
7	Dip	Dip interpolation. The system has detected a dip in the readings and will use interpolation to filter it out.	No
8	Estimate	Estimations. The system has detected that readings were estimated and will replace the estimated readings with a linear curve through the actual readings.	No
9	Duplicate	Duplicate dates. Two readings were found with different readings on the same date. The system will try each of the possibilities and select a suitable one.	No
10	Error	Reading Errors. The system was unable to determine the average water demand for the stand as the readings were totally inconsistent. Both the AADD and TWD are set to 0.	Yes
11	StartDate	Starting Dates. The starting date of the first meter reading is more than three months older than the Readings month specified in the "Calcs" tab.	No
13	RecDate	Inconsistent date recovery. The system found two readings roughly one year apart and Was able to determine an AADD for the plot. The other readings were discarded.	Yes
14	User	The user has adjusted the readings	No

Code No.	Code	Description	Discard Record?
15	TrimRec	Inconsistent reading recovery. The system found a non-consistent reading record, but recovered by trimming records to the first year.	No
17	ZERO UNITS, AADD > 0	Unit_Ts = 0 (number of living units connected with the record is zero in treasury data), Assumed Units = 1 for AADDPerUnit	No (need to be judged for each specific case)
20	PRE-PAID NOT FOUND	Prepaid record was not found in the stands table and hence the readings could not be updated.	Yes
21	PRE-PAID CODE, NO OVERLAP	Prepaid record was found in the stands table but the code did not match the code specified for prepaid records. The period of pre-paid readings did not overlap with the period of the meter readings.	Yes
22	PRE-PAID CODE, OVERLAP	Prepaid record was found in the stands table but the code did not match the code specified for prepaid records. The period of pre-paid readings overlapped with the period of the meter readings	Yes
30	REPLACED	The meter was identified as having been replaced by another meter. The 2 records were concatenated	No
31	OLAP CONS=0	The meter was identified as having been replaced by another meter, but the record periods overlapped. The consumption during the overlapping period = 0	No
32	OLAP CONS>0	The meter was identified as having been replaced by another meter, but the record periods overlapped. The consumption during the overlapping period > 0	No
33	Discarded	The meter was identified as having been replaced by another meter, but the record initial record was more than a user specified number of months old and was discarded.	Yes

The primary data cleaning procedure also ensured that meters with less 12 months of data were excluded from further analysis. Such values would only cover a part of the annual demand pattern and would thus be skewed towards a particular season. In addition, many of the errors, such as data reading, entering and transfer errors, irregular and missing meter readings, and readings not taken on the same day of the month will be substantially reduced by considering a longer consumption period. The next step in the primary data cleaning procedure was to omit all records with zero water consumption and to create a single database from the individual municipal databases in Microsoft Access®.

Secondary Data Cleaning

A secondary data cleaning procedure was carried out on the consolidated database. The aim of the secondary data cleaning procedure was to omit other potentially erroneous records, for example, records with relevant meter readings but flagged as “not metered” or as “vacant”. It could be considered that these records may be acceptable; however it was decided to omit all records with a question mark over them, considering the large number of records available. The secondary cleaning procedure is summarised in four steps:

Step 1: Omit all records flagged as not metered;

Step 2: Omit all records flagged as vacant;

Step 3: Omit all pre-paid metered records. The reason considered for this step is that pre-paid users have totally different consumption patterns than conventional users, which may complicate analysis and increase the data set variation;

Step 4: Remove all duplicate records from the dataset (i.e. records with the same unique identifier in the same municipality, same address but different AADDs). This included removing records where more than one type of land use occurred on a single stand.

The secondary data cleaning procedure was completed by adding the consumptions (AADDs) of all the meters on a stand and a final check to ensure that each record is unique, is a single stand with a single land use.

Consistent Land use Codes

The consumption data originated from various municipalities, each using their own coding system and extracted with various versions of the SWIFT® software. It was thus inevitable that the land use codes of the combined database created with the primary and secondary data cleaning procedures would not be consistent. All the various land use codes contained in the cleaned database were thus reviewed and assigned to the following standardised codes used in this study (Table 3.4):

Table 3-4: Standardised Land Use Codes Used

Land use Code	Description
RES	Residential
BUS_COMM	Business Commercial
IND	Industrial
FARM_AH	Farms
PARKS	Parks
SPORT	Sport facilities
EDU	Educational facilities
GOVT	Government and Institutional Facilities

It may be noted that "Farms" was defined as a land use category and it may seem unusual that farms exist within municipal boundaries. Furthermore, it is questionable whether commercial farms would use municipal water in their operations. However, most of the treasury databases contained "Farms" as a land use category. Judging by the stand sizes, most of these are actually agricultural small holdings, although some small towns especially in the Western Cape, did have significant sized farms within municipal boundaries. It was therefore decided to keep, as part of the analysis, this land use category as a type of use in its own right.

Domestic Group Housing

This study differentiated between single residential stands (i.e. single dwelling per stand) and domestic group housing. Domestic group housing is understood to be flats, townhouses, clusters and duplexes. An attempt was made to create a specific land use category for domestic group housing and to analyse water consumption patterns for this land use category. Group housing is a residential land use type that is increasing in present day South Africa, especially in metropolitan areas where space is limited and security is becoming an issue. It was anticipated that the water demand patterns for group housing would differ significantly from the more conventional single residential housing.

However, in extracting the consumption data from the municipal treasury systems it was noticed that where stands were assigned as group housing, a single meter could be measuring the consumption of the entire development (i.e. block(s) of flats) or a meter per unit could be measuring consumption (e.g. A cluster development with a meter per unit). There was a field in SWIFT® that indicated the number of units attached to a meter. This information was extracted by SWIFT® from the treasury systems and used. The calculated consumption value had to be thus divided with the number of units linked to the water meter for the group housing land use category. However, in analysing the results of this calculation, it was evident by the large number of group housing records, that either had extremely high or extremely low consumption levels, that the possibility existed, that the information on the number of units contained in the treasury systems might not always be very accurate. For this reason, it was considered prudent not to use the group housing data for this study. Nevertheless, there is a possibility to use this data in another study, if proper data cleaning and verification could be done, which was considered to fall outside the scope of this study.

It is however, understood that it was very probable that some group housing were still included in the final dataset that was analysed. This would be houses that were not identified as group housing in the treasury data and thus included. It is expected, that typically this would be cluster type houses or high security residential developments as these developments are mostly considered to be single residential by municipalities with a single meter installed per house.

3.3 Data on Parameters Influencing Water Consumption Patterns

3.3.1 Data Collection

Water demand patterns vary significantly from suburb to suburb, even in relatively cohesive areas such as Gauteng (Stephenson and Turner 1996). Various factors influence every user's decision on how much water he/she wants to use and when. Factors influencing water demand can be found in the answers to questions such as "Does the user have a garden?", "Is it a large garden?", "Has it been a dry summer?", and "Can the user afford the water?". However, to answer these questions for millions of users is an impossible task. The design engineer therefore is faced with the problem of estimating water demand and thus has to, in a sense, predict the user's decision to use or not to use water. Understanding the possible factors that may influence the user's decision on how much water they want to use has been a starting point of this study. The literature review undertaken as part of this study (Chapter 2) indicated that the following factors influence municipal domestic water demand patterns:

- Stand area
- Income level (stand value)
- Climate
- Water tariffs

In the literature review undertaken, very little information was presented as to the possible factors influencing municipal non-domestic water demand patterns.

Data was collected for stand area, income level, climate and other socio-economic factors such as level of unemployment, level of sanitation, household size, etc. At the start of this research, it was anticipated that tariff data in the treasury data would be used for the evaluation of water price on water demand. However, the tariff data extracted from the treasury databases was not as accurate or available as was hoped. For some municipalities tariff data was available but with no explanation of the coding system. For some other municipalities, codes were provided but the database not populated. Some had no data on tariffs at all. Furthermore, the treasury data were obtained in different time periods which would result in using 1998 tariffs for one data set and comparing them with 2005 tariffs in another data set. It was therefore not deemed prudent to use the tariff data that was available.

Table 3.5 indicates the sources of data on parameters influencing water demand.

Table 3-5: Data Sources for Parameters Influencing Water Demand

Factor	Source from where data was collected
Stand area	Treasury data containing records of water consumption and user characteristics for individual connections such as stand area
Income	Treasury data containing records of water consumption and user characteristics for individual connections such as stand value which was used as a proxy for income
Income	Socio-economic statistics per ward was obtained from the South African Municipal Demarcation Board. These statistics include level of unemployment, annual household income, level of sanitation, level of water service, household size
Climatic data (rainfall, average, maximum and minimum daily temperatures)	South African Weather Service
Mean Annual Precipitation Mean Annual Evaporation	Surface Water Resources of South Africa by Midgley et al. (1994)

It is important to observe that the stand value that was used in the analysis, was the sum of the municipal valuations of the value of the stand itself and any improvements made that increased the value of the stand. These values were all obtained from the treasury data.

3.3.2 Description of the Data

<u>*Climatic Data*</u>

The South African Weather Service (SAWS) was approached in June 2005 to supply climatic data for this study. Daily precipitation, daily minimum and maximum temperatures, humidity and evaporation data was requested for the period 01 January 1998 to 31 March 2005 for all the appropriate locations studied in this investigation.

The SAWS supplied the climatic data summarised in Table 3.6 for the time period Jan 1998 to March 2005. The SAWS could not supply data on humidity or evaporation since these measurements are not taken any longer by the service. It was hoped at the beginning of the research to have a higher resolution for climatic data, even as high as on suburb or group of suburbs level in the metropolitan areas. However, there were quite a number of municipalities where no data was available such as Oudtshoorn, Langeberg and Berg River, which cover several towns. The data that was available was evaluated and verified and incorporated in the water consumption database as discussed in the following section (3.3.3).

The WRC report by Midgley et al. (1994) containing extensive information on MAP and MAE was incorporated into the analysis to supplement the information supplied by SAWS. Appendix B contains tables indicating the locations and the MAP and MAE stations used.

Table 3-6: Climatic Data Supplied by the SAWS

Municipality	Rainfall Station Name	Temp Station Name
Alberton Ekurhuleni	ALBERTON	JOHANNESBURG INT WO
Atteridgeville Tshwane	PRETORIA EENDRACHT	PRETORIA EENDRACHT
Beaufort West	BEAUFORT-WES	BEAUFORT-WES
Benoni Ekurhuleni	BENONI - COUNTRY CLUB	SPRINGS
Berg River	not available	not available
Blaauwberg (Cape Town)	CAPE TOWN WO	CAPE TOWN WO
Boksburg Ekurhuleni	BOKSBURG - BRANDWEER	SPRINGS
Brakpan Ekurhuleni	SPRINGS	SPRINGS
Breede River	ROBERTSON	ROBERTSON
Breede Valley	WORCESTER-AWS	WORCESTER-AWS
Buffalo City	Umzoniana	not available
Cape Agulhas	not available	not available
Cederberg	not available	not available
Centurion Tshwane	PRETORIA EENDRACHT	PRETORIA EENDRACHT
Crocodile Tshwane	PRETORIA EENDRACHT	PRETORIA EENDRACHT

Municipality	Rainfall Station Name	Temp Station Name
Drakenstein	MALMESBURY	MALMESBURY
Eastern JHB	JOHANNESBURG INT WO	JOHANNESBURG INT WO
George	GEORGE WO	GEORGE WO
Germiston Ekurhuleni	ALBERTON	JOHANNESBURG INT WO
Helderberg	STRAND	STRAND
Kempton Ekurhuleni	BENONI - COUNTRY CLUB	SPRINGS
Langeberg	not available	not available
Lethabong Ekurhuleni	JOHANNESBURG SANDTON	JOHANNESBURG INT WO
Lethabong JHB	JOHANNESBURG SANDTON	JOHANNESBURG INT WO
Mamelodi Tshwane	PRETORIA EENDRACHT	PRETORIA EENDRACHT
Matzikama	VREDENDAL	VREDENDAL
Midrand JHB	JOHANNESBURG SANDTON	JOHANNESBURG INT WO
MosselBay	not available	not available
Nigel Ekurhuleni	SPRINGS	SPRINGS
Northern JHB	JOHANNESBURG SANDTON	JOHANNESBURG INT WO
NPMSS Tshwane	PRETORIA EENDRACHT	PRETORIA EENDRACHT
Odi Tshwane	PRETORIA EENDRACHT	PRETORIA EENDRACHT
Oostenberg (Cape Town)	CAPE TOWN WO	CAPE TOWN WO
Oudtshoorn	not available	not available
Overstrand	HERMANUS	HERMANUS
Plettenberg Bay	PLETTENBERGBAAI	PLETTENBERGBAAI
Pretoria Tshwane	PRETORIA EENDRACHT	PRETORIA EENDRACHT
Randfontein	JOHANNESBURG INT WO	JOHANNESBURG INT WO
Saldanha	LANGEBAANWEG WO	LANGEBAANWEG WO
Sedibeng	VEREENIGING	VEREENIGING
Southern1 JHB	JOHANNESBURG INT WO	JOHANNESBURG INT WO
Southern2 JHB	JOHANNESBURG INT WO	JOHANNESBURG INT WO
Springs Ekurhuleni	SPRINGS	SPRINGS
Stellenbosch	SOMERSET WEST	STRAND
Swartland	not available	not available
Temba Tshwane	PRETORIA EENDRACHT	PRETORIA EENDRACHT
Theewaterskloof	not available	ROBERTSON
Tygerberg (Cape Town)	CAPE TOWN WO	CAPE TOWN WO
Western JHB	JOHANNESBURG INT WO	JOHANNESBURG INT WO

Socio-Economic Data

The South African Demarcation Board was approached to supply socio-economic statistics in November 2005. The Board supplied all the data requested in electronic format. The following statistics based on the Census 2001 data was supplied for every political ward in South Africa:

- Ward identification
- Number of employed persons
- Number of unemployed persons
- Number of economically active persons
- Number of households living in a caravan or tent

- Number of households living in flats or in blocks of flats
- Number of households living in house or brick structures on a separate stand or yard
- Number of households living in houses/flats/rooms in back yard
- Number of households living in informal dwellings/shacks in back yard
- Number of households living in informal dwellings/shacks NOT in backyard
- Number of households living in private ships/boats
- Number of households living in rooms/flatlets not in back yard but on shared property
- Number of households living in town/cluster/semi-detached houses (simplex, duplex, triplex)
- Number of households living in traditional dwellings/huts/structures made of traditional materials
- Number of households for various intervals of household sizes
- Number of Dwellings (houses) for various intervals of house sizes
- Number of households with no income
- Number of institutions where income is not applicable
- Number of households for various intervals of annual income
- Number of households with no access to piped potable water
- Number of households with access to piped potable water to their dwelling
- Number of households with access to piped potable water to their yard
- Number of households with access to piped potable water more than 200 m from their dwelling
- Number of households with access to piped potable water less than 200 m from their dwelling
- Number of households with no sanitation
- Number of households with bucket latrine sanitation
- Number of households with chemical toilet sanitation
- Number of households with flush sewerage sanitation
- Number of households with flush septic tank sanitation
- Number of households with pit latrine ventilated sanitation
- Number of households with Pit latrine no ventilation sanitation

3.3.3 Data Verification and Linking to Water Consumption Data

Climatic Data

The SAWS provided the available climatic data (precipitation and minimum and maximum temperatures)) for every day for the time period 01 January 1998 to 31 March 2005. The data was provided in text format ('.txt" files) and thus had to be incorporated in Microsoft Access® format. For both the rainfall and temperature datasets, there were a number of days on which measurements were not taken or missing or estimated. The data was evaluated and if the impact of the missing data was considered to be significant the dataset for the month was discarded. The monthly averages were calculated for each available measurement station for both precipitation and temperature. Where estimates or missing data were included in the calculated monthly average, the calculated value was flagged as a possibly unreliable record.

The climatic data was linked to the water consumption database on the basis of municipality. Section 3.2.2 discusses why a higher resolution such as town or suburb could not be used. To link the climatic data to the water consumption database, the time period for the majority readings for a municipality had to be determined. Table 3.1 gives the time period for the majority of readings per municipality. For each municipality, the average annual rainfall and average annual maximum and minimum temperature for the period of consumption measurement were determined from the data of the closest climatic measurement station where available.

The published MAP and MAE values (Midgley et al 1994) were linked to the water consumption database by first establishing the most appropriate MAP and MAE stations for each town and where possible suburb or group of suburbs in metropolitan areas. This was done by using the longitude and latitude information of the climatic stations. The MAP and MAE values were captured manually into the water consumption database by creating a town /suburb table with the relevant corresponding MAP and MAE values and then linking the data automatically to the larger water consumption database.

Socio-Economic Data

From the socio-economic data provided by the South African Demarcation Board (Section 3.2.2) the following averages were calculated per political ward:

- % Unemployment

- % Formal Housing
- Average Household size (weighted average)
- Average House size (weighted average)
- Average Income (weighted average)
- % of dwellings with potable water house connection
- % of dwellings with waterborne sanitation

To link the socio-economic data to the water consumption database was a complicated process. The main problem being that the socio-economic data was made available by the South African Demarcation Board on the basis of political wards. The water consumption data obtained from the municipal treasury systems was based on suburbs. The wards thus had to be superimposed onto the suburbs as accurately as possible. In other words, for each suburb, it had to be decided in which political ward it fell. As the socio-economic database only had ward numbers and special place names for some wards this process was not straightforward at all. Use was made of mapping software provided by the South African Demarcation Board named SA Explorer®. SA Explorer® has a GIS interface (Arcview) and enables the user to view the ward boundaries in relation to major roads, rivers and places. The process of linking the socio-economic data to the water consumption database was thus a manual process. For approximately 13% of the suburbs, it could not be established in which ward they fell. This was mainly because there was limited information on special place names in the wards which could be linked to a suburb or because the suburb fell into two different wards. Another problem encountered with the socio-economic data being only available on the political ward level was that in some wards, high income and low income suburbs or areas had be grouped together in one ward. In such a case, one would not expect the water demand to be homogenous. However, these cases were definitely in the minority.

It should also be noted that the socio-economic data was only linked for domestic consumers, as it was considered that the socio-economic data such as income would have little or no influence on non-domestic use.

4 METHODOLOGY

4.1 Data Filtering

The data cleaning and verification procedure created a single database including data on water consumption and data on parameters that possibly affect water demand (climate and socio-economic data). The database was spilt into 25 separate databases, each database representing a land use type. There were 7 non-domestic user type databases and 18 domestic user type databases. The 18 domestic user databases were created from the domestic data broken down into similar stand areas (creating 8 databases) and similar stand values (creating 10 databases).. Each database comprised some or all the following fields:

- A unique identifier that can be linked back to the treasury data and address and other customer and meter information;
- Land use code;
- Suburb Category that can be linked to a suburb;
- Annual average daily demand (AADD) in kl/day;
- Period of calculation for the AADD (dates);
- Stand Area (m^2);
- Stand Value (R);
- Geographic location of the stand (either “Coastal” or “Inland”) The geographic location was given a numeric value namely Coastal = 1 and Inland = 2;
- Development Level (either City/ Metro or small town/rural) (This field was only included for the non-domestic users as it was expected that the development level may have a significant influence on non-domestic use i.e. larger industries, educational facilities etc in the cities). The development level was also assigned a numeric value, namely City/Metro = 1 and Rural/Small town = 2;
- Mean Annual Precipitation (mm) for the area where the stand is located;
- Mean Annual Evaporation (mm) for the area where the stand is located;
- Average Maximum Temperature for the water consumption measurement period (degree Celsius);

- Average Minimum Temperature for the water consumption measurement period (degree Celsius);
- Unique political ward ID as obtained from the South African Demarcation Board *;
- Percentage of unemployed persons in the economically active population*;
- Percentage of households with formal housing (not informal or traditional dwellings)*;
- Average household size (no. of persons)*;
- Average house size (no. of rooms)*;
- Average household income*;
- Percentage of houses with potable water house connection (i.e. not a stand pipe etc)*;
- Percentage dwellings/houses with water borne sanitation (i.e. flush toilets)*.

Note: **data only available on ward level and only included in the domestic databases as these factors were not considered to have an impact on non-domestic use*

The next step involved selecting only those stands that reasonably represent each land use category. Filters were thus applied to the databases to exclude users with unrealistically low or high stand areas and stand values. These values may not necessarily be wrong, but were omitted on the basis that it is not practical to supply every possible type of consumer through the municipal water distribution system. Table 4.1 summarises the criteria used for the data filters. The cut-off points were decided by pragmatic consideration judging the data range in each case. From the stand area and stand value ranges shown in Table 4.1, every effort was made to retain as many records as was possible.

Table 4-1: Data Filters Applied to Water Consumption Data

Land Use Category	Stand area (m^2)	Stand Value (R)
Single residential	>20 m^2 and <= 4000m^2	>R20,000 and <=R10 million
Business Commercial	>20 m^2 and <= 50 000m^2	>R20,000 and <=R100 million
Industrial	>20 m^2 and <= 50 000m^2	>R20,000 and <=R100 million
Farms	>5000m^2	>R20,000 and <=R10 million
Sport and Parks	>500 m^2 and <= 100 000m^2	>R20,000 and <=R20 million
Education	>100 m^2 and <= 75 000m^2	>R20,000 and <=R20 million
Government and Institutional	>100 m^2 and <= 50 000m^2	>R20,000 and <=R20 million

It may seem that the lower cut-off value for the stand value attribute is unreasonably low for all the land use categories. However it has to be borne in mind that the data for the stand values originated from treasury systems that may not have been updated for many years. Using a higher cut-off value would have resulted in losing a large number of records. Table 4.2 below provides an indication of the distribution of the number of domestic users for various stand value intervals. It is clear that the majority of stands had values less than R 500 000 in the treasury data. It seemed therefore, that the majority of the data might have outdated or lower than expected stand values. It could be argued that using the outdated stand values in combination with more up to date stand values in the data analysis would only result in significant unexplained variation in the final results. Nevertheless, it was decided to retain the available stand value data and review the final results with this fact in mind.

Table 4-2: Final Sizes of the Domestic Datasets for Categories of Stand Value

Dataset Name	Dataset Description	Number of records (stands)
RESVAL50000	Single residential stands with a stand value of greater than or equal to **R 20 000** and smaller than **R 50 000**	211 753
RESVAL100000	Single residential stands with a stand value of greater than or equal to **R 50 000** and smaller than **R 100 000**	200 322
RESVAL250000	Single residential stands with a stand value of greater than or equal to **R 100 000** and smaller than **R 250 000**	298 703
RESVAL500000	Single residential stands with a stand value of greater than or equal to **R 250 000** and smaller than **R 500 000**	227 595
RESVAL750000	Single residential stands with a stand value of greater than or equal to **R 500 000** and smaller than **R 750 000**	49 807
RESVAL1000000	Single residential stands with a stand value of greater than or equal to **R 750 000** and smaller than **R 1 000 000**	14 351
RESVAL2000000	Single residential stands with a stand value of greater than or equal to **R 1 000 000** and smaller than **R 2 000 000**	8 704
RESVAL3000000	Single residential stands with a stand value of greater than or equal to **R 2 000 000** and smaller than **R 3 000 000**	1 286
RESVAL4000000	Single residential stands with a stand value of greater than or equal to **R 3 000 000** and smaller than **R 4 000 000**	428
RESGreaterThan4000000	Single residential stands with a stand value of greater than or equal to **R 4 000 000**	1 275

Dataset Name	Dataset Description	Number of records (stands)
TOTAL		**1 014 224**

Table 4.3 lists the number of records in each domestic stand area category dataset after the data cleaning processes and after the data filters were applied. The final numbers also exclude records where any one or more of the fields (or parameters) were not populated. For example for some records it could not be established in which ward they fell and subsequently the socio-economic data could not be populated and thus these records had to be excluded from further analysis.

Table 4-3: Final Sizes of the Domestic Datasets for Categories of Stand Area

Dataset Name	Dataset Description	Number of Records (stands)
RES500	Single residential stands with stand areas of greater than or equal to **20** m^2 and smaller than **500** m^2	433 104
RES750	Single residential stands with stand areas of greater than or equal to **500** m^2 and smaller than **750** m^2	112 443
RES1000	Single residential stands with stand areas of greater than or equal to **750** m^2 and smaller than **1 000** m^2	136 019
RES1500	Single residential stands with stand areas of greater than or equal to **1 000** m^2 and smaller than **1 500** m^2	184 497
RES2000	Single residential stands with stand areas of greater than or equal to **1 500** m^2 and smaller than **2 000** m^2	49 952
RES2500	Single residential stands with stand areas of greater than or equal to **2 000** m^2 and smaller than **2 500** m^2	17 506
RES3000	Single residential stands with stand areas of greater than or equal to **2 500** m^2 and smaller than **3 000** m^2	6 884
RES4000	Single residential stands with stand areas of greater than or equal to **3 000** m^2 and smaller than **4 000** m^2	6 168
TOTAL		**946 573**

Table 4.4 lists the number of records in each non-domestic dataset after the data cleaning processes and after the data filters were applied:

Table 4-4: Final Sizes of the Non-Domestic Datasets.

Data Set Name	Number of records
Business Commercial	24 810
Education	1 447
Farms	8 895
Government & Institutions	8 259
Industrial	8 000
Parks	367
Sport	68
Total	**51 846**

The number of records in the Sport category was very small and it was expected that no statistical significant results would be obtained for this category.

4.2 Data Analysis and Demand Estimation

4.2.1 Domestic Water Consumption – Assessment of the Existing Design Guidelines

The first step in the data analysis was a direct comparison of the domestic data with the "Red Book" (CSIR, 2003) demand estimate guideline for domestic use. To do this comparison, it had to be assumed that the climatic and socio-economic factors in a suburb were fairly homogenous. The average stand area and the average AADD for a suburb were thus determined and these values were plotted against the upper and lower limits given as a demand estimate guideline in the "Red Book" (CSIR, 2003).

A single variable regression analysis was also done with ln(StandArea) specified as the independent variable and ln(AADD) as the dependent variable using all the domestic data (i.e. not categorising the data in datasets of different intervals for stand area). SPS statistical analysis software was used to perform this regression analysis with 1 091 685 domestic records countrywide.

The assessment of the Red Book design guidelines were further refined by categorising the data (1 091 685 records) in two categories for geographic location namely coastal or inland. A single variable regression analysis was performed with stand area specified as the independent variable for each category. The definition of 'coastal' used was a town or area directly adjacent to the sea. For example the geographic location of Lamberts Bay (0 km from the coast) would be "coastal" and Vredendal (approximately 50 km from the coast) would be "inland".

4.2.2 Assessment of Factors Influencing Domestic Water Consumption

The next step was to also perform a step-wise multiple variable regression analysis for each of the domestic stand area datasets with ln(AADD) specified as the dependent variable to determine which factors correlates best with domestic consumption. A stepwise regression for more than one independent variable will always select the variable that correlates best with the dependent variable first and then the next best correlation and so forth. Thus multiple variable models were obtained for various stand area and stand value categories of domestic demand. This provided models of the combined effect of the parameters influencing domestic water demand. The regression statistics for each of the regressions are discussed in detail in Chapter 5.

The stepwise regression models informed the decision as to which two variables influence water demand the most in each land use category. In most cases, it was the ln(StandValue) and the ln(StandArea). Thus for these two parameters, single variable regression models were calculated with 95 % confidence intervals for each land use category. The single variable regression models were further refined by distinguishing between the geographic location (e.g. inland or coastal). Figure 4.1 portrays schematically the number of single variable regression models calculated for domestic user categories. Thus for every domestic user category 4 single variable regression models were created and with 95 % confidence limits for the mean of the AADD. The regression results of each of these models are discussed in detail in Chapter 5.

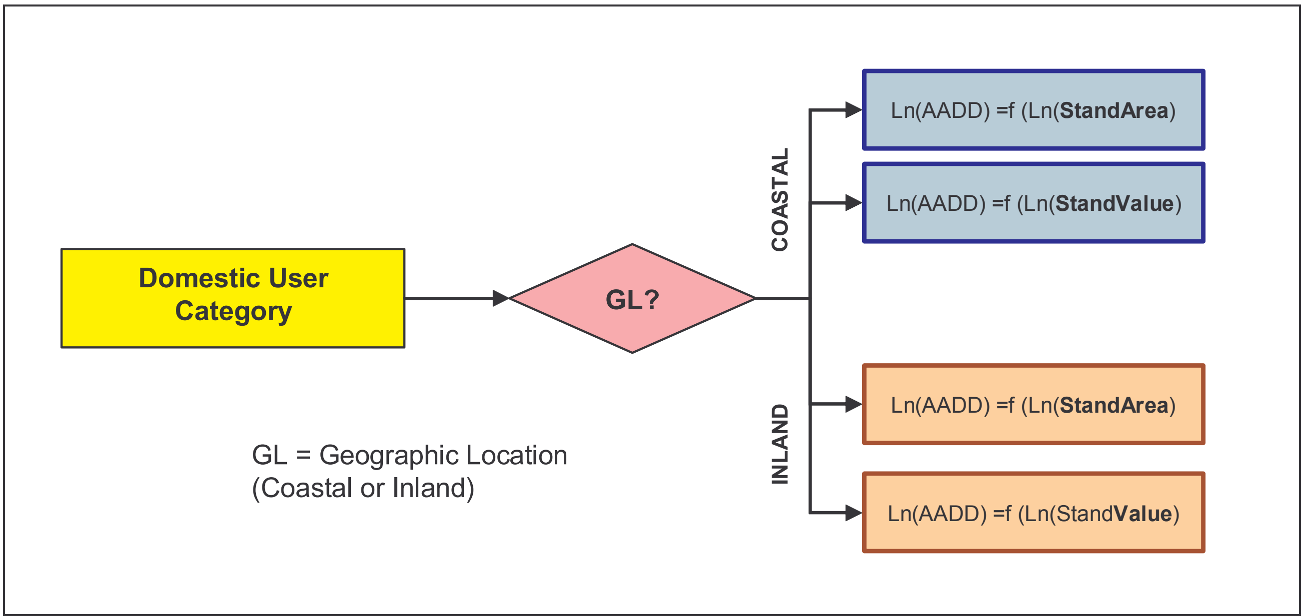


Figure 4-1: Single Variable Regression Models for Domestic User Categories

4.2.3 Proposed Revised Domestic Demand Estimation Guideline

Following the outcome of the multiple variable and single variable regression analyses it was considered constructive to generate a curve that could serve as an indication of how a revised guideline for domestic demand estimation might look if the results of this study were used. A single curve indicating AADD for various stand areas was determined with the 25%, 75% and 95% confidence intervals for the data that could serve as design envelopes. In using this curve, a user would have to interpret the influence of income.

4.2.4 Non-Domestic Water Consumption

The first step was to calculate frequency distributions and descriptive statistics for each of the datasets to better understand the nature of each dataset. In the literature review undertaken with this study very little previous work was found on water consumption estimation of non-domestic users for municipal services design purposes. Therefore this step was considered necessary. The frequency distributions were calculated for intervals of the natural logarithm of the demand (AADD).

The next step was to perform a step-wise multiple variable regression analysis for each land use dataset with ln(AADD) specified as the dependent variable. By doing a stepwise regression it could be established which variables (or parameters) influences non-domestic water demand the most and also obtain a model of the combined effect of the parameters influencing water demand. The regression statistics for each of the regressions are discussed in detail in Chapter 5.

As was the case with the domestic analysis the stepwise regression models informed the decision as to which two variables influences water demand the most in each land use category and consequently single variable regression models with 95 % confidence intervals were also calculated with stand area and stand value specified as independent variables for each land use category. The single variable regression models were further refined by distinguishing between the geographic location (e.g. inland or coastal) and the development level (e.g. metro and city or small town rural).

Figure 4.2 portrays schematically the number of single variable regression models calculated for non-domestic user categories . Thus for every non-domestic user category 8 single variable regression models with 95 % confidence limits for the mean of the AADD were calculated. The regression results of each of these models are discussed in detail in Chapter 5.

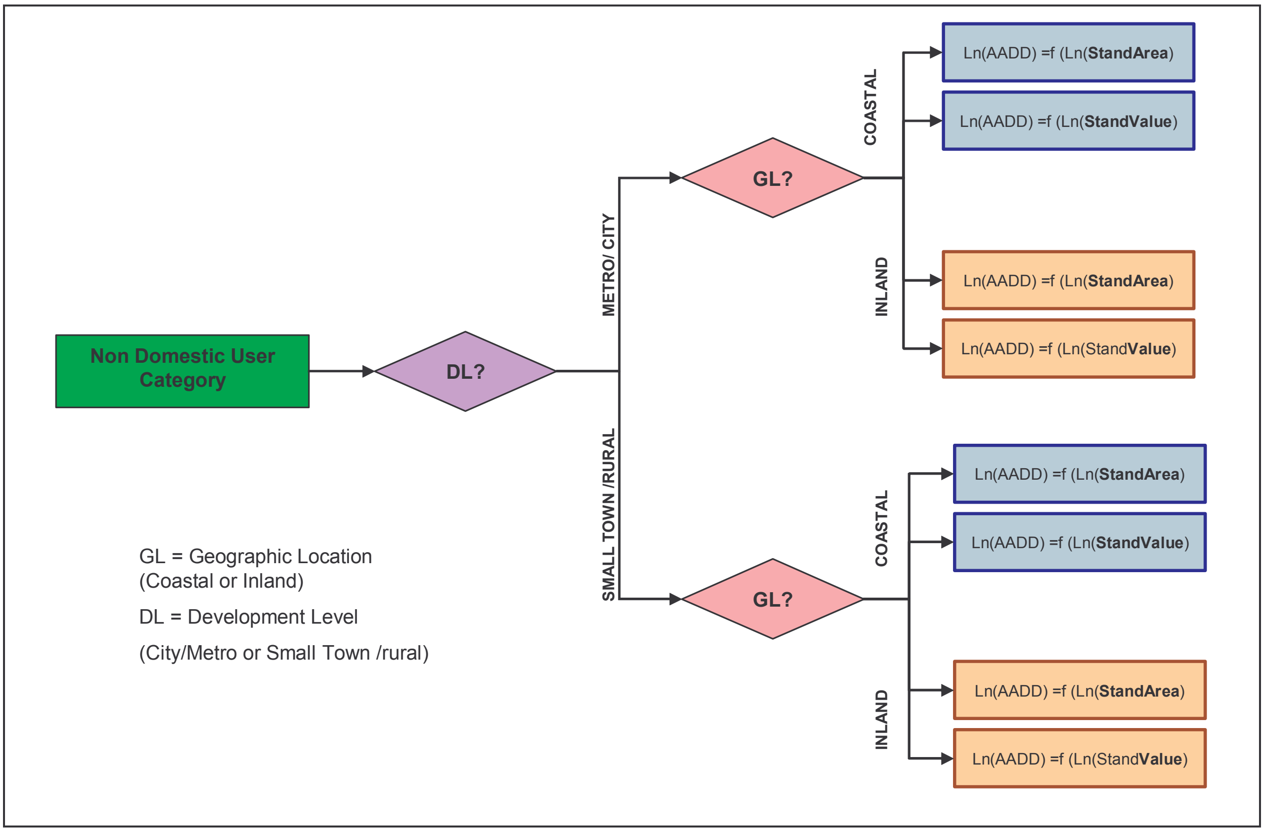


Figure 4-2: Single Variable Regression Models for Non-Domestic User Categories

4.2.5 Proposed Non-Domestic Demand Estimation Guideline

The existing demand estimation guidelines for non-domestic water use are limited to specific types of non-domestic use and are usually based on the floor area of the non-domestic development. The floor area of a non-domestic development usually has to be estimated or guessed by the design engineer. It was considered that an ideal guideline would be one where no such additional information of the attributes of the non-domestic users were required, as it is generally very difficult to obtain such data. It was therefore considered constructive to present a possible revision to the existing non-domestic demand estimation guidelines for the user categories defined in this study by using the available data.

A possibility investigated with this study, was to use the 95% tolerance limit for the mean of the log normal distribution that was fitted to the population of each non-domestic user category. In other words, the independent variable for the guideline curve would be the number of users. The curve would give the 95% confidence limits for the average water demand of X number of users. For the purposes of demand estimation it is proposed to use only the upper 95% tolerance limit. The design engineer would have to select or know the number of non-domestic users of the specific user category and would read the estimated AADD from the guideline curve. The curve has an asymptote in the mean of the log normal distribution.

5 RESULTS

5.1 Domestic Water Consumption

5.1.1 Assessment of the Existing Design Guidelines

The first step in evaluating domestic demand was to make a direct comparison with the "Red Book' (CSIR, 2003) guidelines for estimating domestic water demand. This constituted an assessment of the design guidelines using all the available data for domestic users (i.e. not categorising the data) and considering stand area as the only variable influencing water demand. The cleaned data for all the domestic users was grouped together by suburb. Suburbs with less than 50 stands were excluded from this analysis. In this analysis, there were 1188 suburbs represented in various cities and towns across South Africa. An average AADD value and an average stand area value per suburb were calculated. Figure 5.1 depicts the outcome of this preliminary assessment by plotting the suburb averages with the upper and lower envelope curves of the "Red Book" (CSIR, 2003).

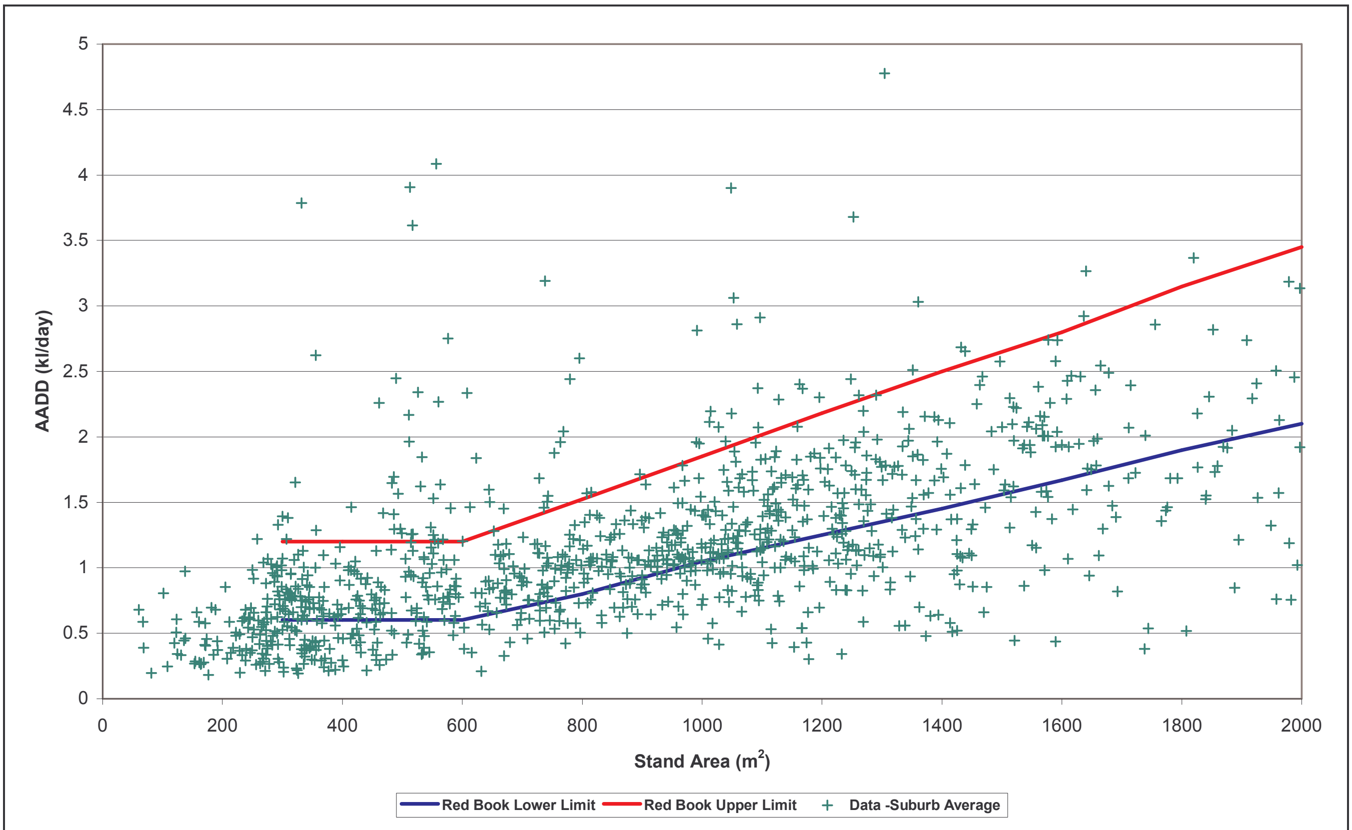


Figure 5-1: Preliminary Assessment of the Existing Demand Estimation Guidelines

From Figure 5.1 it can be seen that the majority of the suburb averages fell well within the two envelope curves of the design guideline. Of the total 1188 data points 461 (38.8%) fell below the lower envelope curve and *100* (8.4 %) above the upper envelope curve. Thus a total of *47.2 %* of the data points fell outside of the envelope curves of the design guideline.

An analysis of the suburbs that fell outside the upper guideline curve was made and found that 57% of the data points represented suburbs with average stand areas smaller than 800m^2. The suburbs that had consumption values above the upper guideline curve were mostly a combination of middle income, smaller single residential developments such as Cason in Boksburg, Brixton and Yeoville in Johannesburg and low income or township developments such as Orange Grove (Johannesburg) and Ncala in Germiston. However, a number of high income, high security developments such as Kyalami Estates, Dainfern, Douglasdale and Sunning Hill in the northern suburbs were also found among the suburbs with higher water consumptions. Approximately 23 of the 100 suburbs, with greater water consumption than predicted by the "Red Book", could be classified as high income and typically high security developments.

The single variable regression analysis was done for all 1 091 685 domestic stands countrywide, with stand area specified as the independent variable. The format of the regression model that was used with 97.5% confidence intervals for the average, is given below:

$$y = a + bx \pm (1.96)s\sqrt{\frac{1}{N} + \frac{(x-\bar{x})^2}{(N+1)(\sigma)}}$$

Where
a and b = regression coefficients
s = standard error of the estimate
N= number of datapoints
σ = variance

The following regression model was the result of this analysis:

$$\ln(AADD) = -1.610 + 0.297\ln(StdArea) \pm 0.860\sqrt{9.16\times10^{-7} + \frac{(\ln StdArea - 6.4124)^2}{666977}} \qquad (5.1)$$

Where

StdArea = Stand Area in m^2

AADD = Annual Average Daily Demand (kl/day)

The first part of the equation (before ±) describes the average water demand curve, and the second part the 97.5 % confidence interval. The regression model had an adjusted R square value (square of the Pearson product moment correlation coefficient) of 0.218, which implied that 21.8 % of the variability in the data could be explained by this equation. An adjusted R square value of more than 20 % is considered good when predicting human behaviour as was the case with this study. Numerous factors contribute to the variability in the data and given the large dataset, even greater variation could be expected. The t-statistics for the regression coefficients indicated that they were significantly different from zero. The detailed regression statistics are included in Appendix C. A logarithmic regression model was used as this gave the best fit to the data. Figure 5.2 below plots the regression equation against the Red Book design guideline envelope curves for domestic water use. The guideline curves of the Red Book were extrapolated up to stand areas of 4 000 m^2 in order to make the comparison. This study had more than 30 000 residential stands with stand areas greater than 2000 m^2. The Red Book only gives the guideline curves up to a stand area of 2 000m^2.

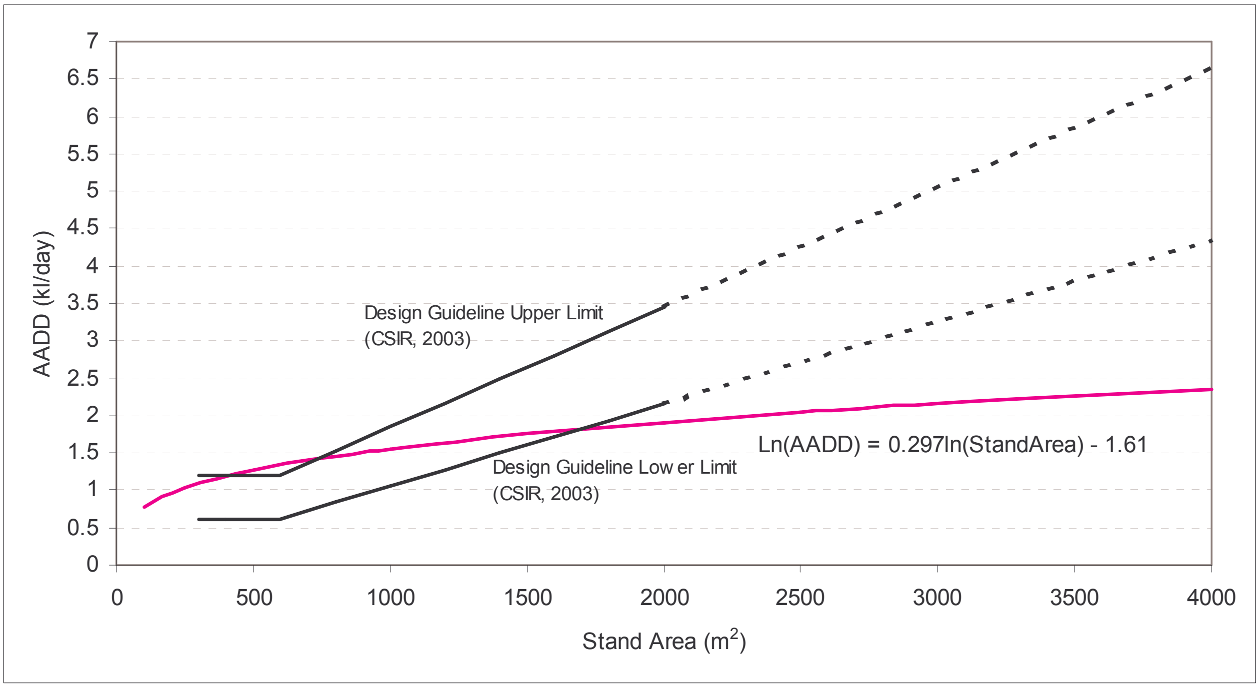


Figure 5-2: Single Variable Regression Model with Stand Area for All Domestic Data

It can be seen from Figure 5.2 that the regression model predicts AADD values within the range of the guideline envelope curves up to stand areas of approximately 1 700 m^2. The regression line lies very close to the upper limit for stand areas smaller than 750 m^2. In fact, the model predicts slightly higher AADD values for stand areas between 500 m^2 and 750 m^2. For stand areas larger than 1 700 m^2 the model predicts much lower AADD values. This result corresponds with the preliminary evaluation that was made by comparing the suburb averages with the "Red Book" guideline, namely that the majority of the stands that exhibited higher consumption than what the guideline predicted were smaller stands (stand areas less than 800m^2).

Figure 5.3 plots the regression equation with its 97.5 % confidence limits. The 97.5 % confidence limits lie very close to the regression line, so much so that the difference can hardly be distinguished. The 97.5 % confidence limits indicate the 97.5 % probability that the average AADD of all stands will lie within these lower and upper limits. The extremely large dataset is responsible for the fact that the confidence intervals lie so close to the regression line.

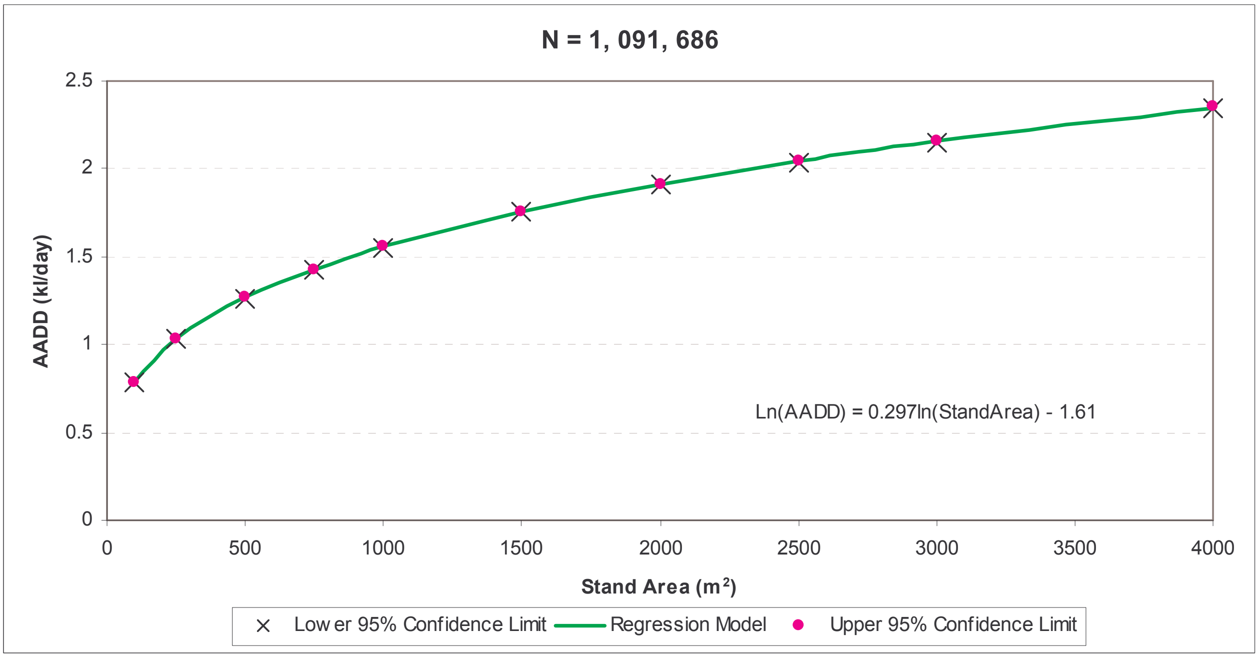


Figure 5-3: Confidence Intervals for the Single Variable Regression Model for Domestic Use (all data points)

As discussed in Chapter 4 (Methodology) the assessment of the "Red Book" design guidelines were further refined by categorising the data in two categories for geographic location namely coastal or inland. A single variable regression model was determined for the average demand as a function of the stand area for each category. The results are as follows:

Inland

$$\ln(AADD) = -0.1691 + 0.314\ln(StdArea) \pm 0.884\sqrt{1.234\times10^{-6} + \frac{(\ln(StdArea) - 6.487)^2}{480302}} \tag{5.2}$$

Adjusted R^2 = 0.223; N = 810 595

Coastal

$$\ln(AADD) = -1.124 + 0.204\ln(StdArea) \pm 0.743\sqrt{3.56\times10^{-6} + \frac{(\ln(StdArea) - 6.198)^2}{169287}} \tag{5.3}$$

Adjusted R^2 = 0.148 N = 281 091

The coastal dataset had a greater unexplained variability compared with the inland dataset. The regression models had adjusted R square values (square of the Pearson product moment correlation coefficient) of 0.223 for the inland dataset and 0.148 for the coastal dataset. This implies that 22.3% of the variability in the data could be explained by the equation for inland users and only 14.8% of the variability by the coastal equation. As mentioned, an adjusted R square value of more than 20 % is considered good when predicting human behaviour. Numerous factors contribute to the variability in the data. The greater unexplained variability of the coastal dataset, may possibly be due to the large number of traditional seaside holiday villages included in the dataset. The demand patterns of these towns and villages exhibit seasonal variations that would cause significant variations the calculated AADD values used in this analysis. It is not recommended that the coastal relationship is used, as presented, to estimate water demand in coastal towns. It is merely presented here to highlight the difference in predicted water demand when compared to inland stands of the same stand area. The t-statistics for the regression coefficients indicated that they were significantly different from zero. The detailed regression statistics for both these regression models are contained in Appendix C. Figures 5.4 plots the regression models against the "Red Book" guideline curves for the inland and coastal categories. The 97.5 % confidence intervals for the average AADD were once again quite close to the regression line due to the large dataset in each case and hence were not distinguishable from the average curve for each category.

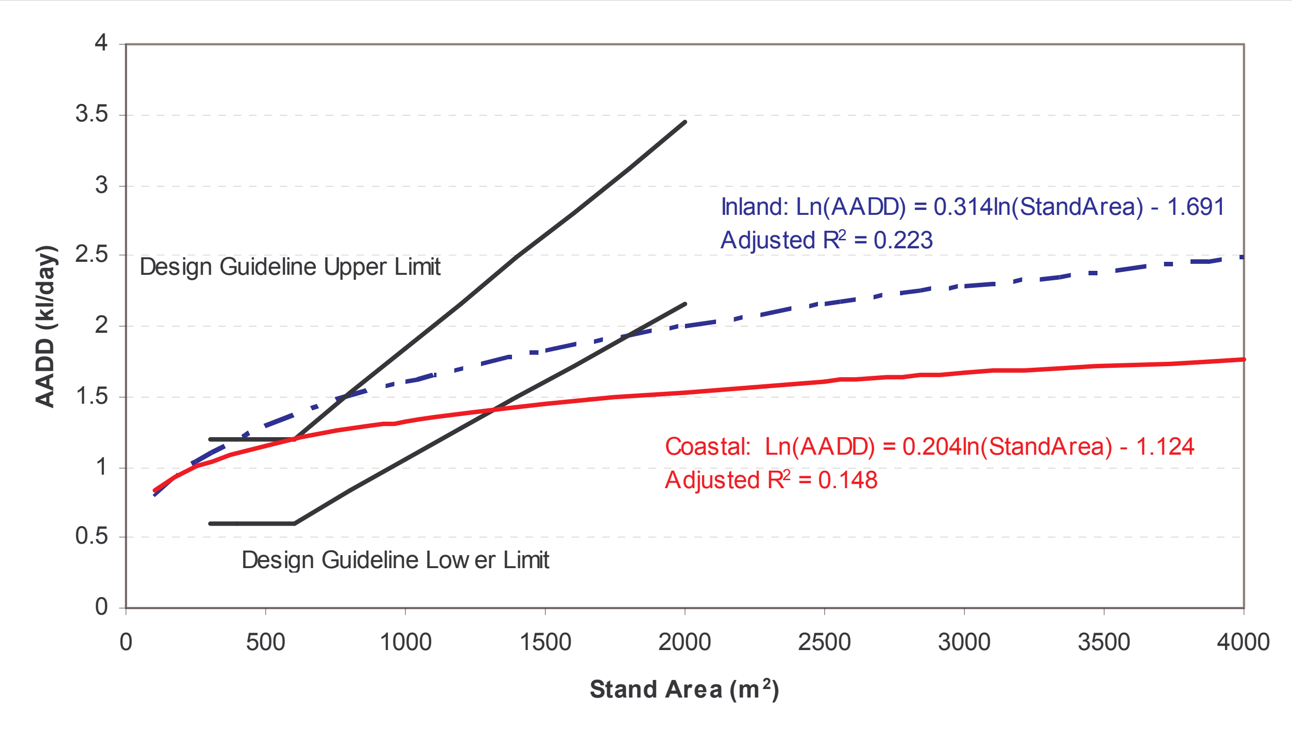


Figure 5-4: Single Variable Regression Models for the Inland and Coastal Categories

From Figure 5.4, it can be seen that the regression models predict that domestic consumptions in inland areas are generally higher than that of coastal areas. For stand areas less than 250 m^2 the geographic location makes no significant difference in the domestic water consumption. A possible explanation for the seemingly higher consumption at inland location may be that quite a number of the coastal towns studied constitute the typical holiday seaside town with a number of residences only being occupied during holiday seasons. Not much gardening is done at these houses and thus it could be expected that the water demand would be less when compared with an 'inland' house.

The regression model for coastal domestic consumption predicts that the average AADD will be within the "Red Book" design guideline curves up to stand areas of approximately 1300 m^2. For stand areas larger than 1300 m^2 the regression model indicates that the average domestic consumption is likely to be less than the minimum "Red Book" curve. For inland domestic consumption it seems that the model predicts the average AADD to fall well within the guideline curves (excluding stand areas of about 500m^2 to 800m^2) up to stand areas of approximately 1800 m^2. Thereafter the model predicts lower AADD values.

5.1.2 Assessment of Factors Influencing Domestic Water Consumption

All the domestic data was divided into 8 categories of similar stand area and also into 10 categories of similar stand value, as discussed in detail in Chapter 4. On each of these 18 categories, a stepwise multiple regression analysis was performed. Table 5.3 summarises the outcome of the stepwise regression analyses on each category by listing the independent variables influencing water consumption in each category by order of best correlation with AADD. The number of data points in each analysis is also indicated and the adjusted R-square value of the each stepwise regression model. The detailed regression statistics are contained in Appendix C, including the regression coefficients for each regression model (or equation).

Table 5-1: Summary of the Outcome of the Stepwise Multi-Variable Regression Analyses on Domestic Categories

Category	Description of the Category	Number of points	Number of Variables	Independent Variables by order of Best Correlation with Ln (AADD)	Cumulative Adjusted R^2
STAND AREA CATEGORIES					
RES500	Single residential stands with stand areas of greater than or equal to 20 m^2 and smaller than 500 m^2	433 104	14	Stand Area	0.053
				Geographic Location	0.066
				Ln (Stand Value)	0.079
				Mean Annual Evaporation	0.082
				Ave Min Temperature	0.087
				% Unemployment	0.089
				Ave Household Income	0.09
				% Waterborne Sanitation	0.09
				Ave Max Temperature	0.091
				Mean Annual Precipitation	0.091
				% Water Connection	0.091
				Ave Household Size	0.091
				% Formal Housing	0.091
				Ave House Size	0.091
RES750	Single residential stands with stand areas of greater than or equal to 500 m^2 and smaller than 750 m^2	112 443	14	Ln(Stand Value)	0.037
				Geographic Location	0.062
				Ave Household Income	0.071
				Mean Annual Precipitation	0.078
				Ave House Size	0.085
				Stand Area	0.088
				% Water connection	0.09
				Ave Max Temperature	0.094
				Mean Annual Evaporation	0.097
				Stand Value	0.099
				% unemployed	0.101
				Ave Min Temperature	0.101
				Ave Household Size	0.102
				Ln(Stand Area)	0.102
RES1000	Single residential stands	136 019	14	Ln(Stand Value)	0.067

Category	Description of the Category	Number of points	Number of Variables	Independent Variables by order of Best Correlation with Ln (AADD)	Cumulative Adjusted R^2
	with stand areas of greater than or equal to 750 m^2 and smaller than 1 000 m^2			Geographic Location	0.086
				Ave Min Temperature	0.101
				Mean Annual Evaporation	0.113
				Mean Annual Precipitation	0.12
				Ave Household Income	0.124
				Ave House Size	0.132
				Stand Value	0.134
				Stand Area	0.135
				Ave Max Temperature	0.137
				%water connection	0.139
				% unemployed	0.14
				% Formal Housing	0.14
				%Waterborne Sanitation	0.141
RES1500	Single residential stands with stand areas of greater than or equal to 1 000 m^2 and smaller than 1 500 m^2	184 497	15	Ln(Stand Value)	0.092
				Ave Household Income	0.117
				Ave House Size	0.139
				Geographic Location	0.15
				Mean Annual Precipitation	0.161
				Stand Area	0.167
				Ave Min Temperature	0.171
				Mean Annual Evaporation	0.172
				Stand Value	0.174
				% Formal Housing	0.175
				%Waterborne Sanitation	0.176
				Ave Max Temperature	0.176
				% unemployed	0.176
				Ave Household size	0.177
				%house water connection	0.177
RES2000	Single residential stands with stand areas of greater than or equal to 1 500 m^2 and smaller than 2 000 m^2	49 952	11	Ln(Stand Value)	0.105
				Ave Household Income	0.132
				Ave House Size	0.155
				Geographic Location	0.163
				Mean Annual Precipitation	0.169
				%Formal Housing	0.171
				Stand Value	0.173
				Stand Area	0.173
				Mean Annual Evaporation	0.174
				Ln(Stand Area)	0.174
				%house water connection	0.174
RES2500	Single residential stands with stand areas of greater than or equal to 2 000 m^2 and smaller than 2 500 m^2	17 506	11	Ln(Stand Value)	0.092
				Geographic Location	0.125
				Ave Household Income	0.136
				Ave House Size	0.144
				%Formal Housing	0.146
				%unemployed	0.147
				% water connection	0.148
				Mean Annual Precipitation	0.149
				Stand Value	0.15
				Ave Household size	0.15
				Mean Annual Evaporation	0.15
RES3000	Single residential stands	6 884	8	Ln(Stand Value)	0.161

Category	Description of the Category	Number of points	Number of Variables	Independent Variables by order of Best Correlation with Ln (AADD)	Cumulative Adjusted R^2
	with stand areas of greater than or equal to 2 500 m^2 and smaller than 3 000 m^2			Geographic Location	0.187
				Stand Area	0.194
				Ave Household Income	0.199
				Ave House Size	0.206
				Mean Annual Precipitation	0.207
				%Formal Housing	0.208
				%unemployed	0.208
RES4000	Single residential stands with stand areas of greater than or equal to 3 000 m^2 and smaller than 4 000 m^2	6 168	8	Ln (Stand Value)	0.164
				Geographic Location	0.196
				Stand Value	0.207
				Mean Annual Evaporation	0.213
				Ave Household Income	0.218
				Ave Max Temperature	0.22
				Mean Annual Precipitation	0.222
				%Waterborne Sanitation	0.223
STAND VALUE CATEGORIES					
RESVAL50000	Single residential stands with a stand value of greater than or equal to R 20 000 and smaller than R 50 000	211 753	16	Ln(Stand Value)	0.029
				Ln(Stand Area)	0.041
				Geographic Location	0.05
				Stand Value	0.056
				Ave Household Income	0.058
				% Unemployed	0.062
				Ave Min Temperature	0.063
				Ave Max Temperature	0.067
				Mean Annual Evaporation	0.07
				%Waterborne Sanitation	0.071
				%Formal Housing	0.072
				Stand Area	0.073
				Ave Household Size	0.073
				Mean Annual Precipitation	0.073
				% water connection	0.073
				Ave House Size	0.073
RESVAL100000	Single residential stands with a stand value of greater than or equal to R 50 000 and smaller than R 100 000	200 322	14	Ln(Stand Area)	0.033
				Geographic Location	0.049
				Mean Annual Evaporation	0.055
				Ave Min Temperature	0.058
				Ln(Stand Value)	0.06
				Ave Max Temperature	0.061
				Ave Household Income	0.062
				%Unemployed	0.065
				% Formal Housing	0.066
				Ave House Size	0.066
				Stand Area	0.066
				Mean Annual Precipitation	0.066
				Stand Value	0.066
				%water connection	0.066
RESVAL250000	Single residential stands with a stand value of greater than or equal to R 100 000 and smaller	298 703	15	Ln(Stand Area)	0.065
				Geographic Location	0.089
				Ave Household Income	0.099
				Ave House Size	0.106

Category	Description of the Category	Number of points	Number of Variables	Independent Variables by order of Best Correlation with Ln (AADD)	Cumulative Adjusted R^2
	than R 250 000			Stand Value	0.111
				%water connection	0.114
				%Unemployed	0.116
				Mean Annual Precipitation	0.118
				Ave Max Temperature	0.119
				Ave Min Temperature	0.124
				Mean Annual Evaporation	0.126
				%Formal Housing	0.127
				%Waterborne Sanitation	0.128
				Ave Household size	0.128
				Stand Area	0.128
RESVAL500000	Single residential stands with a stand value of greater than or equal to R 250 000 and smaller than R 500 000	227 595	16	Ln(Stand Area)	0.107
				Ave Household Income	0.134
				Ave House Size	0.152
				Geographic Location	0.161
				Stand Value	0.17
				Mean Annual Precipitation	0.178
				%water connection	0.181
				Ave Household size	0.182
				Ave Max Temperature	0.184
				Ave Min Temperature	0.185
				Mean Annual Evaporation	0.187
				Stand Area	0.188
				%Waterborne Sanitation	0.188
				%Unemployed	0.188
				%Formal Housing	0.188
				Ln(Stand Value)	0.188
RESVAL750000	Single residential stands with a stand value of greater than or equal to R 500 000 and smaller than R 750 000	49 807	13	Ln(Stand Area)	0.102
				Geographic Location	0.148
				Ave Household Income	0.165
				Ave House Size	0.179
				Mean Annual Precipitation	0.185
				Stand Area	0.189
				Ln(Stand Value)	0.192
				%Waterborne Sanitation	0.194
				%Unemployed	0.197
				Ave Max Temperature	0.198
				%water connection	0.199
				Mean Annual Evaporation	0.199
				%Formal Housing	0.2
RESVAL1000000	Single residential stands with a stand value of greater than or equal to R 750 000 and smaller than R 1 000 000	14 351	13	Geographic Location	0.087
				Ln(Stand Area)	0.133
				Mean Annual Precipitation	0.139
				Ave Max Temperature	0.146
				Ave Min Temperature	0.149
				Ave Household Income	0.15
				%Unemployed	0.159
				%water connection	0.165
				Ave House Size	0.166
				Stand Area	0.167
				Ln(Stand Value)	0.167

Category	Description of the Category	Number of points	Number of Variables	Independent Variables by order of Best Correlation with Ln (AADD)	Cumulative Adjusted R^2
				Mean Annual Evaporation	0.168
				%Formal Housing	0.168
RESVAL2000000	Single residential stands with a stand value of greater than or equal to R 1 000 000 and smaller than R 2 000 000	8 704	12	Geographic Location	0.079
				Ln(Stand Area)	0.118
				Ave Min Temperature	0.141
				%Unemployed	0.147
				Ave Household Income	0.16
				%water connection	0.169
				Stand Value	0.175
				Stand Area	0.18
				Mean Annual Evaporation	0.184
				Ave House Size	0.185
				Ln(Stand Value)	0.185
				Mean Annual Precipitation	0.186
RESVAL3000000	Single residential stands with a stand value of greater than or equal to R 2 000 000 and smaller than R 3 000 000	1 286	9	Ave Min Temperature	0.261
				%water connection	0.356
				Ave Max Temperature	0.388
				Mean Annual Precipitation	0.424
				%Waterborne Sanitation	0.448
				Ave Household Income	0.462
				Mean Annual Evaporation	0.469
				Ave House Size	0.471
				Stand Area	0.472
RESVAL4000000	Single residential stands with a stand value of greater than or equal to R 3 000 000 and smaller than R 4 000 000	428	3	Ave Min Temperature	0.195
				Geographic Location	0.291
				Stand Area	0.298
RESGT4000000	Single residential stands with a stand value of greater than or equal to R 4 000 000	1 275	7	Mean Annual Evaporation	0.274
				%Unemployed	0.366
				Ave Min Temperature	0.39
				Ave Household Income	0.402
				Ave House Size	0.414
				Ave Max Temperature	0.427
				Mean Annual Precipitation	0.429

When the data was categorised by stand area, it lessened to an extent or removed the effect of stand area on water consumption, in the regression analyses. Consequently, stand value emerged as the most significant parameter influencing domestic water consumption for all the stand area categories, apart from the smallest stand area category (stands with areas between 20 and 500 m^2). For the smallest stand area category, stand area came out as the most significant parameter affecting water consumption. A possible reason for this is that the smallest stand area category had quite a wide range of stand values, from low income houses in partly informal settlements to high income security residential developments.

For six of the eight stand area categories, geographic location (inland vs. coastal) came forward as the parameter that correlated with ln(AADD) next best to stand value.

Similarly, as was the case with the stand area categories, when the data was categorised by stand value, it removed to an extent the effect of stand value on the water consumption in the regression analysis and thus it can be seen from Table 5.1, stand area appeared as the most significant parameter influencing consumption for four of the stand value categories and as the second most significant parameter for most of the remaining categories. For stands with stand values of less than R50 000, stand value emerged as the most significant parameter influencing water consumption and not stand area as expected. These stands could be considered to be mostly in low income areas. A possible explanation is that in these areas not much gardening is undertaken and hence stand area would not be expected to have the most significant influence on water consumption but rather stand value as a proxy for income.

The water consumption of stands that have values greater than R750 000 but smaller than R 2 million was predicted by the multiple variable model to be mostly influenced by geographic location (inland vs. coastal). While the water consumption of stands with values between R 2 million and R4 million was dictated by the annual average minimum temperature. Mean Annual Evaporation was indicated to be the most significant parameter influencing water consumption for stands with values greater than R4 million.

Geographic location also emerged as a significant parameter influencing water demand for most stand value categories.

The adjusted R-square values for most of the regression models appear to be quite low. However, it should be kept in mind that human behaviour is being studied and an adjusted R-square value of greater than 20 % is considered a good fit when predicting human behaviour. Some R^2 values are still considerably smaller than 20 %. This is especially true of the larger categories such as RES500 and RESVAL50000. It is true that more data points will have greater inherent variability and thus lower adjusted R^2 values. Once again, attention is drawn to the fact that the data for stand values originated from different treasury systems covering different time periods over the last eight years (1998 -2005). Furthermore, the stand value data might not have been updated regularly by treasuries resulting in out of date stand values. This might result in, for example, a 1999 -2002 dataset still containing stand values for some records relating to say early nineties. This could greatly contribute to the noted variability in the data. The regression models are all however, considered to be statistically significant given the large dataset and the significant t- and F-statistics.

It should also be noted that given the nature of the stand value data discussed above, this study considered the stand values as a relative indication of the true values only. This stand value was considered as a proxy indicator for income level defined as follows:

Table 5-2: Definition of Income Level Used in the Study

Stand Value Range	Income Level
R 20 000 – R 50 000	Very Low Income
R 50 000 – R100 000	Low Income
R100 000 – R250 000	Low Middle Income
R250 000 – R500 000	High Middle Income
R500 000 – R750 000	High Income
More than R 750 000	Very High Income

Table 5.3 and Table 5.4 below summarise the first step regression equations (i.e. first step of the regression with the independent variable that influenced AADD the most) that were obtained for the stand area and stand value categories respectively. Appendix C can be consulted for the remaining models for each domestic category.

Table 5-3: Regression Results (First Step) of the Stepwise Multiple Variable Regression Analyses for Categories of Stand Area

Stand Area Category	Regression Equation
>=20m^2 and < 500m^2	$\ln AADD = 0.001 StdArea - 0.195$
>=500m^2 and < 750m^2	$\ln AADD = 0.099 \ln(StdValue) - 0.920$
>=750m^2 and < 1000m^2	$\ln AADD = 0.181 \ln(StdValue) - 1.824$
>=1000m^2 and < 1500m^2	$\ln AADD = 0.251 \ln(StdValue) - 2.595$
>=1500m^2 and < 2000m^2	$\ln AADD = 0.295 \ln(StdValue) - 3.050$
>=2000m^2 and < 2500m^2	$\ln AADD = 0.279 \ln(StdValue) - 2.788$
>=2500m^2 and < 3000m^2	$\ln AADD = 0.348 \ln(StdValue) - 3.670$

Stand Area Category	Regression Equation
>=3000m^2 and < 4000m^2	$\ln AADD = 0.331\ln(StdValue) - 3.453$

Table 5-4: Regression Results (First Step) of the Stepwise Multiple Variable Regression Analyses for Categories of Stand Value

Stand Value Category	Regression Equation
>=R20,000 and < R50,000	$\ln AADD = 0.266\ln(StdValue) - 2.736$
>=R50,000 and < R100,000	$\ln AADD = 0.133\ln(StdArea) - 0.661$
>=R100,000 and < R250,000	$\ln AADD = 0.186\ln(StdArea) - 0.910$
>=R250,000 and < R500,000	$\ln AADD = 0.286\ln(StdArea) - 1.455$
>=R500,000 and < R750,000	$\ln AADD = 0.266\ln(StdArea) - 1.159$
>=R750,000 and < R1,000,000	$\ln AADD = -0.427(GL) + 1.398$
>=R1,000,000 and < R2,000,000	$\ln AADD = -0.446(GL) + 1.534$
>=R2,000,000 and < R3,000,000	$\ln AADD = -0.172(MinTemp) + 2.890$
>=R3,000,000 and < R4,000,000	$\ln AADD = -0.123(MinTemp) + 2.986$
>=R4,000,000	$\ln AADD = 0.0027(MAE) - 3.336$

Where:

StdValue = Stand Value in R

StdArea = Stand Area in m^2

GL = Geographic Location where Inland = 2 and Coastal = 1

MinTemp =Annual Average Minimum Daily Temperature in degree Celsius

MAE =Mean Annual Evaporation in mm

Figure 5.5 portrays these regression equations for the various stand area categories graphically where stand value was identified as the most significant independent variable (i.e. for seven of the eight categories).

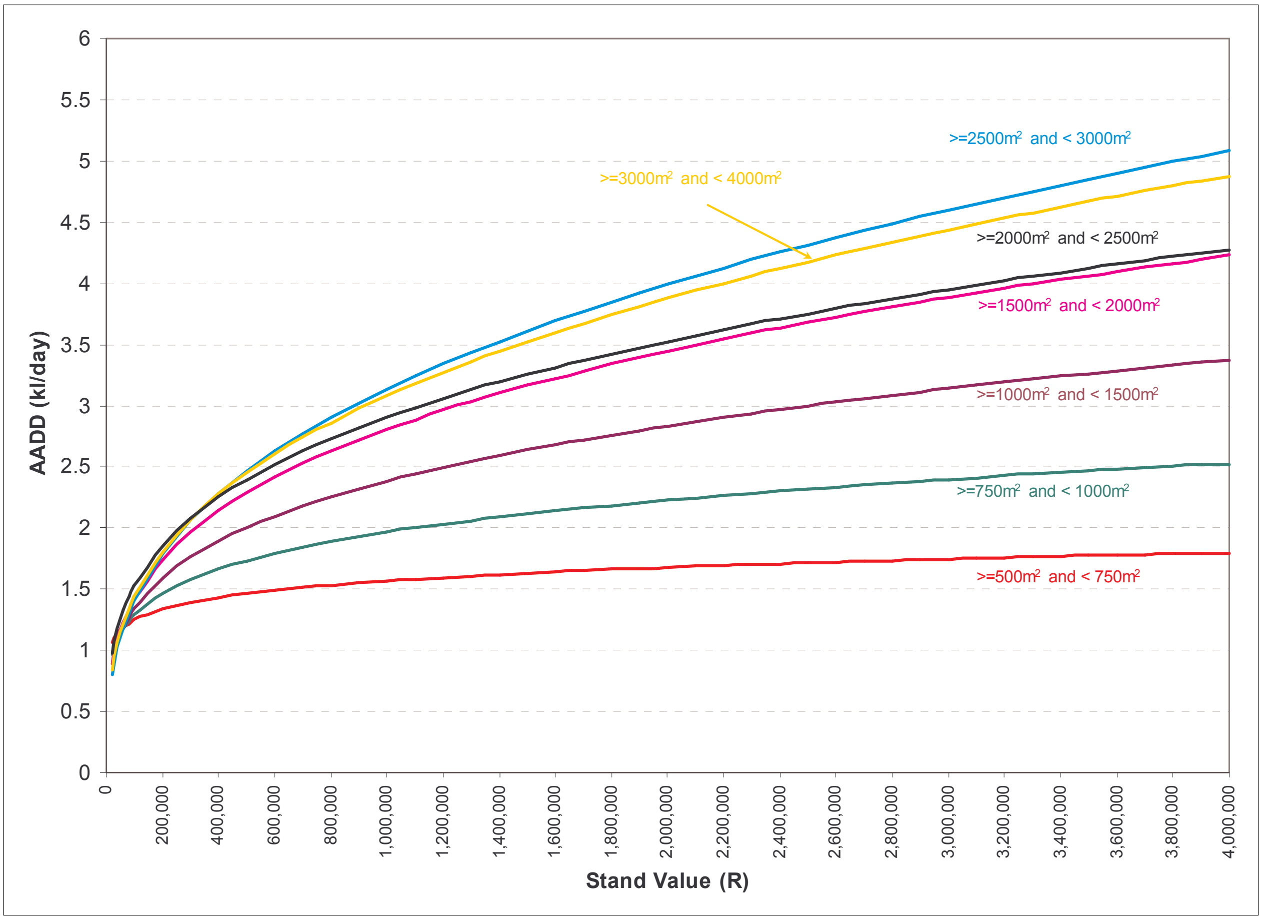


Figure 5-5: Regression Results (First Step) of the Stepwise Multiple Variable Regression for Stand Area Categories

Figure 5.5 shows the marked impact of stand area on domestic water consumption. For the same stand value or income level, larger stands will consume more water than smaller stands. This analysis thus confirms stand area as the most significant parameter that influences domestic water consumption as indicated by the "Red Book" design guideline.

Figure 5.6 depicts the first step of the multiple regression analysis for various categories of stand value where stand area emerged as the most significant parameter (i.e. first independent variable). The "Red Book" upper and lower guideline curves are also shown on Figure 5.6. The regression models predicted that higher income levels would consume more water than lower income levels for the same stand area.

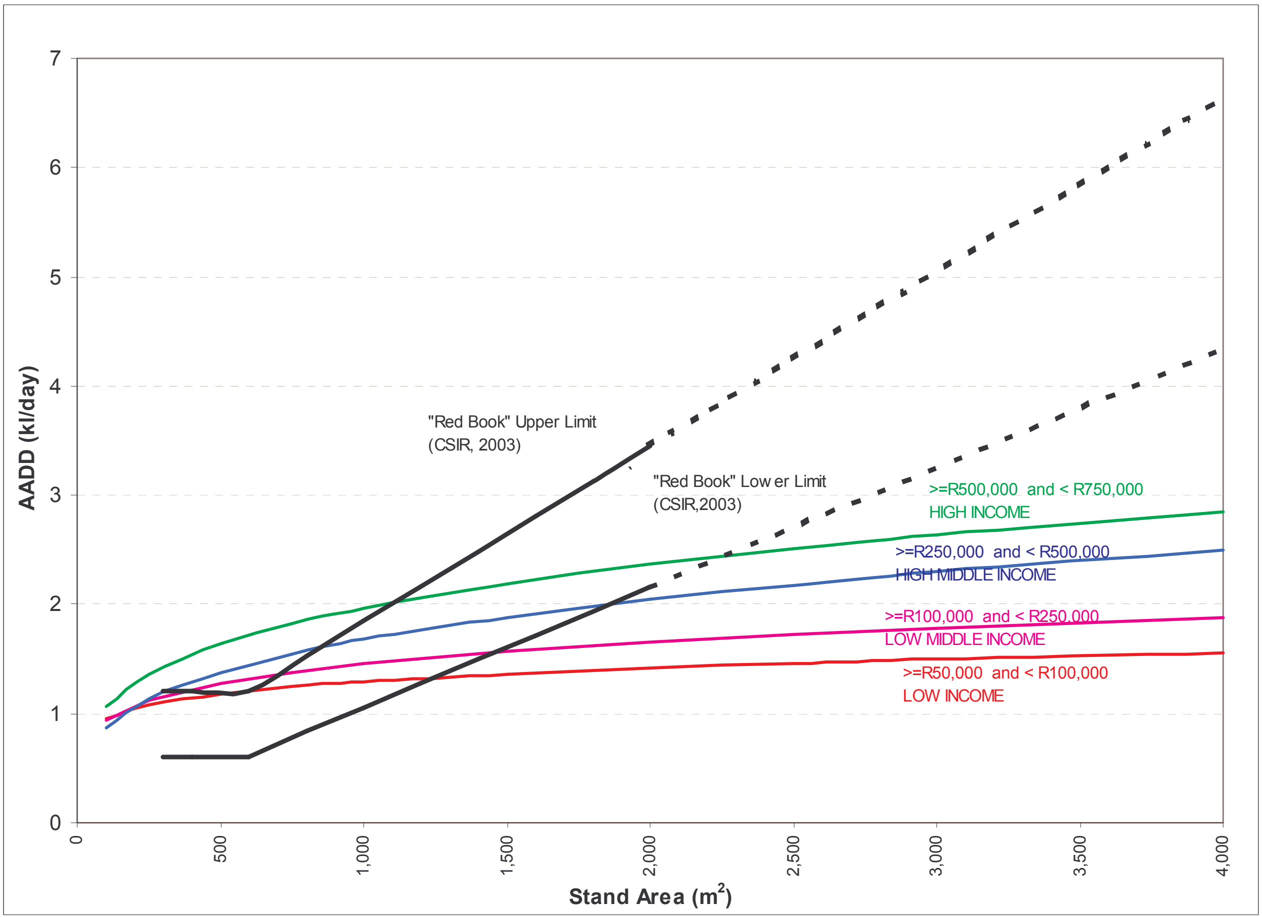


Figure 5-6: Regression Results (First Step) of the Stepwise Multiple Variable Regression for Stand Value Categories

The outcome of the stepwise multiple variable regression analyses, apart from providing a combined model of the parameters influencing water consumption, was that the top three parameters that influence domestic water consumption the most were:

- Stand Area
- Stand Value (as a proxy for income)
- Geographic Location (Inland or Coastal)

As discussed in Chapter 4 (Methodology) the next step in the data analysis was to perform single variable regressions for each of these parameters. A summary of the regression equations with the 97.5 % confidence intervals are contained in tabular format below. The detail regression statistics are contained in Appendix C.

Table 5-5: Single Variable Regression Results for Domestic Categories

Category	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
RES500	Inland	Stand Area	$\ln(AADD) = 0.317\ln(StdArea) - 1.703 \pm 0.824\sqrt{3\times10^{-6} + \frac{(\ln StdArea - 5.746)^2}{24707}}$	333 880	0.041
		Stand Value	$\ln(AADD) = 0.148\ln(StdValue) - 1.505 \pm 0.817\sqrt{3\times10^{-6} + \frac{(\ln StdValue - 10.965)^2}{156256}}$	333 880	0.056
	Coastal	Stand Area	$\ln(AADD) = 0.16\ln(StdArea) - 0.899 \pm 0.607\sqrt{6.9\times10^{-6} + \frac{(\ln StdArea - 5.634)^2}{26683}}$	145 019	0.047
		Stand Value	$\ln(AADD) = 0.072\ln(StdValue) - 0.798 \pm 0.612\sqrt{6.9\times10^{-6} + \frac{(\ln StdValue - 11.216)^2}{86141}}$	145 019	0.03
RES750	Inland	Stand Area	$\ln(AADD) = 0.436\ln(StdArea) - 2.547 \pm 0.829\sqrt{1.2\times10^{-5} + \frac{(\ln StdArea - 6.422)^2}{1294}}$	86 291	0.016
		Stand Value	$\ln(AADD) = 0.153\ln(StdValue) - 1.531 \pm 0.770\sqrt{1.2\times10^{-5} + \frac{(\ln StdValue - 11.791)^2}{41605}}$	80 473	0.073
	Coastal	Stand Area	$\ln(AADD) = 0.269\ln(StdArea) - 1.549 \pm 0.713\sqrt{1.7\times10^{-5} + \frac{(\ln StdArea - 6.425)^2}{755}}$	58 088	0.07
		Stand Value	$\ln(AADD) = 0.089\ln(StdValue) - 0.904 \pm 0.698\sqrt{1.8\times10^{-5} + \frac{(\ln StdValue - 12.137)^2}{36686}}$	56 354	0.039
RES1000	Inland	Stand Area	$\ln(AADD) = 0.453\ln(StdArea) - 2.667 \pm 0.813\sqrt{8.2\times10^{-6} + \frac{(\ln StdArea - 6.805)^2}{977}}$	122 181	0.01
		Stand Value	$\ln(AADD) = 0.230\ln(StdValue) - 2.402 \pm 0.776\sqrt{8.2\times10^{-6} + \frac{(\ln StdValue - 12.268)^2}{38120}}$	122 181	0.095
	Coastal	Stand Area	$\ln(AADD) = 0.269\ln(StdArea) - 1.533 \pm 0.786\sqrt{2.8\times10^{-5} + \frac{(\ln StdArea - 6.77)^2}{285}}$	35 667	0.03
		Stand Value	$\ln(AADD) = 0.147\ln(StdValue) - 1.546 \pm 0.757\sqrt{2.8\times10^{-5} + \frac{(\ln StdValue - 12.509)^2}{19510}}$	35 667	0.073
RES1500	Inland	StandArea	$\ln(AADD) = 0.418\ln(StdArea) - 2.399 \pm 0.864\sqrt{5.7\times10^{-6} + \frac{(\ln StdArea - 7.067)^2}{2802}}$	175 128	0.014

Category	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
		Stand Value	$\ln(AADD) = 0.292\ln(StdValue) - 3.090 \pm 0.817\sqrt{5.7\times10^{-6} + \frac{(\ln StdValue - 12.481)^2}{47985}}$	175 128	0.118
	Coastal	StandArea	$\ln(AADD) = 0.233\ln(StdArea) - 1.266 \pm 0.835\sqrt{3.3\times10^{-5} + \frac{(\ln StdArea - 7.037)^2}{426}}$	30 410	0.004
		StandValue	$\ln(AADD) = 0.193\ln(StdValue) - 2.055 \pm 0.790\sqrt{3.3\times10^{-5} + \frac{(\ln StdValue - 12.596)^2}{16452}}$	30 410	0.11
RES2000	Inland	StandArea	$\ln(AADD) = 0.335\ln(StdArea) - 1.764 \pm 0.982\sqrt{2.0\times10^{-5} + \frac{(\ln StdArea - 7.459)^2}{488}}$	48 846	0.005
		Stand Value	$\ln(AADD) = 0.317\ln(StdValue) - 3.308 \pm 0.923\sqrt{2.0\times10^{-5} + \frac{(\ln StdValue - 12.741)^2}{14996}}$	48 846	0.122
	Coastal	StandArea	$\ln(AADD) = 0.293\ln(StdArea) - 1.688 \pm 1.037\sqrt{1.9\times10^{-4} + \frac{(\ln StdArea - 7.447)^2}{41.2}}$	5 150	0.002
		StandValue	$\ln(AADD) = 0.217\ln(StdValue) - 2.292 \pm 0.982\sqrt{1.9\times10^{-4} + \frac{(\ln StdValue - 12.830)^2}{3229}}$	5 150	0.105
RES2500	Inland	StandArea	$\ln(AADD) = -0.102\ln(StdArea) + 1.601 \pm 1.09\sqrt{5.8\times10^{-5} + \frac{(\ln StdArea - 7.678)^2}{68.7}}$ #	17 175	0.000
		Stand Value	$\ln(AADD) = 0.339\ln(StdValue) - 3.529 \pm 1.011\sqrt{5.8\times10^{-5} + \frac{(\ln StdValue - 12.848)^2}{6134}}$	17 175	0.133
	Coastal	StandArea	$\ln(AADD) = -0.642\ln(StdArea) + 5.485 \pm 1.19\sqrt{4\times10^{-4} + \frac{(\ln StdArea - 7.691)^2}{9.9}}$	2 477	0.005
		StandValue	$\ln(AADD) = 0.211\ln(StdValue) - 2.179 \pm 1.12\sqrt{4\times10^{-4} + \frac{(\ln StdValue - 12.906)^2}{2471}}$	2 477	0.121
RES3000	Inland	StandArea	$\ln(AADD) = 1.502\ln(StdArea) - 11.037 \pm 1.219\sqrt{1.4\times10^{-4} + \frac{(\ln StdArea - 7.904)^2}{28.6}}$	7 153	0.022
		Stand Value	$\ln(AADD) = 0.371\ln(StdValue) - 3.944 \pm 1.105\sqrt{1.4\times10^{-4} + \frac{(\ln StdValue - 12.887)^2}{4027}}$	7 153	0.196

Category	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
	Coastal	StandArea	$\ln(AADD) = -0.091\ln(StdArea) + 1.329 \pm 1.321\sqrt{1.1\times10^{-3} + \frac{(\ln StdArea - 7.906)^2}{2.745}}$ #	915	-0.01
		StandValue	$\ln(AADD) = 0.195\ln(StdValue) - 1.941 \pm 1.252\sqrt{1.1\times10^{-3} + \frac{(\ln StdValue - 13.086)^2}{1095}}$	915	0.099
RES4000	Inland	StandArea	$\ln(AADD) = 0.474\ln(StdArea) - 2.896 \pm 1.364\sqrt{1.7\times10^{-4} + \frac{(\ln StdArea - 8.178)^2}{47.3}}$	5 915	0.004
		Stand Value	$\ln(AADD) = 0.363\ln(StdValue) - 3.842 \pm 1.266\sqrt{1.7\times10^{-4} + \frac{(\ln StdValue - 13.286)^2}{3111}}$	5 915	0.142
	Coastal	StandArea	$\ln(AADD) = 0.046\ln(StdArea) + 0.06 \pm 1.462\sqrt{8.9\times10^{-4} + \frac{(\ln StdArea - 8.150)^2}{9}}$ #	1 127	-0.001
		StandValue	$\ln(AADD) = 0.175\ln(StdValue) - 1.802 \pm 1.358\sqrt{8.9\times10^{-4} + \frac{(\ln StdValue - 12.765)^2}{2811}}$	1 127	0.137

Stand Value Categories

Category	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
RESV50000	Inland	StandArea	$\ln(AADD) = 0.11\ln(StdArea) - 0.598 \pm 0.829\sqrt{5.8\times10^{-6} + \frac{(\ln StdArea - 5.712)^2}{22414}}$	171 578	0.009
		Stand Value	$\ln(AADD) = 0.283\ln(StdValue) - 2.895 \pm 0.788\sqrt{4.7\times10^{-6} + \frac{(\ln StdValue - 10.276)^2}{18561}}$	211 842	0.041
	Coastal	StandArea	$\ln(AADD) = 0.092\ln(StdArea) - 0.565 \pm 0.619\sqrt{1.8\times10^{-5} + \frac{(\ln StdArea - 5.678)^2}{19480}}$	56 049	0.028
		StandValue	$\ln(AADD) = 0.154\ln(StdValue) - 1.645 \pm 0.633\sqrt{1.7\times10^{-5} + \frac{(\ln StdValue - 10.418)^2}{3589}}$	59 784	0.013
RESV100000	Inland	StandArea	$\ln(AADD) = 0.115\ln(StdArea) - 0.527 \pm 0.806\sqrt{6.8\times10^{-6} + \frac{(\ln StdArea - 5.938)^2}{36382}}$	146 424	0.019

Category	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
		Stand Value	$\ln(AADD)=0.178\ln(StdValue)-1.83\pm0.809\sqrt{6.8\times10^{-6}+\frac{(\ln StdValue-11.152)^2}{5733}}$	147 553	0.007
	Coastal	StandArea	$\ln(AADD)=0.101\ln(StdArea)-0.552\pm0.623\sqrt{1.5\times10^{-5}+\frac{(\ln StdArea-5.769)^2}{28561}}$	66 168	0.042
		StandValue	$\ln(AADD)=0.124\ln(StdValue)-1.336\pm0.655\sqrt{1.4\times105+\frac{(\ln StdValue-11.152)^2}{2578}}$	70 281	0.005
RESV250000	Inland	StandArea	$\ln(AADD)=0.148\ln(StdArea)-0.632\pm0.809\sqrt{4.0\times10^{-6}+\frac{(\ln StdArea-6.698)^2}{69973}}$	252 741	0.034
		Stand Value	$\ln(AADD)=0.310\ln(StdValue)-3.359\pm0.809\sqrt{4.0\times10^{-6}+\frac{(\ln StdValue-12.01)^2}{16480}}$	235 444	0.035
	Coastal	StandArea	$\ln(AADD)=0.142\ln(StdArea)-0.767\pm0.655\sqrt{1.4\times10^{-5}+\frac{(\ln StdArea-6.301)^2}{26627}}$	71 795	0.063
		StandValue	$\ln(AADD)=0.059\ln(StdValue)-0.583\pm0.684\sqrt{1.4\times10^{-5}+\frac{(\ln StdValue-11.999)^2}{5454}}$	72 644	0.002
RESV500000	Inland	StandArea	$\ln(AADD)=0.226\ln(StdArea)-0.996\pm0.849\sqrt{5.5\times10^{-6}+\frac{(\ln StdArea-7.068)^2}{44990}}$	181 410	0.063
		Stand Value	$\ln(AADD)=0.364\ln(StdValue)-4.026\pm0.868\sqrt{5.5\times10^{-6}+\frac{(\ln StdValue-12.721)^2}{6833}}$	181 566	0.025
	Coastal	StandArea	$\ln(AADD)=0.206\ln(StdArea)-1.081\pm0.733\sqrt{1.6\times10^{-5}+\frac{(\ln StdArea-6.627)^2}{12912}}$	62 371	0.059
		StandValue	$\ln(AADD)=0.359\ln(StdValue)-4.298\pm0.747\sqrt{1.6\times10^{-5}+\frac{(\ln StdValue-12.75)^2}{2405}}$	62 611	0.033
RESV750000	Inland	StandArea	$\ln(AADD)=0.210\ln(StdArea)-0.692\pm0.958\sqrt{2.6\times10^{-5}+\frac{(\ln StdArea-7.29)^2}{16278}}$	38 207	0.073
		Stand Value	$\ln(AADD)=0.383\ln(StdValue)-4.256\pm0.994\sqrt{2.6\times10^{-5}+\frac{(\ln StdValue-13.286)^2}{488}}$	38 230	0.007

Category	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
	Coastal	StandArea	$\ln(AADD) = 0.272\ln(StdArea) - 1.416 \pm 0.909\sqrt{6.4\times10^{-5} + \frac{(\ln StdArea - 6.95)^2}{4298}}$	15 593	0.086
		StandValue	$\ln(AADD) = 0.172\ln(StdValue) - 1.812 \pm 0.955\sqrt{6.4\times10^{-5} + \frac{(\ln StdValue - 13.282)^2}{200}}$	15 651	0.002
RESV1million	Inland	StandArea	$\ln(AADD) = 0.161\ln(StdArea) - 0.240 \pm 1.107\sqrt{8.6\times10^{-5} + \frac{(\ln StdArea - 7.473)^2}{6240}}$	11 614	0.042
		Stand Value	$\ln(AADD) = 0.421\ln(StdValue) - 4.775 \pm 1.131\sqrt{8.6\times10^{-5} + \frac{(\ln StdValue - 13.651)^2}{76.8}}$	11 625	0.003
	Coastal	Stand Area	$\ln(AADD) = 0.239\ln(StdArea) - 1.192 \pm 1.103\sqrt{2.4\times10^{-4} + \frac{(\ln StdArea - 7.034)^2}{2118}}$	4 109	0.085
		Stand Value	$\ln(AADD) = 0.027\ln(StdValue) + 0.133 \pm 1.17\sqrt{2.4\times10^{-4} + \frac{(\ln StdValue - 13.647)^2}{28.17}}$ #	4 125	0.000
RESV2million	Inland	Stand Area	$\ln(AADD) = 0.099\ln(StdArea) - 0.315 \pm 1.325\sqrt{1.5\times10^{-4} + \frac{(\ln StdArea - 7.667)^2}{5343}}$	6 685	0.017
		Stand Value	$\ln(AADD) = 0.196\ln(StdValue) - 1.684 \pm 1.335\sqrt{1.5\times10^{-4} + \frac{(\ln StdValue - 14.039)^2}{224.4}}$	6 702	0.003
	Coastal	StandArea	$\ln(AADD) = 0.249\ln(StdArea) - 1.146 \pm 1.27\sqrt{3.5\times10^{-4} + \frac{(\ln StdArea - 7.174)^2}{2171}}$	2 855	0.101
		StandValue	$\ln(AADD) = 0.496\ln(StdValue) - 6.330 \pm 1.356\sqrt{3.5\times10^{-4} + \frac{(\ln StdValue - 14.068)^2}{103.2}}$	2 888	0.018
RESV3million	Inland	Stand Area	$\ln(AADD) = -0.331\ln(StdArea) + 3.97 \pm 2.185\sqrt{1.2\times10^{-3} + \frac{(\ln StdArea - 8.588)^2}{748.3}}$	856	0.07
		Stand Value	$\ln(AADD) = -2.316\ln(StdValue) + 35.19 \pm 2.189\sqrt{1.2\times10^{-3} + \frac{(\ln StdValue - 14.713)^2}{14.64}}$	866	0.067
	Coastal	StandArea	$\ln(AADD) = -0.139\ln(StdArea) + 1.654 \pm 1.713\sqrt{1.3\times10^{-3} + \frac{(\ln StdArea - 8.278)^2}{2254}}$	771	0.068

Category	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
		Stand Value	$\ln(AADD) = -1.808\ln(StdValue) + 27.12 \pm 1.717\sqrt{1.3\times10^{-3} + \frac{(\ln StdValue - 14.721)^2}{13.13}}$	777	0.066
RESV4million	Inland	Stand Area	$\ln(AADD) = -0.054\ln(StdArea) + 2.455 \pm 2.121\sqrt{3.6\times10^{-3} + \frac{(\ln StdArea - 8.568)^2}{272.63}}$ #	281	-0.001
		Stand Value	$\ln(AADD) = 1.843\ln(StdValue) - 25.701 \pm 2.095\sqrt{3.5\times10^{-3} + \frac{(\ln StdValue - 15.03)^2}{2.09}}$ #	282	0.018
	Coastal	Stand Area	$\ln(AADD) = 0.109\ln(StdArea) + 0.027 \pm 2.017\sqrt{3.4\times10^{-3} + \frac{(\ln StdArea - 7.992)^2}{387.4}}$ #	296	0.011
		Stand Value	$\ln(AADD) = -1.065\ln(StdValue) + 16.926 \pm 2.033\sqrt{3.3\times10^{-3} + \frac{(\ln StdValue - 15.045)^2}{2.18}}$ #	302	0.004
RESVGT4million	Inland	Stand Area	$\ln(AADD) = 0.070\ln(StdArea) + 1.327 \pm 3.138\sqrt{1.3\times10^{-3} + \frac{(\ln StdArea - 9.347)^2}{986.2}}$ #	789	0.001
		Stand Value	$\ln(AADD) = -0.375\ln(StdValue) + 7.926 \pm 3.136\sqrt{1.2\times10^{-3} + \frac{(\ln StdValue - 15.945)^2}{217.7}}$	805	0.013
	Coastal	Stand Area	$\ln(AADD) = -0.075\ln(StdArea) + 1.394 \pm 2.603\sqrt{9.3\times10^{-4} + \frac{(\ln StdArea - 9.321)^2}{1692}}$	1 074	0.004
		Stand Value	$\ln(AADD) = -0.197\ln(StdValue) + 3.883 \pm 2.626\sqrt{9.3\times10^{-4} + \frac{(\ln StdValue - 16.161)^2}{586.1}}$	1 079	0.011

 # Equation coefficients do not differ significantly from zero

The adjusted R-square values for a number of the regression equations were quite low. A possible reason is the size of the datasets and the inherent variability normally observed in the data attempting to describe human behaviour. However, the t-statistics and P-values for the regression coefficients were evaluated to ensure that the coefficients were significantly different from zero (i.e. if the P-value is smaller than 0.05 the regression coefficient is significantly different from zero). Where this was not the case it is indicated in Table 5.5. Where the coefficients were not significantly different from zero the equation could not be used in further analysis.

Figure 5.7 and Figure 5.8 plot the single variable regression models, with stand area as independent variable, for the stand value categories (up to stand values of R 2 million), on a single graph for the two geographic locations (inland and coastal) respectively. The regression models predicted inland water consumption to be consistently higher than coastal consumption for the same stand area and the same stand value category. Appendix C contains separate graphs for each stand value category and there the distinct difference between inland and coastal consumption can be easily compared. The 97.5% confidence intervals for the average AADD are also indicated on the separate graphs in Appendix C.

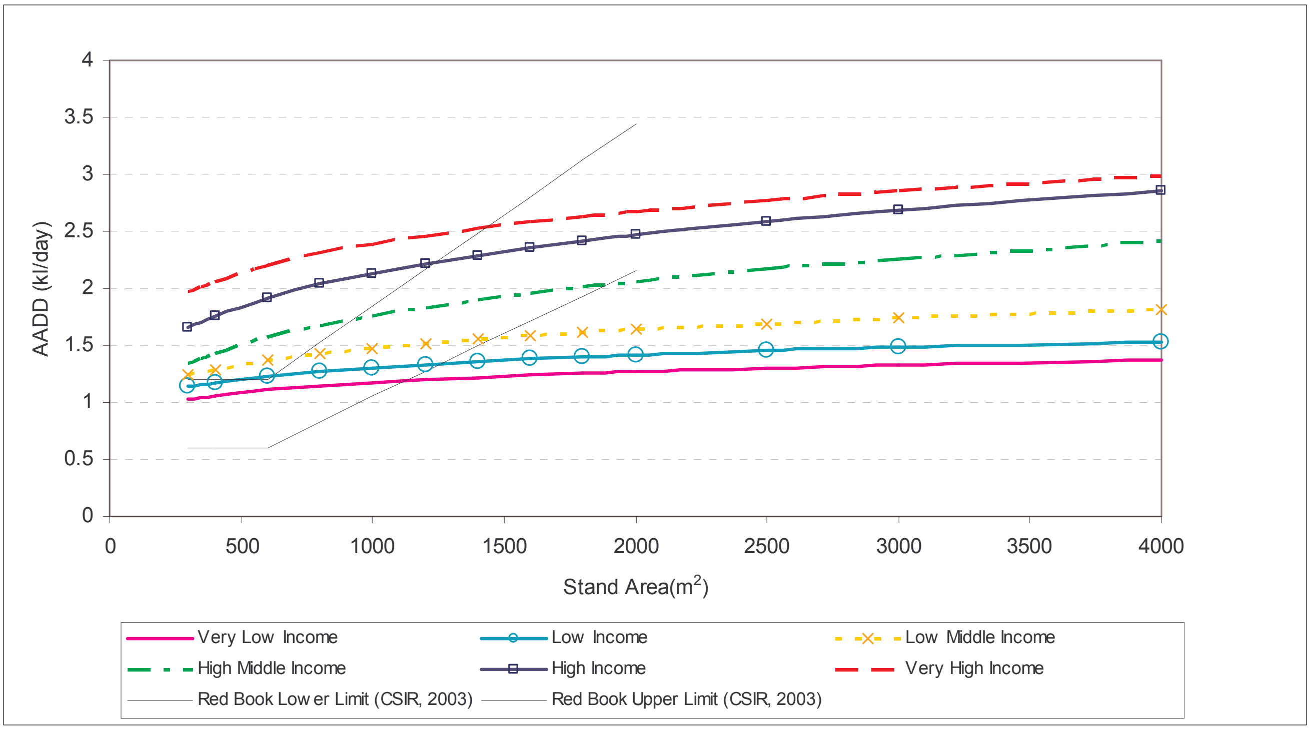


Figure 5-7: Single Variable Regression Analysis (AADD vs. Stand Area) for Categories of Stand Value (Income) and Geographic Location is Inland

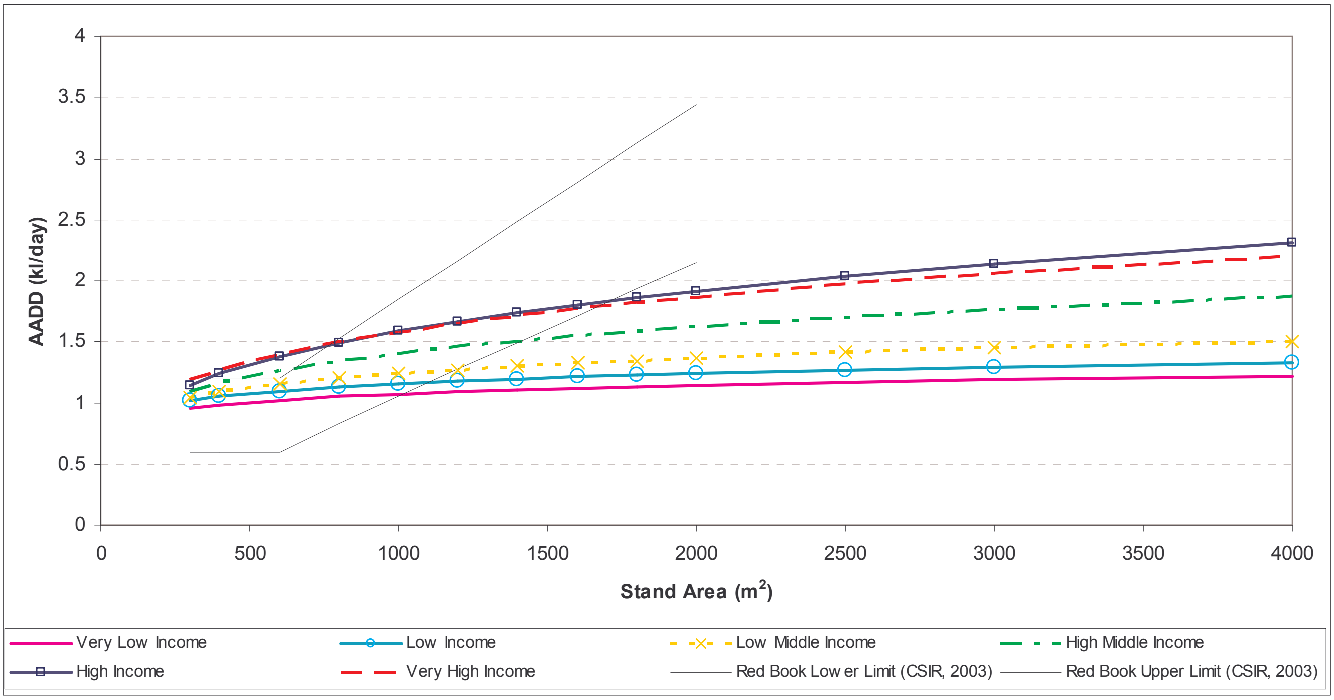


Figure 5-8: Single Variable Regression Analysis (AADD vs. Stand Area) for Categories of Stand Value (Income) and Geographic Location is Coastal

Figure 5.7 and 5.8 also indicate how the models could assist in interpreting the "Red Book" (CSIR, 2003) guideline curves when estimating domestic demand and taking income level into account. However, the regression models predicted much higher water consumption for smaller stands (500 to 1000 m^2) for especially inland domestic consumers (Figure 5.7), than what the Red Book guideline proposes. Only very low income and low income stands, fell within the guideline curves when considering inland users. It could be argued that high income users seldom have such small stand areas. However, these high income users are thought to be possibly high income residential security developments such as Dainfern in the northern suburbs of Johannesburg. From these results it is apparent that income has a significant impact on domestic water consumption and that it is imperative to jointly take income and stand area into account when estimating domestic demand.

Figure 5.9 and Figure 5.10 plot the single variable regression models, with stand value as independent variable, for the stand area categories, on a single graph, for the two geographic locations (inland and coastal) respectively. Once again, as was the case with the stand value categories, the regression models predicted inland water consumption to be consistently higher than coastal consumption for the same stand value and the same stand area category. Appendix C contains separate graphs for each stand area category and there the distinct difference between inland and coastal consumption can be easily compared. The 97.5% confidence intervals for the average AADD are also indicated on the separate graphs in Appendix C.

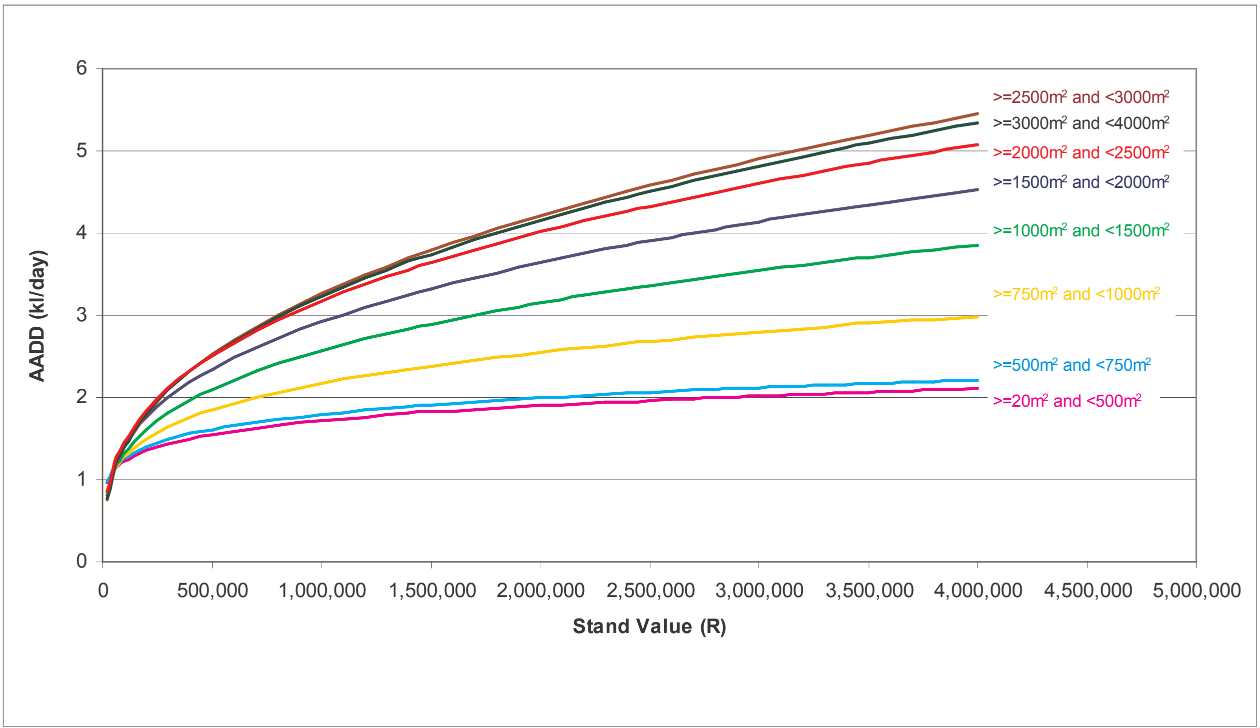


Figure 5-9:Single Variable Regression Analysis (AADD vs. Stand Value) for Categories of Stand Value (Income) and Geographic Location is Inland

The impact of stand area on domestic consumption is evident in Figures 5.9 and 5.10. Larger domestic users with larger stands will consume more water than domestic users with smaller stands for the stand income level (stand value). The only exception was found for extreme large stands (greater than 2500 m^2). For both inland and coastal locations, the regression models predicted that stands with areas between 2500 m^2 and 3000m^2 will consume more water than stands with areas between 3000m^2 and 4000 m^2. This was an unexpected result and a possible reason may be that the sample sizes for the larger stands categories (greater than 2500 m^2) were quite small compared to the other categories and might have assisted in giving this unanticipated result.

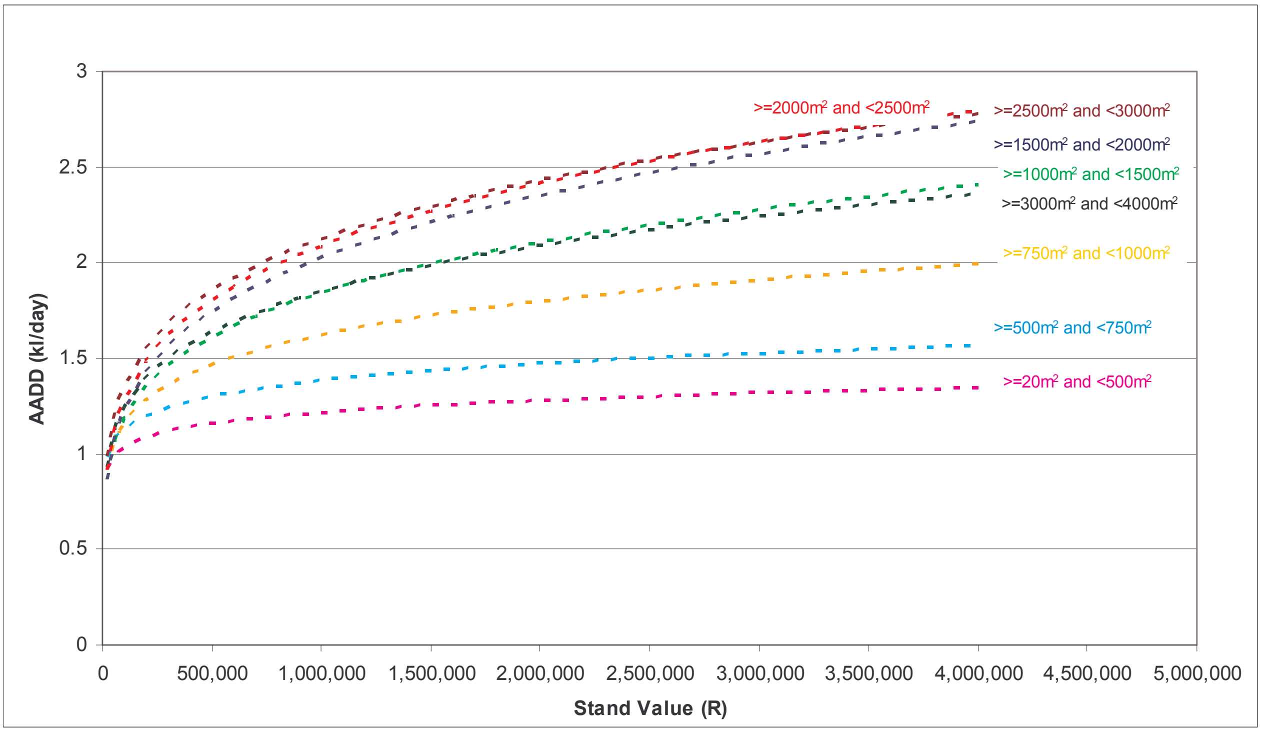


Figure 5-10: Single Variable Regression Analysis (AADD vs. Stand Value) for Categories of Stand Value (Income) and Geographic Location is Coastal

The single variable regression models confirm the outcome of the multiple variable regression models. Both stand area and stand value (as proxy for income level) have a positive effect on domestic water consumption. Geographic location is also an important factor to consider with domestic water demand estimation. The results of this study indicated that coastal stands consistently consume less water than inland stands for the same stand area and value.

5.1.3 Proposal for Revising the Existing Domestic Demand Estimation Guideline

The results of this study confirmed the Red Book (CSIR, 2003) guideline, in that stand area was found to be one of the most significant parameters influencing domestic water demand. Stand value and geographic location were also found to significantly influence domestic water demand. It was considered constructive to investigate possibilities for revising the existing demand estimation guideline, given the results of this study.

A two-variable model with stand area and stand value specified as independent variables could be developed. However, it was considered that information on stand value is generally difficult to obtain and is usually not data that a design engineer would easily have access to if stand value is included as a variable in the proposed revised guideline. Also stand values change over time and would be significantly different from the values used in this study for the exact same suburb. Stand values are also many times based on perception rather than true values.

Another possibility was to develop three income models and to use stand value as a proxy for income as discussed in previous chapters of this study. In other words, categorise all the domestic data in three income categories namely low, middle and high income, based on stand value data and then generate three regression models with only stand area specified as the independent variable. In essence, this was done with the single variable regression analysis for the various stand value categories (discussed in Section 5.1.2 and presented in Figures 5.7 and 5.8). It was felt, however, that the categorisation of stand values into income levels could be considered subjective, when used for the development of a guideline curve. It was therefore deemed that a guideline curve that allows the user to interpret income levels himself, would be better.

It was therefore decided to develop a single guideline curve for AADD vs. Stand Area using all the available domestic data (1 091 685 domestic stands) with design envelopes around it that are based on the 25%, 75% and 95% confidence intervals for the number of datapoints. The curve is the same curve that is presented in Figures 5.2 and 5.3, but the confidence intervals are different. The confidence intervals presented in Figure 5.3 is the intervals for the average AADD. The confidence intervals for the number of datapoints indicate the limits wherein x% of the datapoints would fall. The format of the equation is given below as:

$$y = a + bx \pm (Z\alpha)s\sqrt{1 + \frac{(x - \bar{x})^2}{(N+1)(\sigma)}}$$

Where

a and b = regression coefficients

s = standard error of the estimate

N= number of datapoints

σ = variance

Zα = critical value of the normal distribution

Figure 5.11 presents the regression model using all the available domestic data and the various confidence intervals for the number of datapoints that could be used as an indication for a possible design guideline for domestic use. The equations for the curves are given below.

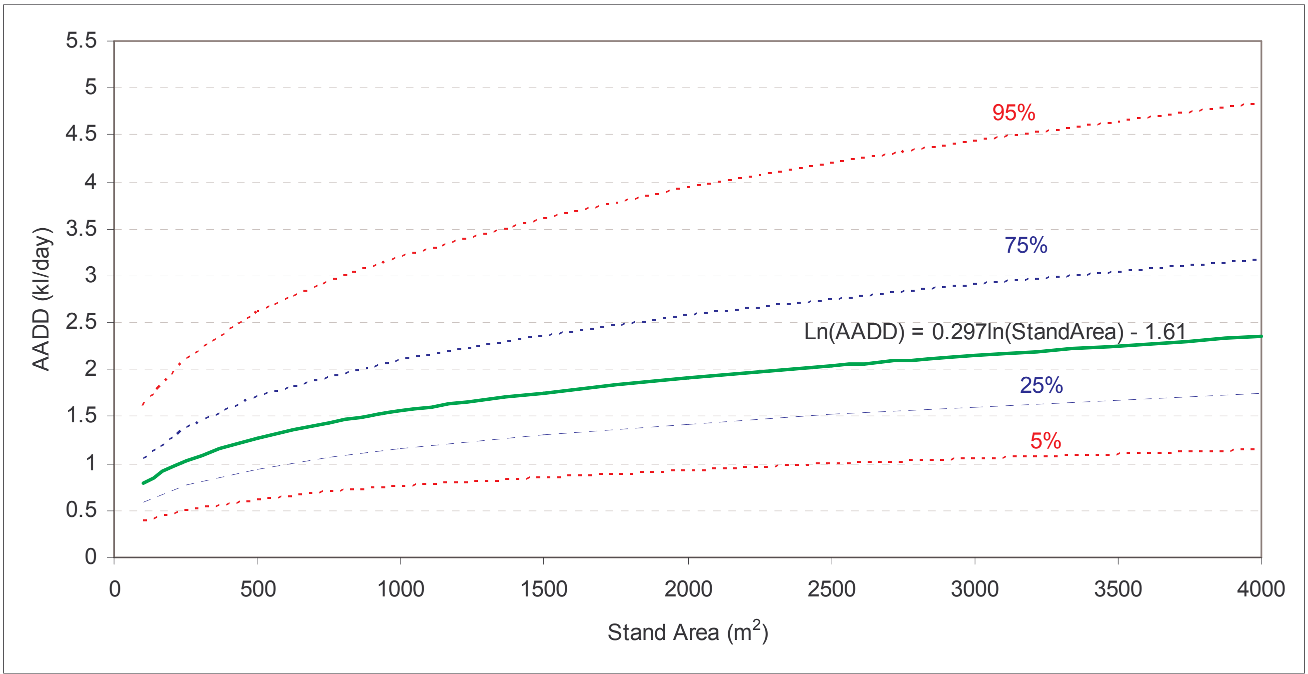


Figure 5-11: Proposal for a Guideline to Estimate Domestic Water Demand with Design Envelopes

<u>95% / 5% Confidence Intervals for the Number of Datapoints:</u>

$$\ln(AADD) = -1.610 + 0.297\ln(StdArea) \pm (1.6449)(0.43865)\sqrt{1+\frac{(\ln StdArea - 6.4124)^2}{666977}}$$

<u>75% /25% Confidence Intervals for the Number of Datapoints:</u>

$$\ln(AADD) = -1.610 + 0.297\ln(StdArea) \pm (0.6745)(0.43865)\sqrt{1+\frac{(\ln StdArea - 6.4124)^2}{666977}}$$

The curve in Figure 5.11 could be used as an indication of the typical domestic water demand in South Africa for varying stand areas and the confidence intervals could be used as design envelopes. The user thus has to consider geographic location and income level in selecting an AADD value within the design envelopes for a given stand area.

Figure 5.12 indicates the proposed new guideline curve and compares it to the existing Red Book (CSIR, 2003) guideline. Figure 5.12 also plots the average AADD and average stand area for each suburb studied.

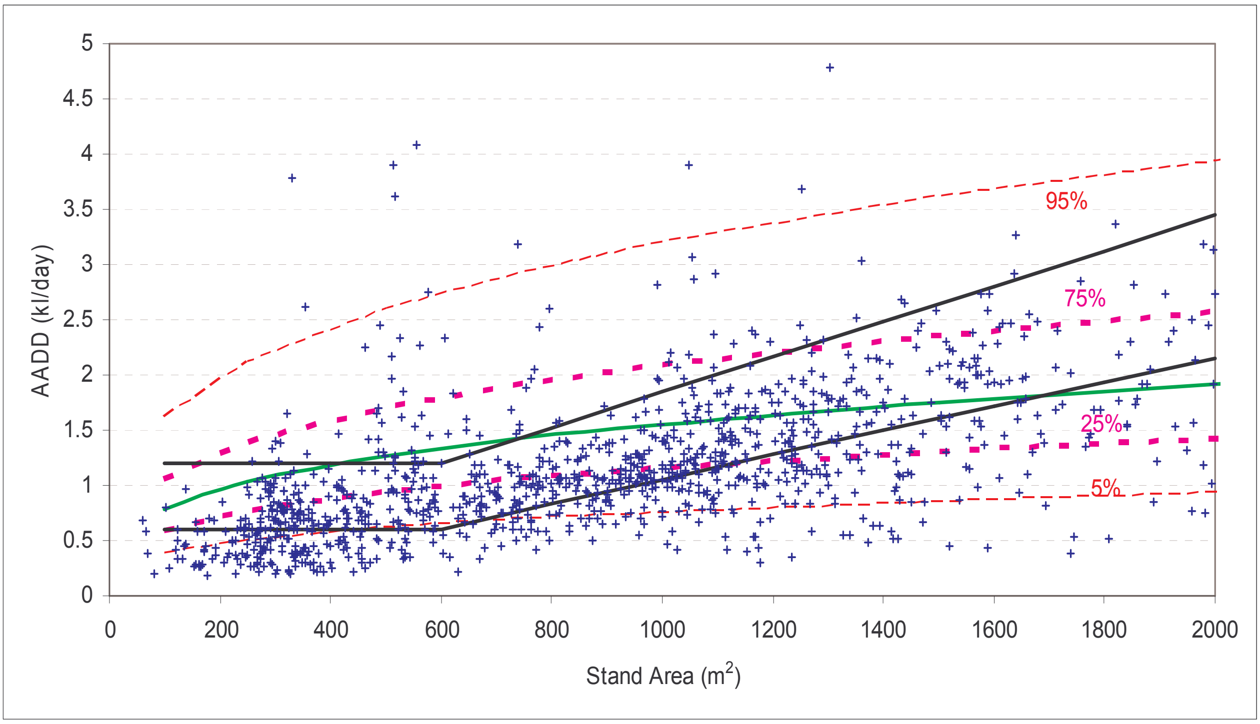


Figure 5-12: Proposed Revision to the Existing Guideline

5.2 Non-Domestic Water Consumption

5.2.1 Frequency Distribution of Non-Domestic Water Consumption Data

As discussed in the methodology (Chapter 4), the first step was to calculate frequency distributions and descriptive statistics for each of the non-domestic datasets to better understand the nature of each dataset. The frequency distributions were calculated for intervals of the natural logarithm of the demand (AADD). The frequency distributions generated for most of the non-domestic user categories displayed a clear log-normal distribution trend with the Business Commercial, Industrial and Farms categories being the most noteworthy. Figures 5.13 to 5.18 show the frequency distributions obtained for ln(AADD) for each of the non-domestic user categories.

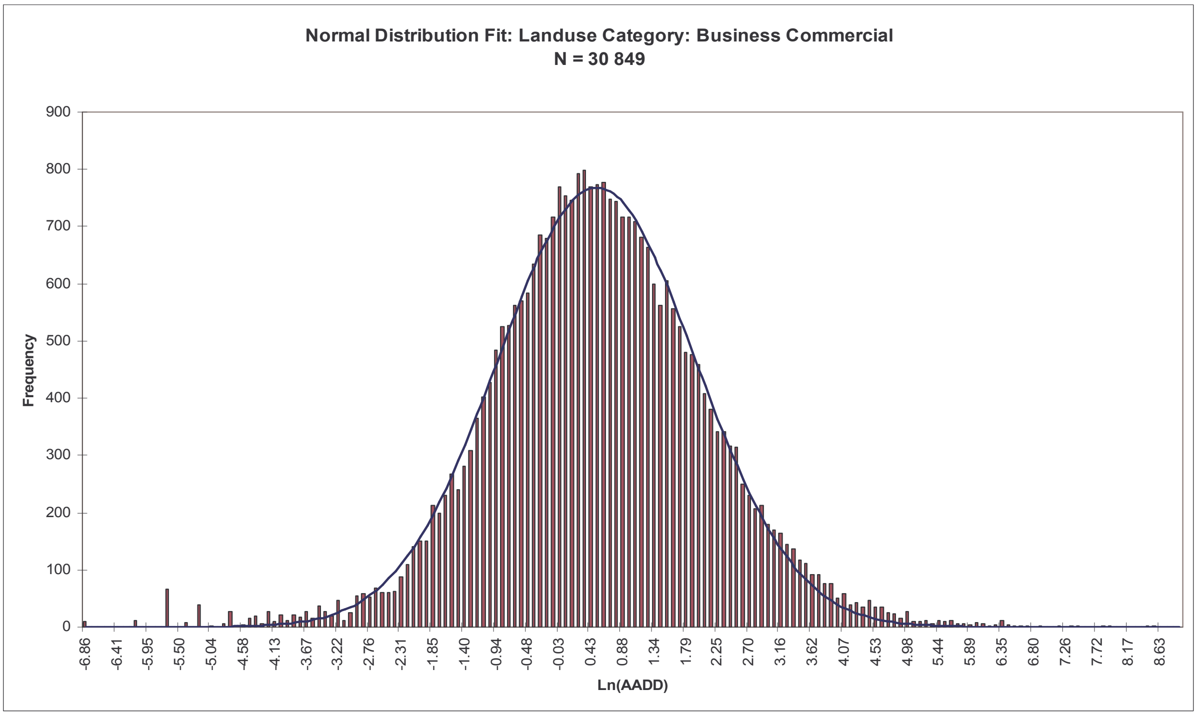


Figure 5-13: Normal Distribution Fit for Frequency Distribution for Business Commercial Users

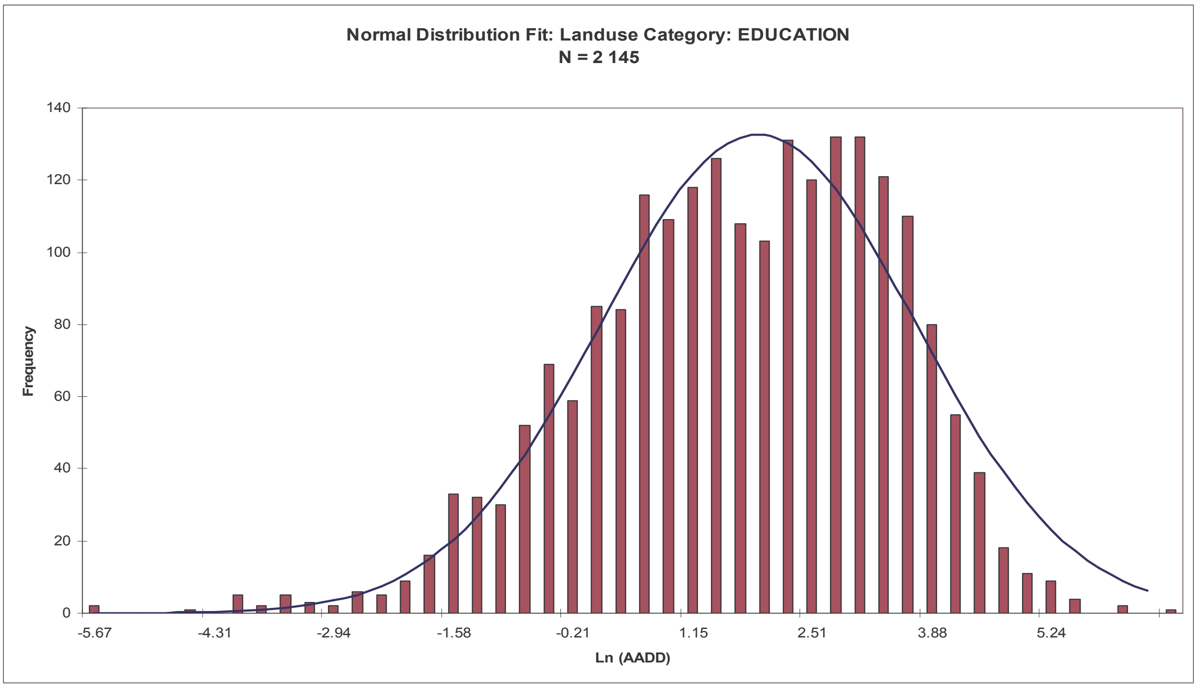


Figure 5-14: Normal Distribution Fit for Frequency Distribution for Educational Users

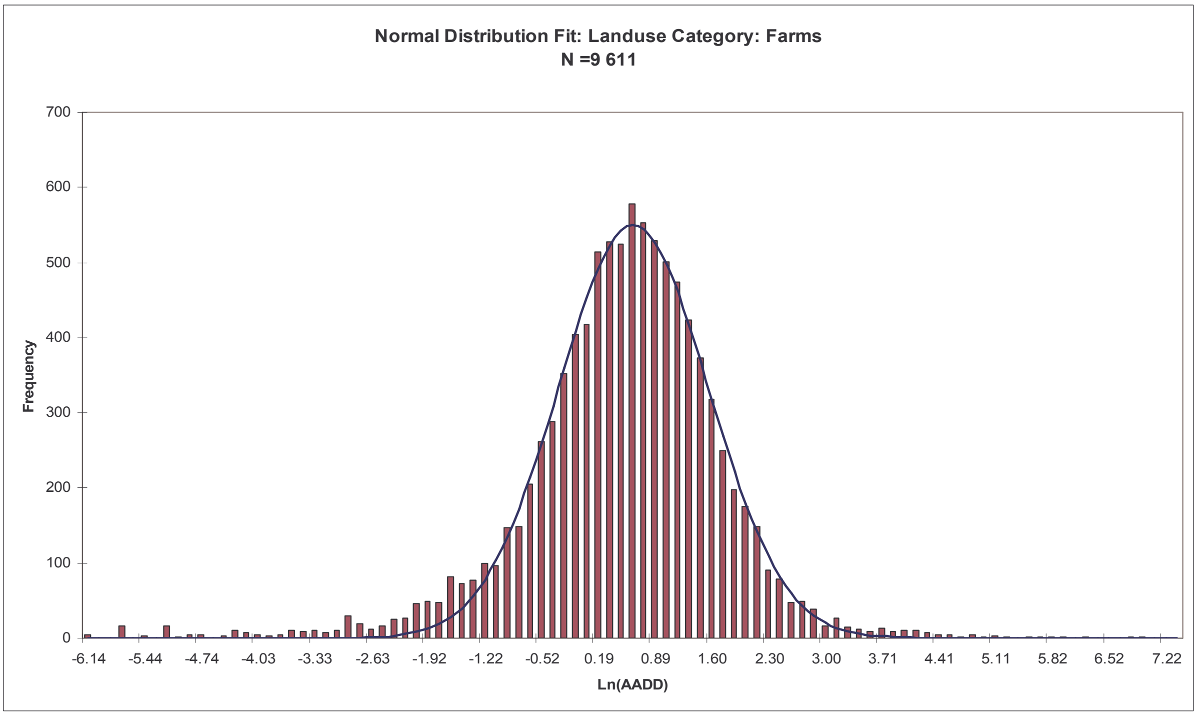


Figure 5-15: Normal Distribution Fit for Frequency Distribution for Farm Users

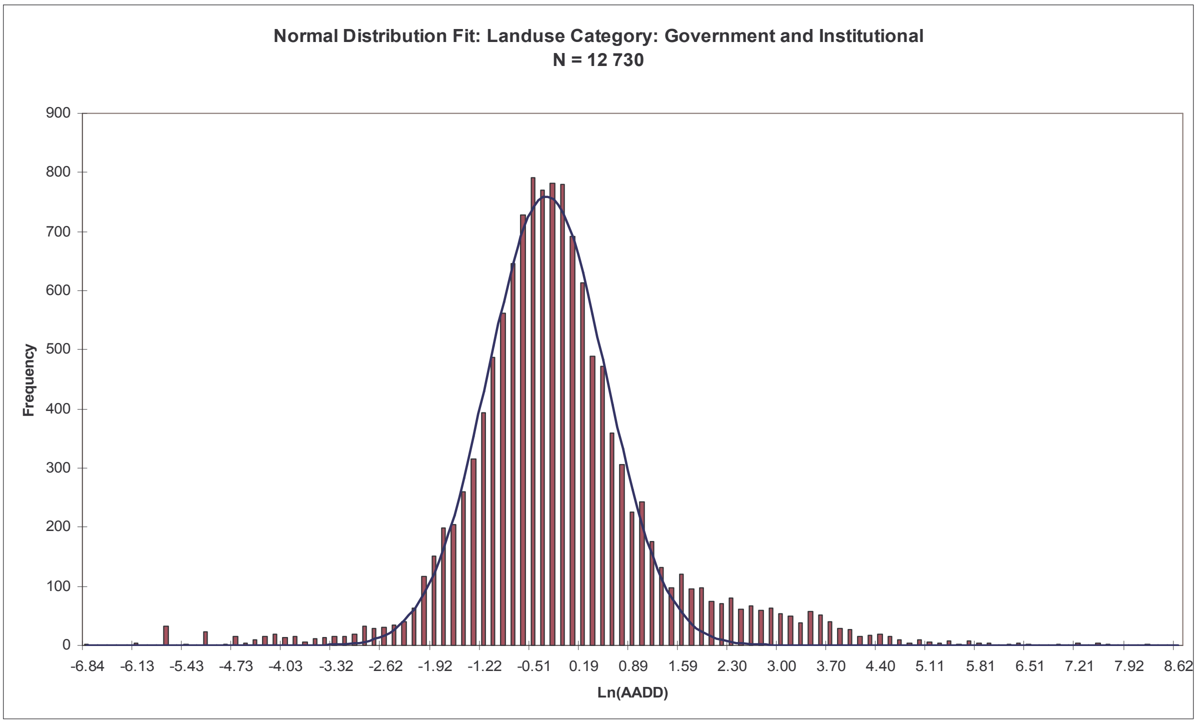


Figure 5-16: Normal Distribution Fit for Frequency Distribution for Government and Institutional Users

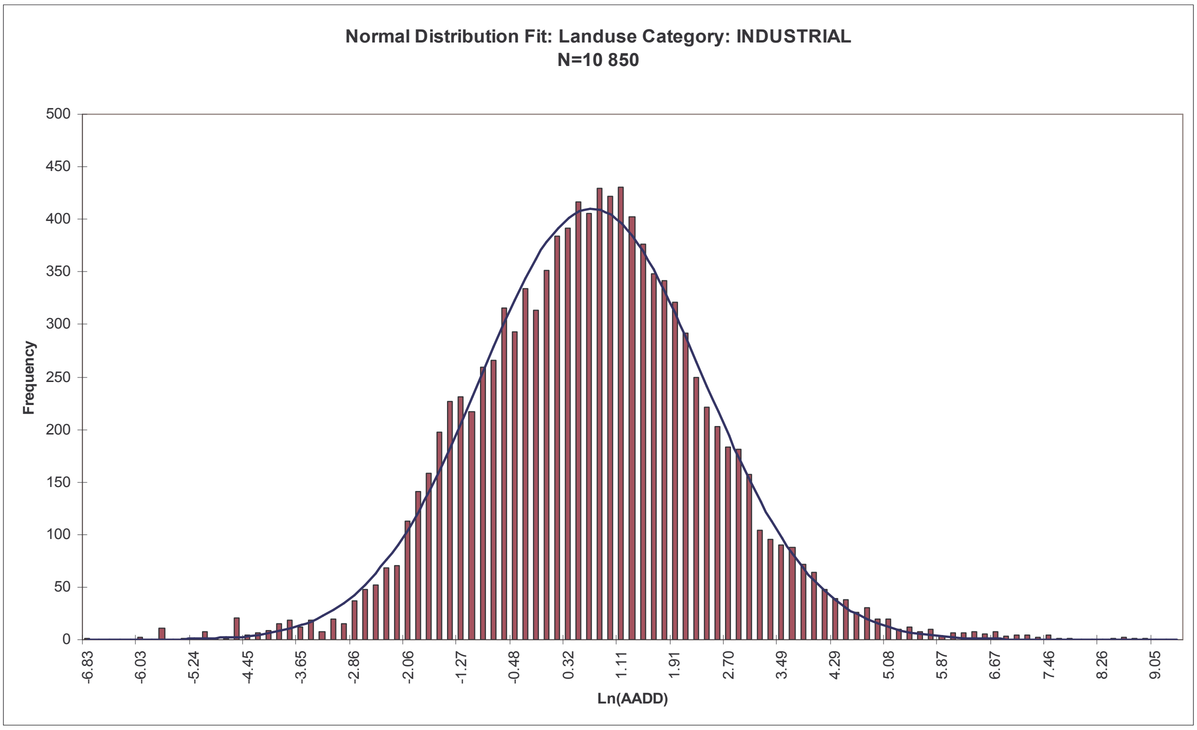


Figure 5-17: Normal Distribution Fit for Frequency Distribution for Industrial Users

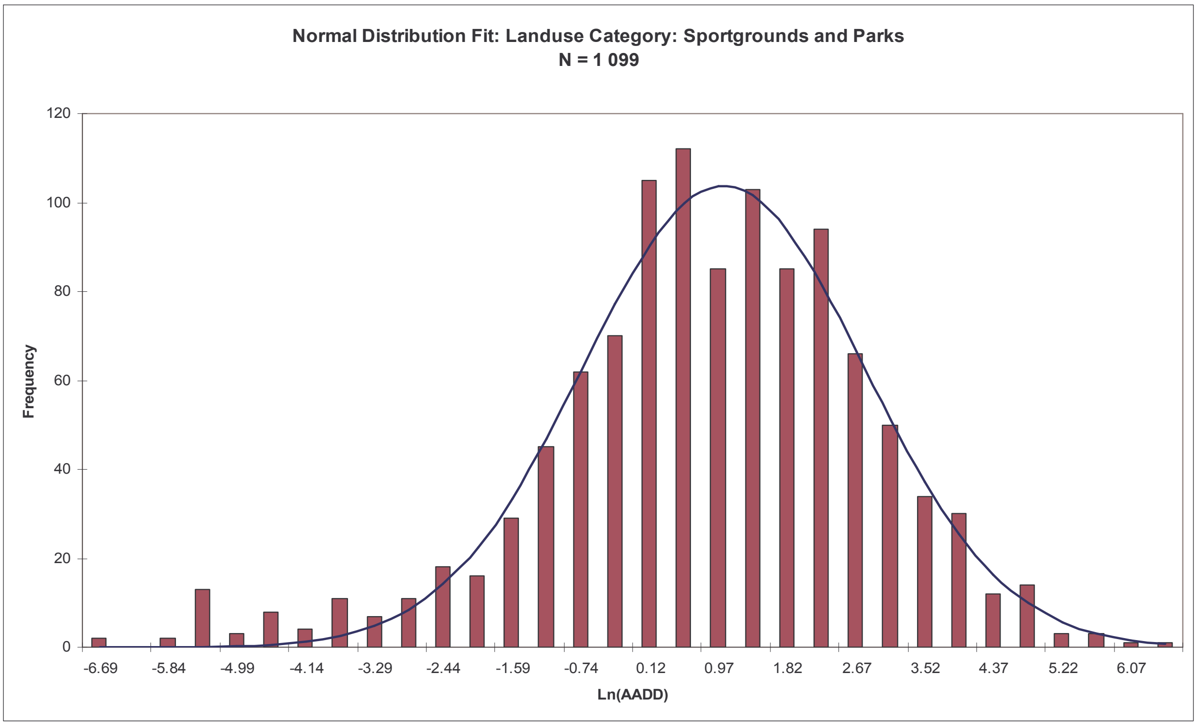


Figure 5-18: Normal Distribution Fit for Frequency Distribution for Sport grounds and Parks

The fits for the Government and Institutional, and Educational categories were less good. A possible reason for this, was that the analysis relied heavily on the land use codes assigned by treasury systems. In both these datasets, it was noted with the data cleaning exercise that the assignment of land use codes was not consistent for all the municipalities. In other words, what one municipality regarded as an educational facility was regarded as a government institution by another. It is possible, that further refinement of the land use codes may improve the normal distribution fits for these two datasets. A direct comparison between the “Red Book” guidelines and the non-domestic user categories of this study was problematic because the user categories are not consistent with the “Red Book” categories. Furthermore, the “Red Book” guideline gives the demand estimate in a unit of gross floor area (m^2) and therefore the mean values of the log-normal distributions could not be compared directly.

The clear log-normal trend, observed in the distributions of the non-domestic user categories, led to the use of logarithmic regression models during further analysis of the data.

5.2.2 Assessment of Factors Influencing Non-Domestic Water Consumption

There are seven non-domestic user categories and a dataset for each category. On each of these datasets, a stepwise multiple variable regression analysis was performed. Table 5.6 below summarises the outcome of the stepwise regression analyses on each category by listing the independent variables influencing water consumption in each category by order of best correlation with the AADD. The number of data points in each analysis is also indicated and the adjusted R-square value of the each stepwise regression model. The detailed regression statistics are contained in Appendix C, including the regression coefficients for each regression model (equation).

Table 5-6: Summary of the Outcome of the Stepwise Multiple Variable Regression Analysis of Non-Domestic Users Categories

Category	Number of data points	Number of Variables	Independent Variables by order of Best Correlation with ln (AADD)	Adjusted R^2
Business Commercial	24 810	8	LN(Stand Value)	0.23
			LN(Stand Area)	0.283
			Geographic Location	0.303
			Stand Value	0.31
			Development Level	0.317
			Ave Max Temperature	0.326
			Stand Area	0.33
			Ave Min Temperature	0.332
Education	1 447	5	Ln(Stand Area)	0.408
			Development Level	0.422
			Stand Value	0.437
			Ave Max Temperature	0.455
			Ln(Stand Value)	0.458
Farms	8 895	6	LN(Stand Value)	0.096
			Ave Min Temperature	0.113
			Stand Area	0.118
			Stand Value	0.121
			Mean Annual Evaporation	0.122
			Mean Annual Precipitation	0.122
Government and	8 259	8	Ln(Stand Area)	0.336
			LN(Stand Value)	0.373

Category	Number of data points	Number of Variables	Independent Variables by order of Best Correlation with ln (AADD)	Adjusted R^2
Institutional			Stand Value	0.387
			Geographic Location	0.392
			Development Level	0.4
			Stand Area	0.401
			Ave Max Temperature	0.401
			Mean Annual Precipitation	0.401
Industrial	8 000	8	Ln(Stand Area)	0.313
			Stand value	0.33
			Geographic Location	0.345
			Mean Annual Precipitation	0.35
			Stand Area	0.354
			Ln(StandValue)	0.36
			Ave Max Temperature	0.362
			Development Level	0.364
Parks	367	4	Ln(Stand Area)	0.137
			Ave Min Temperature	0.176
			Development Level	0.191
			Stand Value	0.207
Sports	68	2	Ave Min Temperature	0.083
			Ln(Stand Area)	0.13

As the regression statistics indicate (adjusted R-square values, P-values and t-values) strong correlations were found between the AADD and the various parameters listed in Table 5.6. The adjusted R-square values were better on the whole, than those obtained with the multiple variable regression analyses for domestic use. A possible reason may be that the non-domestic sets were significantly smaller than the domestic datasets. The Sports category is quite a small dataset and may be too small to provide statistically significant results.

Stand area emerged as the factor influencing water consumption most for four of the seven non-domestic categories namely Education, Government and Institutional, Industrial and Parks. Figure 5.19 depicts the regression equations resulting from the first step of the stepwise regression analyses for these four categories. The results made sense, in that compared to the other categories, Industrial users consumed more water for the same stand area. Education users consume less than

Industrial users but significantly more than Government and Institutional users and Parks. Parks consume less than any of the other three for the same stand area.

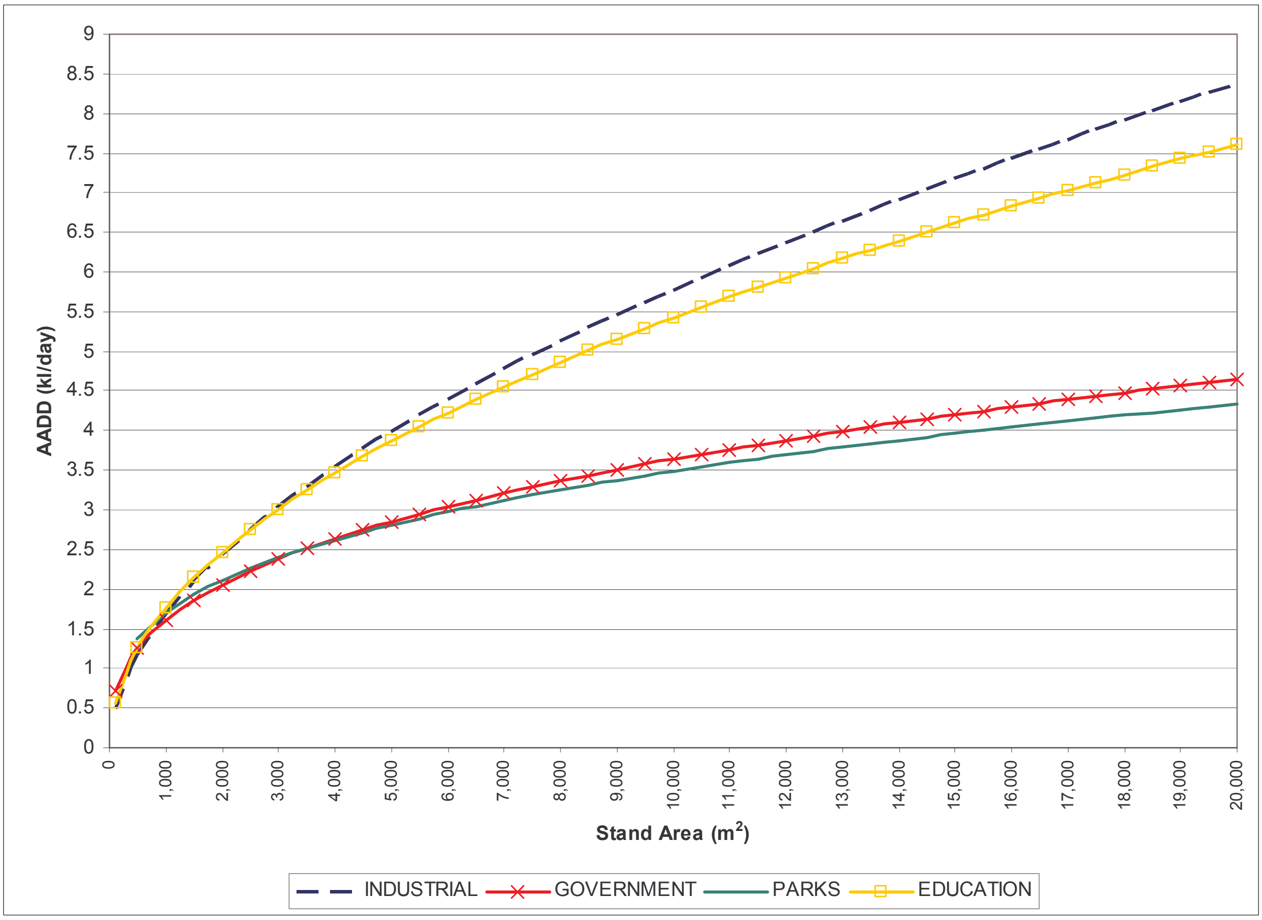


Figure 5-19: Comparison of the First Step of the Stepwise Multiple Variable Regression Models for Industrial, Education, Government and Parks User Categories

Stand value was the most significant variable influencing business commercial and farm water consumption. It must be kept in mind that the farms category might have contained agricultural smallholdings, which is in a sense, domestic water use rather than non-domestic water use. However, there was no clear and objective way to distinguish between small holdings and working farms in the dataset thus the category was analysed as is. Figure 5.20 below portrays Business Commercial AADD and Farm AADD against stand value as predicted by the first-step model of the multi-variable regression analyses. The regression equations predict that business commercial users and farms consume approximately the same amount of water especially for higher stand values. At lower stand values, farms consume more water than business commercial users.

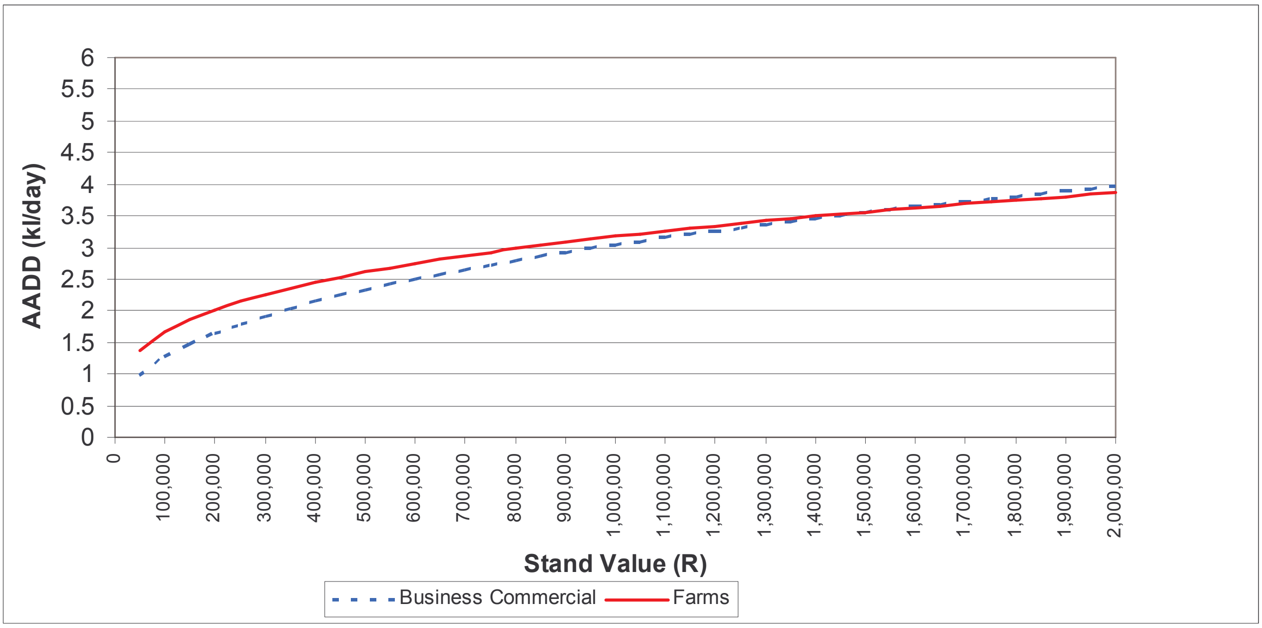


Figure 5-20: Comparison of the First Step of the Stepwise Multiple Variable Regression Analysis for Business Commercial and Farms User Categories

It should be noted that, in contrast to the domestic water consumption analysis, in the non-domestic consumption analysis, stand value was not considered as a proxy for income, but the stand values were used as an indication of the true monetary value of the sites.

Geographic location (i.e. inland or coastal) did not have the marked impact on non-domestic use as was the case with domestic use. For only three categories namely, business commercial, government and institutions and industrial, did geographic location have a significant impact on the water consumption. The development level (city vs. small town) however, came out as a factor affecting water consumption for seven of the eight user categories.

As discussed in Chapter 4 (Figure 4.2), the next step in the data analysis was to perform single variable regressions for stand area, stand value, geographic location and development level. A summary of the regression equations for the average of all demands with the 97.5 % confidence intervals are contained in tabular format below. The detail regression statistics are contained in Appendix C.

Table 5-7: Single Variable Regression Results for Non-Domestic User Categories with Distinction Made for Development Level and Geographic Location

Category	Development Level	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
Business Commercial	Metro/City	Inland	Stand Area	$\ln(AADD) = 0.435\ln(StdArea) - 2.260 \pm 1.77\sqrt{5.1\times10^{-5} + \frac{(\ln StdArea - 7.248)^2}{26022}}$	19,610	0.236
			Stand Value	$\ln(AADD) = 0.409\ln(StdValue) - 4.427 \pm 1.74\sqrt{5.1\times10^{-5} + \frac{(\ln StdValue - 13.032)^2}{32670}}$	19,610	0.260
		Coastal	Stand Area	$\ln(AADD) = 0.501\ln(StdArea) - 2.778 \pm 1.64\sqrt{4.6\times10^{-4} + \frac{(\ln StdArea - 7.076)^2}{1984}}$	2,192	0.244
			Stand Value	$\ln(AADD) = 0.401\ln(StdValue) - 4.547 \pm 1.627\sqrt{4.6\times10^{-4} + \frac{(\ln StdValue - 13.247)^2}{3279}}$	2,192	0.258
	Rural/Small Town	Inland	Stand Area	$\ln(AADD) = 0.312\ln(StdArea) - 2.382 \pm 1.595\sqrt{5.1\times10^{-4} + \frac{(\ln StdArea - 7.328)^2}{2042}}$	1,973	0.209
			Stand Value	$\ln(AADD) = 0.386\ln(StdValue) - 4.328 \pm 1.6\sqrt{5.1\times10^{-4} + \frac{(\ln StdValue - 12.860)^2}{2302}}$	1,973	0.207
		Coastal	Stand Area	$\ln(AADD) = 0.340\ln(StdArea) - 2.002 \pm 1.53\sqrt{3.8\times10^{-4} + \frac{(\ln StdArea - 7.321)^2}{2268}}$	2,613	0.141
			Stand Value	$\ln(AADD) = 0.319\ln(StdValue) - 3.814 \pm 1.5\sqrt{3.8\times10^{-4} + \frac{(\ln StdValue - 13.478)^2}{3269}}$	2,613	0.179
Education	Metro/City	Inland	Stand Area	$\ln(AADD) = 0.487\ln(StdArea) - 2.739 \pm 1.825\sqrt{1.1\times10^{-3} + \frac{(\ln StdArea - 8.54)^2}{2297}}$	942	0.400
			Stand Value	$\ln(AADD) = 0.466\ln(StdValue) - 4.786 \pm 2.03\sqrt{1.1\times10^{-3} + \frac{(\ln StdValue - 13.33)^2}{1600}}$	942	0.254
		Coastal	Stand Area	$\ln(AADD) = 0.502\ln(StdArea) - 2.766 \pm 2.156\sqrt{5.5\times10^{-3} + \frac{(\ln StdArea - 9.396)^2}{335}}$	181	0.277
			Stand Value	$\ln(AADD) = 0.469\ln(StdValue) - 4.750 \pm 2.156\sqrt{5.5\times10^{-3} + \frac{(\ln StdValue - 14.292)^2}{375}}$	181	0.270
	Rural/Small Town	Inland	Stand Area	$\ln(AADD) = 0.531\ln(StdArea) - 3.338 \pm 1.67\sqrt{4.4\times10^{-3} + \frac{(\ln StdArea - 8.19)^2}{742}}$	229	0.556

Category	Development Level	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
			Stand Value	$\ln(AADD) = 0.543\ln(StdValue) - 6.0 \pm 1.65\sqrt{4.4\times10^{-3} + \frac{(\ln StdValue - 12.919)^2}{725}}$	229	0.567
		Coastal	Stand Area	$\ln(AADD) = 0.389\ln(StdArea) - 2.337 \pm 2.08\sqrt{8.6\times10^{-3} + \frac{(\ln StdArea - 9.382)^2}{323}}$	116	0.267
			Stand Value	$\ln(AADD) = 0.362\ln(StdValue) - 3.864 \pm 2.136\sqrt{8.6\times10^{-3} + \frac{(\ln StdValue - 14.308)^2}{315}}$	116	0.224
Farms	Metro/City	Inland	Stand Area	$\ln(AADD) = -0.065\ln(StdArea) + 1.404 \pm 1.41\sqrt{1.2\times10^{-4} + \frac{(\ln StdArea - 9.787)^2}{3561}}$	8,075	0.003
			Stand Value	$\ln(AADD) = 0.245\ln(StdValue) - 2.298 \pm 1.362\sqrt{1.2\times10^{-4} + \frac{(\ln StdValue - 12.520)^2}{5385}}$	8,085	0.076
		Coastal	Stand Area	$\ln(AADD) = 0.357\ln(StdArea) - 2.559 \pm 1.81\sqrt{3.4\times10^{-3} + \frac{(\ln StdArea - 9.815)^2}{659}}$	298	0.247
			Stand Value	$\ln(AADD) = 0.710\ln(StdValue) - 7.504 \pm 1.7\sqrt{3.3\times10^{-3} + \frac{(\ln StdValue - 11.906)^2}{221}}$	303	0.327
	Rural/Small Town	Inland	Stand Area	$\ln(AADD) = 0.373\ln(StdArea) - 2.449 \pm 2.43\sqrt{7.8\times10^{-3} + \frac{(\ln StdArea - 10.319)^2}{372}}$	128	0.204
			Stand Value	$\ln(AADD) = 0.604\ln(StdValue) - 6.562 \pm 2.39\sqrt{7.6\times10^{-3} + \frac{(\ln StdValue - 13.08)^2}{156}}$	132	-.220
		Coastal	Stand Area	$\ln(AADD) = 0.149\ln(StdArea) - 0.855 \pm 1.92\sqrt{2.5\times10^{-3} + \frac{(\ln StdArea - 11.568)^2}{2082}}$	402	0.105
			Stand Value	$\ln(AADD) = 0.437\ln(StdValue) - 4.745 \pm 1.86\sqrt{2.4\times10^{-3} + \frac{(\ln StdValue - 12.817)^2}{347.3}}$	411	0.150
Government and Institutional	Metro/City	Inland	Stand Area	$\ln(AADD) = 0.345\ln(StdArea) - 1.843 \pm 1.815\sqrt{4.8\times10^{-4} + \frac{(\ln StdArea - 7.207)^2}{5150}}$	2,081	0.255
			Stand Value	$\ln(AADD) = 0.424\ln(StdValue) - 4.429 \pm 1.77\sqrt{4.8\times10^{-4} + \frac{(\ln StdValue - 11.956)^2}{3842}}$	2,081	0.288

Category	Development Level	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
		Coastal	Stand Area	$\ln(AADD) = 0.394\ln(StdArea) - 2.331 \pm 1.5\sqrt{1.4\times10^{-3} + \frac{(\ln StdArea - 7)^2}{1604}}$	720	0.370
			Stand Value	$\ln(AADD) = 0.395\ln(StdValue) - 4.361 \pm 1.57\sqrt{1.4\times10^{-3} + \frac{(\ln StdValue - 12.129)^2}{1341}}$	720	0.310
	Rural/Small Town	Inland	Stand Area	$\ln(AADD) = 0.336\ln(StdArea) - 1.881 \pm 0.853\sqrt{1.6\times10^{-4} + \frac{(\ln StdArea - 6.28)^2}{5307}}$	6,149	0.34
			Stand Value	$\ln(AADD) = 0.3\ln(StdValue) - 3.231 \pm 0.851\sqrt{1.6\times10^{-4} + \frac{(\ln StdValue - 11.553)^2}{1806}}$	6,419	0.343
		Coastal	Stand Area	$\ln(AADD) = 0.405\ln(StdArea) - 2.548 \pm 1.715\sqrt{9.2\times10^{-3} + \frac{(\ln StdArea - 7.9)^2}{214}}$	109	0.291
			Stand Value	$\ln(AADD) = 0.311\ln(StdValue) - 3.475 \pm 1.856\sqrt{9.2\times10^{-3} + \frac{(\ln StdValue - 13.283)^2}{218}}$	109	0.171
Industrial	Metro/City	Inland	Stand Area	$\ln(AADD) = 0.509\ln(StdArea) - 2.882 \pm 1.817\sqrt{1.9\times10^{-4} + \frac{(\ln StdArea - 7.892)^2}{7015}}$	5,347	0.283
			Stand Value	$\ln(AADD) = 0.406\ln(StdValue) - 4.285 \pm 1.93\sqrt{1.9\times10^{-4} + \frac{(\ln StdValue - 13.336)^2}{7566}}$	5,347	0.194
		Coastal	Stand Area	$\ln(AADD) = 0.648\ln(StdArea) - 4.159 \pm 1.8\sqrt{5.8\times10^{-4} + \frac{(\ln StdArea - 7.91)^2}{1648}}$	1,720	0.324
			Stand Value	$\ln(AADD) = 0.336\ln(StdValue) - 3.643 \pm 2.03\sqrt{5.8\times10^{-4} + \frac{(\ln StdValue - 13.712)^2}{2547}}$	1,720	0.134
	Rural/Small Town	Inland	Stand Area	$\ln(AADD) = 0.587\ln(StdArea) - 3.775 \pm 2.13\sqrt{3.7\times10^{-3} + \frac{(\ln StdArea - 7.849)^2}{514}}$	269	0.357
			Stand Value	$\ln(AADD) = 0.656\ln(StdValue) - 7.623 \pm 2.05\sqrt{3.7\times10^{-3} + \frac{(\ln StdValue - 12.895)^2}{462}}$	269	0.401
		Coastal	Stand Area	$\ln(AADD) = 0.33\ln(StdArea) - 2.088 \pm 1.46\sqrt{1.3\times10^{-3} + \frac{(\ln StdArea - 7.13)^2}{1426}}$	799	0.259

Category	Development Level	Geographic Location	Independent Variable	Equation	N	Adjusted R^2
			Stand Value	$\ln(AADD) = 0.284\ln(StdValue) - 3.232 \pm 1.46\sqrt{1.3\times10^{-3} + \frac{(\ln StdValue - 12.339)^2}{1904}}$	799	0.255
Parks	Metro/City	Inland	Stand Area	$\ln(AADD) = 0.284\ln(StdArea) - 1.325 \pm 2.319\sqrt{5.3\times10^{-3} + \frac{(\ln StdArea - 8.930)^2}{420}}$	190	0.109
			Stand Value	$\ln(AADD) = 0.205\ln(StdValue) - 1.383 \pm 2.385\sqrt{4.5\times10^{-3} + \frac{(\ln StdValue - 12.262)^2}{468}}$ #	223	0.052
		Coastal	Stand Area	$\ln(AADD) = 0.289\ln(StdArea) - 1.3 \pm 2.18\sqrt{7.6\times10^{-3} + \frac{(\ln StdArea - 8.462)^2}{206}}$ #	131	0.09
			Stand Value	$\ln(AADD) = 0.211\ln(StdValue) - 1.542 \pm 2.134\sqrt{6\times10^{-3} + \frac{(\ln StdValue - 12.063)^2}{360}}$	167	0.07
	Rural/Small Town	Inland	Stand Area	$\ln(AADD) = 0.328\ln(StdArea) - 2.335 \pm 1.24\sqrt{1.96\times10^{-2} + \frac{(\ln StdArea - 8.21)^2}{90.73}}$	51	0.314
			Stand Value	$\ln(AADD) = 0.260\ln(StdValue) - 2.822 \pm 1.39\sqrt{1.89\times10^{-2} + \frac{(\ln StdValue - 12.157)^2}{72.13}}$	53	0.142
		Coastal	Stand Area	Sample size too small to be statistically significant	3	n/a
			Stand Value	Sample size too small to be statistically significant	5	n/a
Sports	Metro/City	Inland	Stand Area	$\ln(AADD) = 0.193\ln(StdArea) - 0.6 \pm 2.13\sqrt{1.72\times10^{-2} + \frac{(\ln StdArea - 8.642)^2}{140.24}}$	58	0.055
			Stand Value	$\ln(AADD) = 0.224\ln(StdValue) - 1.877 \pm 2.16\sqrt{1.72\times10^{-2} + \frac{(\ln StdValue - 13.145)^2}{70.7}}$ #	58	0.032
		Coastal	Stand Area	Sample size too small to be statistically significant	8	n/a
			Stand Value	Sample size too small to be statistically significant	8	n/a
	Rural/Small Town	Inland	Stand Area	Sample size too small to be statistically significant	3	n/a
			Stand Value	Sample size too small to be statistically significant	3	n/a
		Coastal	Stand Area	Sample size too small to be statistically significant	1	n/a
			Stand Value	Sample size too small to be statistically significant	1	n/a

 # Equation coefficients do not differ significantly from zero.

The adjusted R-square values for most of the regression equations indicate strong correlations between the independent variables (stand area and stand value) and the AADD, considering that an adjusted R-square value of more than 20 % is considered good, when predicting human behaviour. The t-statistics and P-values for all the regression coefficients were evaluated to ensure that the coefficients differ significantly from zero. Where this was not the case, it is indicated in Table 5.7. These regression models with insignificant coefficients were omitted from any further analysis. The number of data points (or records) in some categories for the Parks and Sports datasets were considered to be insufficient to provide statistically significant regression models. The cut-off point for the minimum number of data points used in this study was 50 data points.

Figures 5.21 to 5.33 represent the regression models for each non-domestic user category. Appendix C contains separate graphs for each model with 97.5 % confidence intervals for the average AADD indicated.

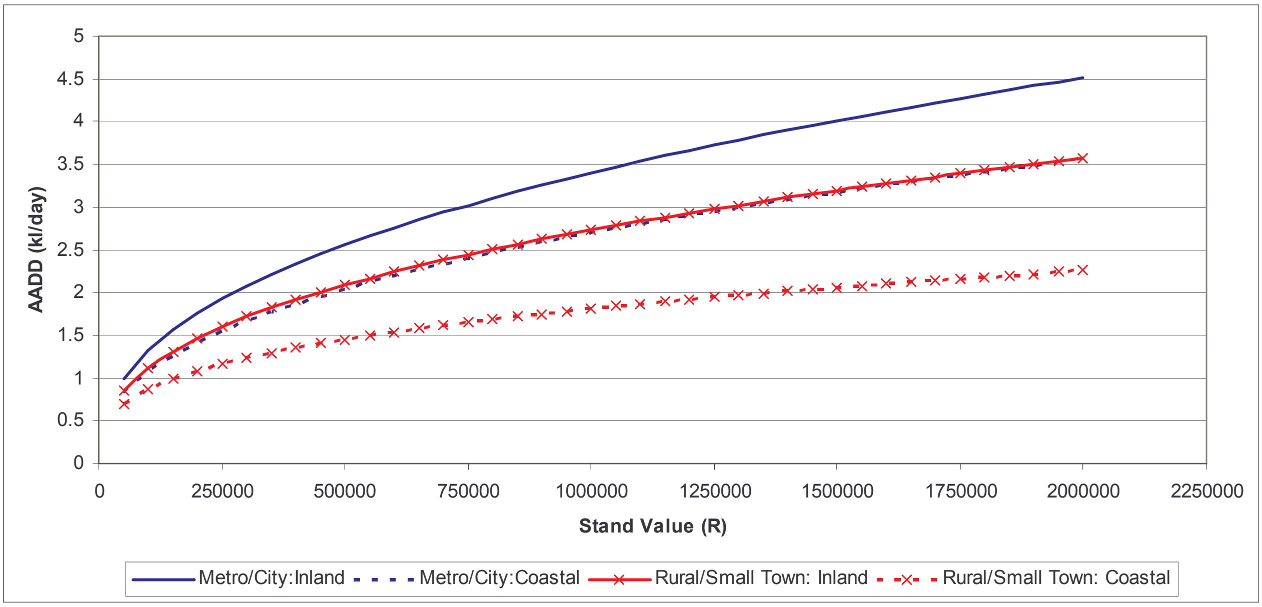


Figure 5-21: Single Variable Regression Results: AADD vs. Stand Value for Business Commercial Users

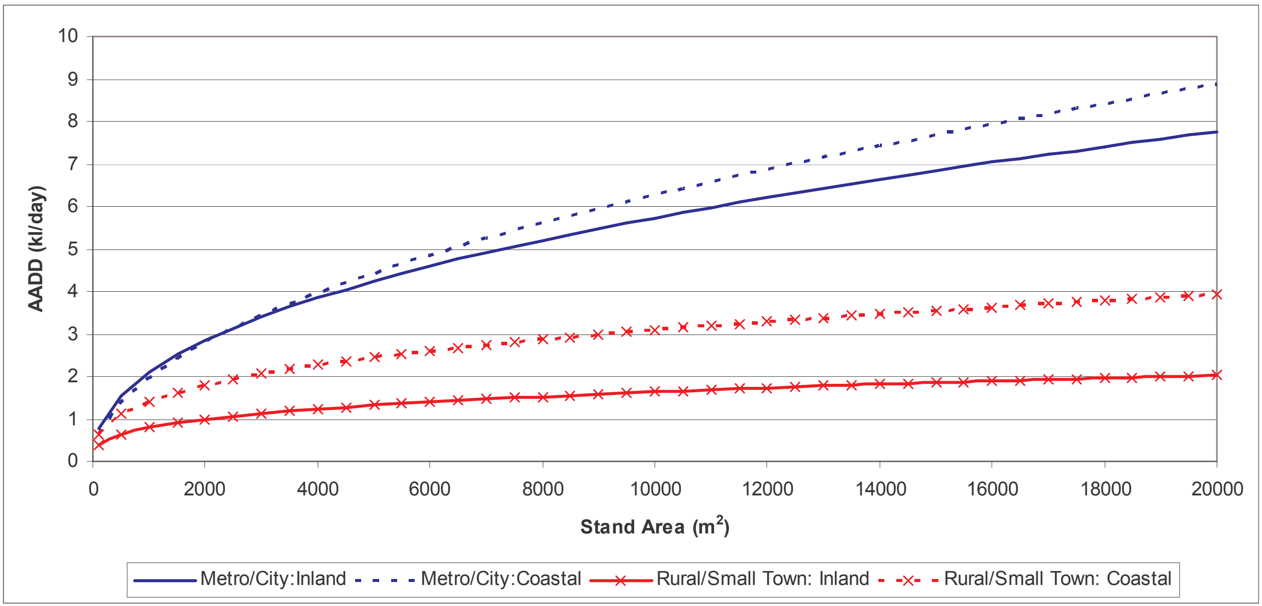


Figure 5-22: Single Variable Regression Results: AADD vs. Stand Area for Business Commercial Users

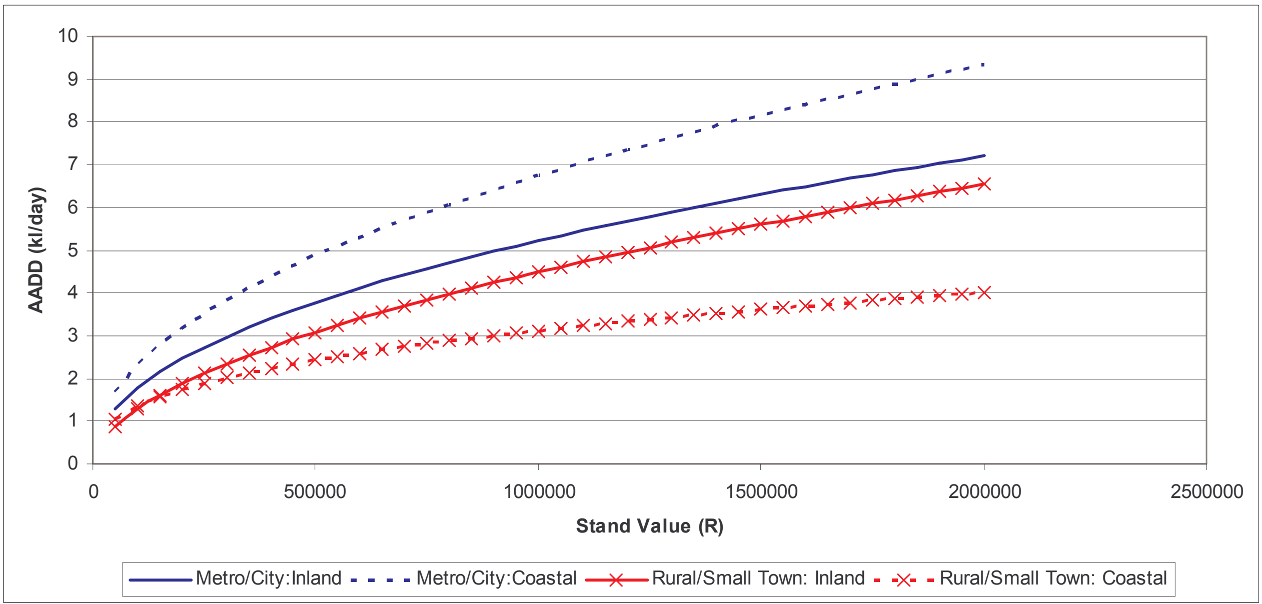


Figure 5-23: Single Variable Regression Results: AADD vs. Stand Value for Educational Users

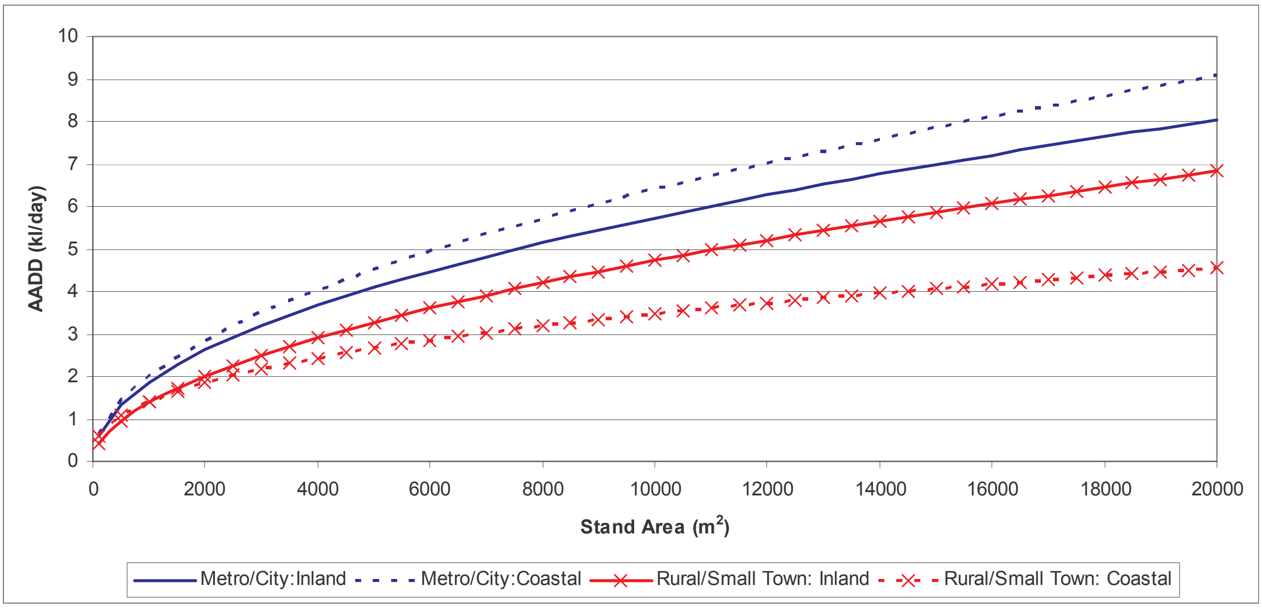


Figure 5-24: Single Variable Regression Results: AADD vs. Stand Area for Educational Users

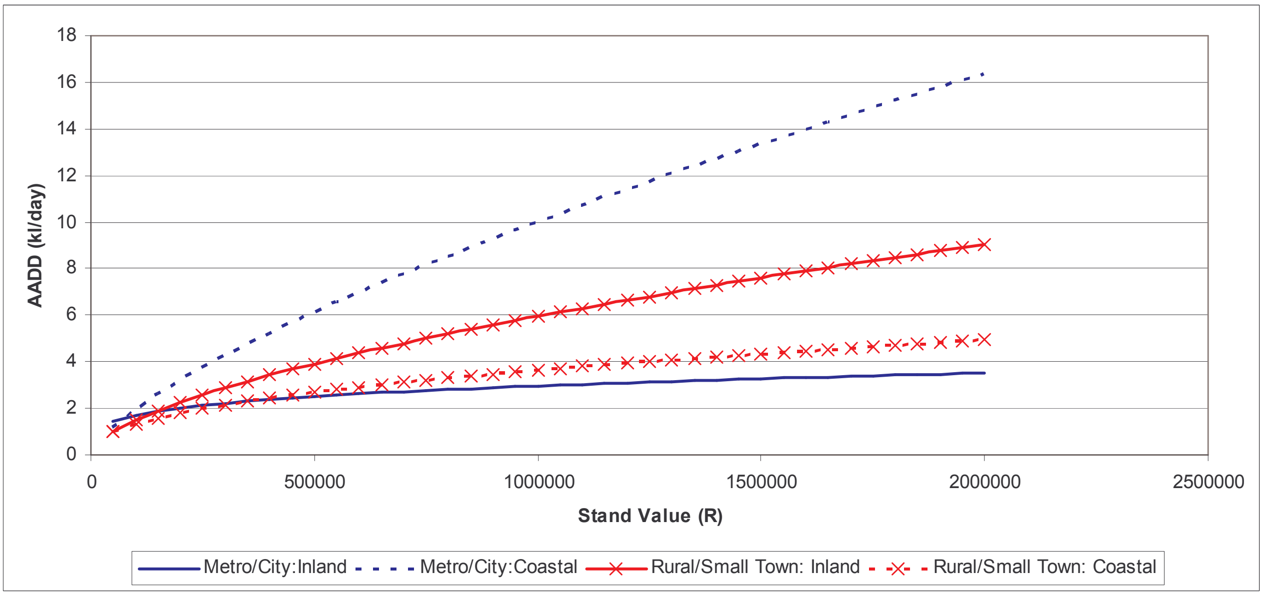


Figure 5-25: Single Variable Regression Results: AADD vs. Stand Value for Farm Users

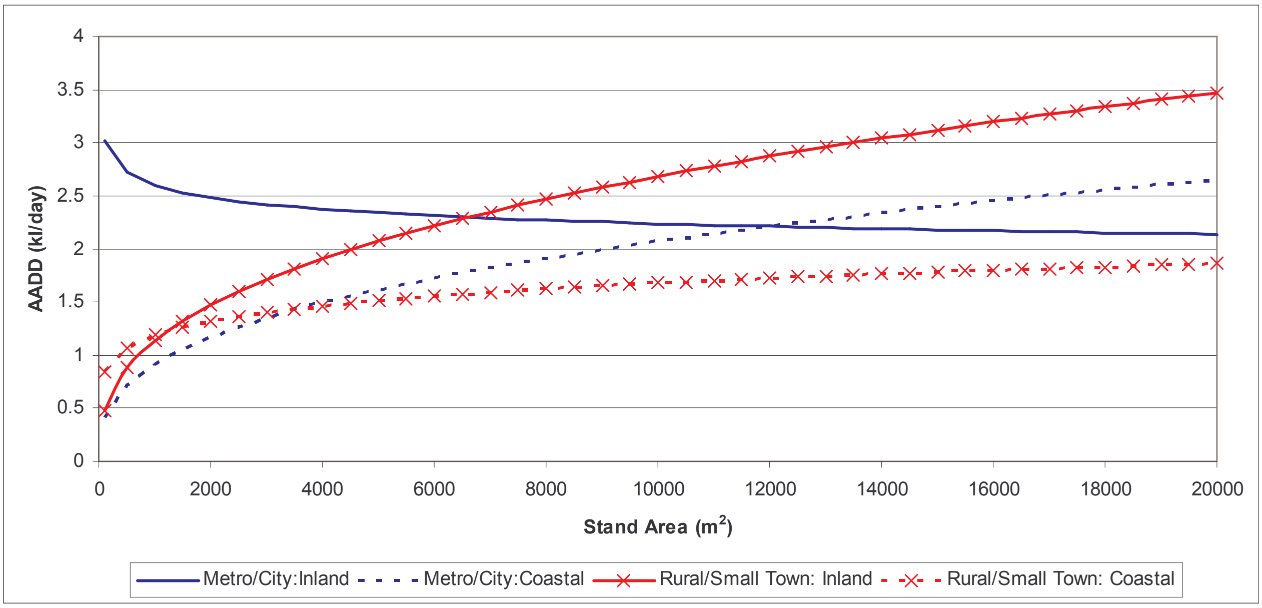


Figure 5-26: Single Variable Regression Results: AADD vs. Stand Area for Farm Users

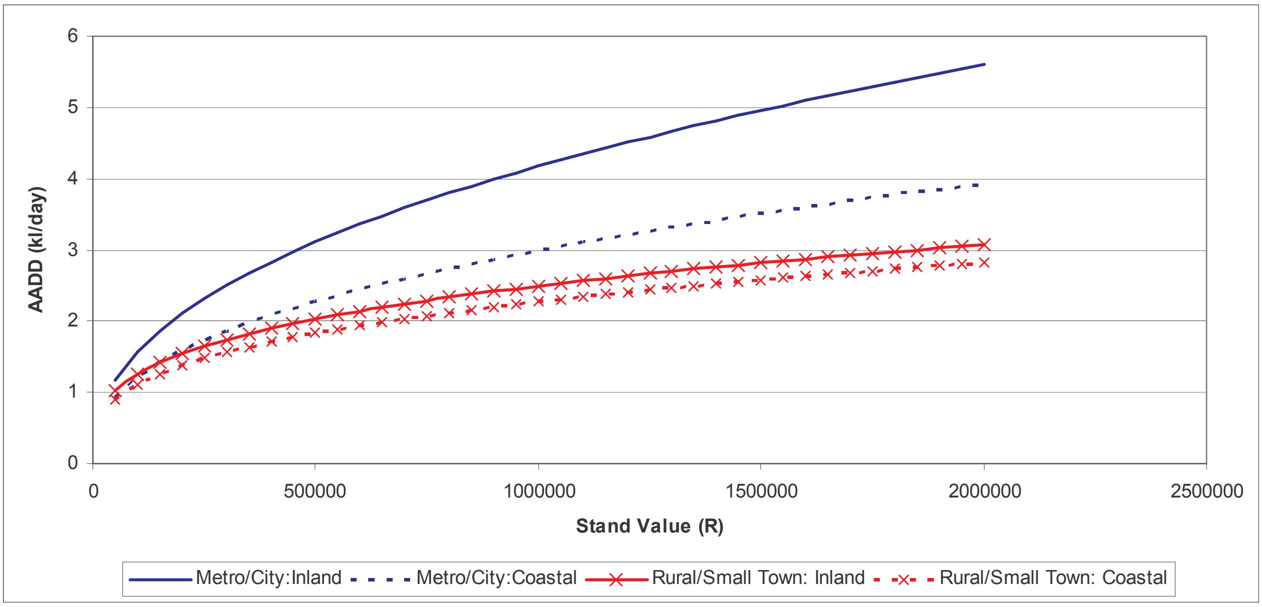


Figure 5-27: Single Variable Regression Results: AADD vs. Stand Value for Government and Institutional Users

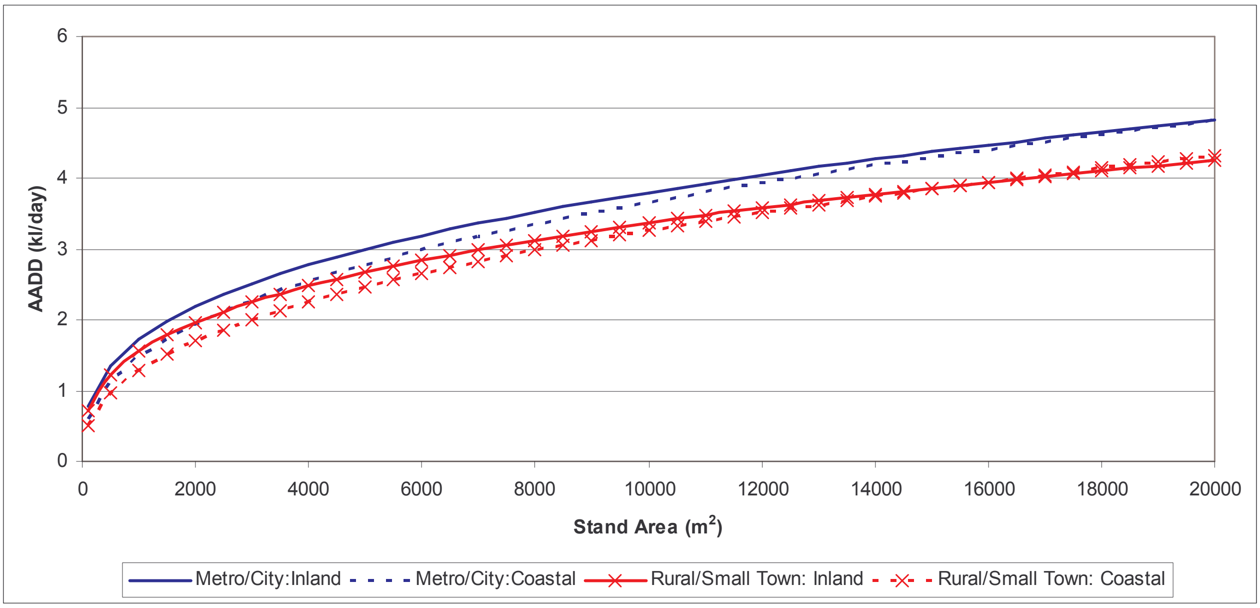


Figure 5-28: Single Variable Regression Results: AADD vs. Stand Area for Government and Institutional Users

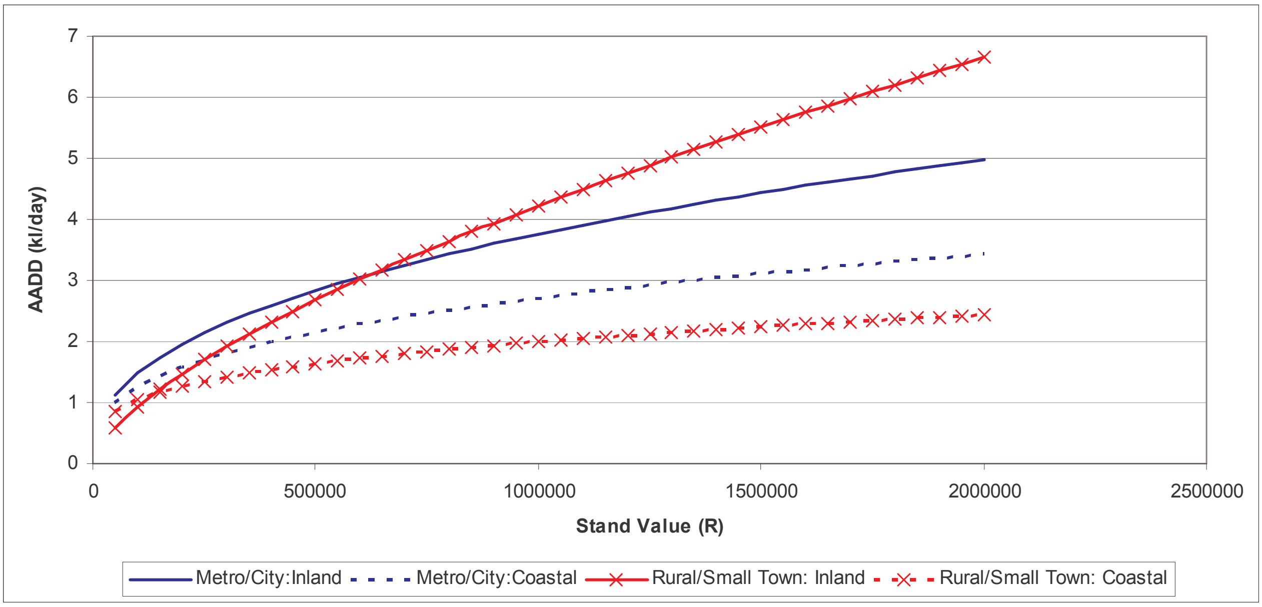


Figure 5-29: Single Variable Regression Results: AADD vs. Stand Value for Industrial Users

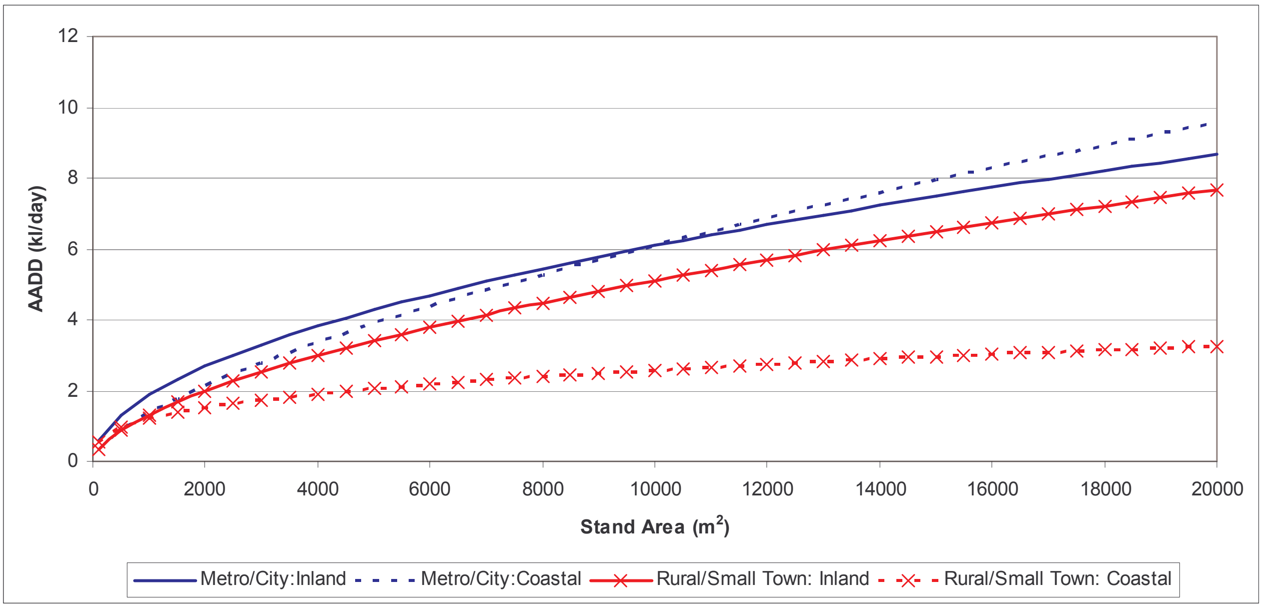


Figure 5-30: Single Variable Regression Results: AADD vs. Stand Area for Industrial Users

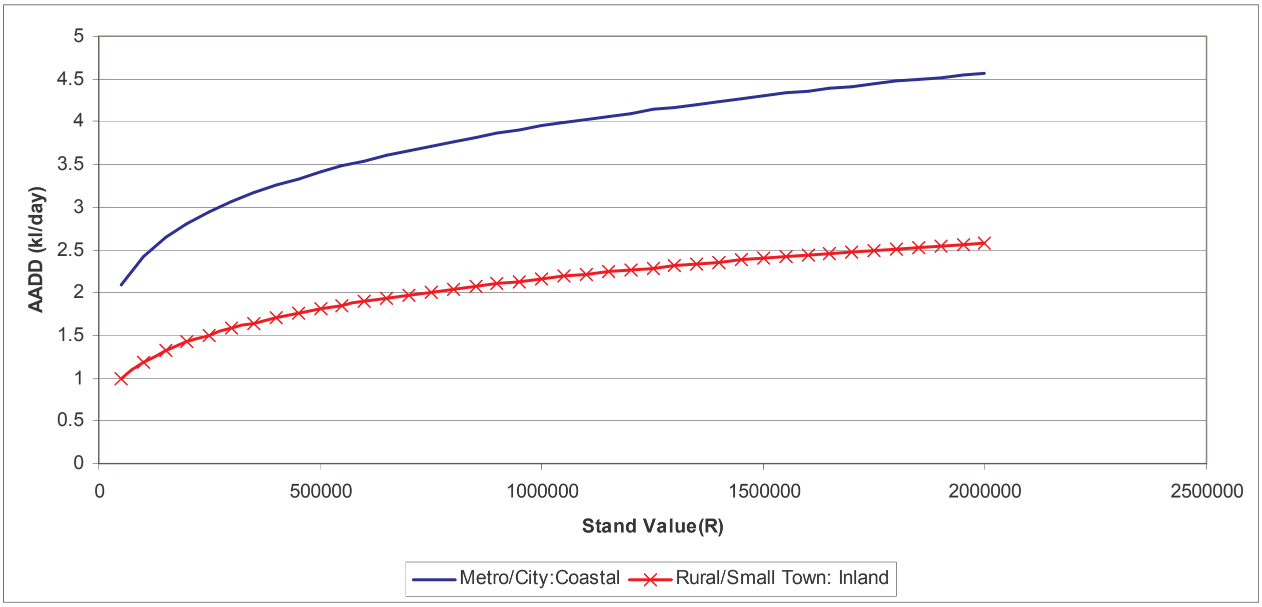


Figure 5-31: Single Variable Regression Results: AADD vs. Stand Value for Parks Users

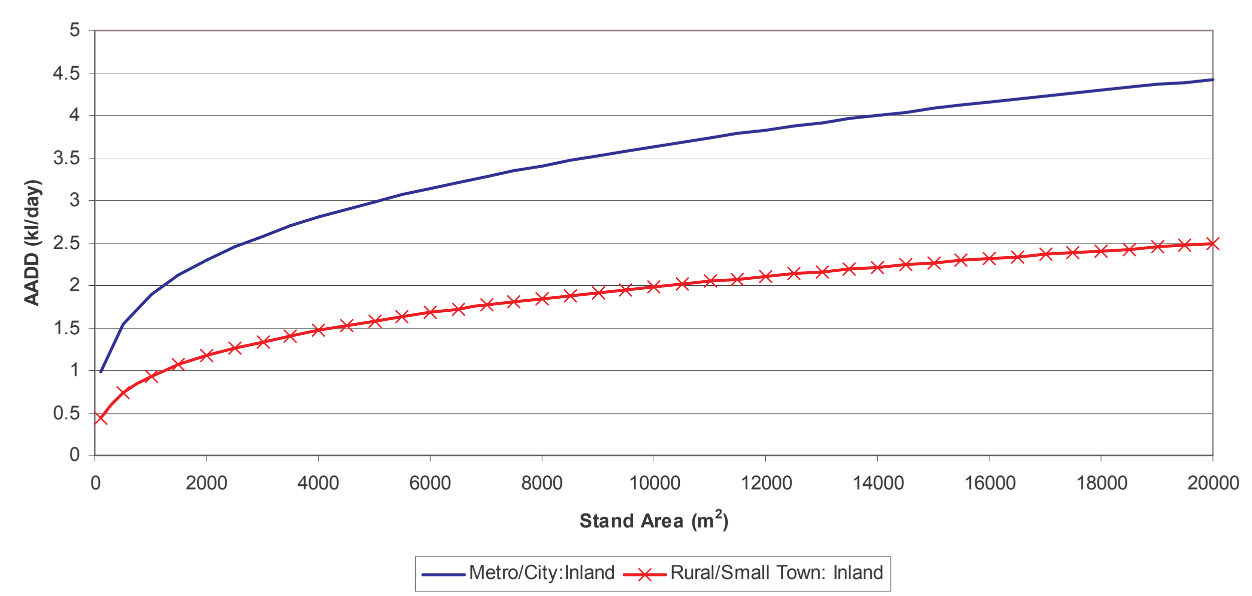


Figure 5-32: Single Variable Regression Results: AADD vs. Stand Area for Parks Users

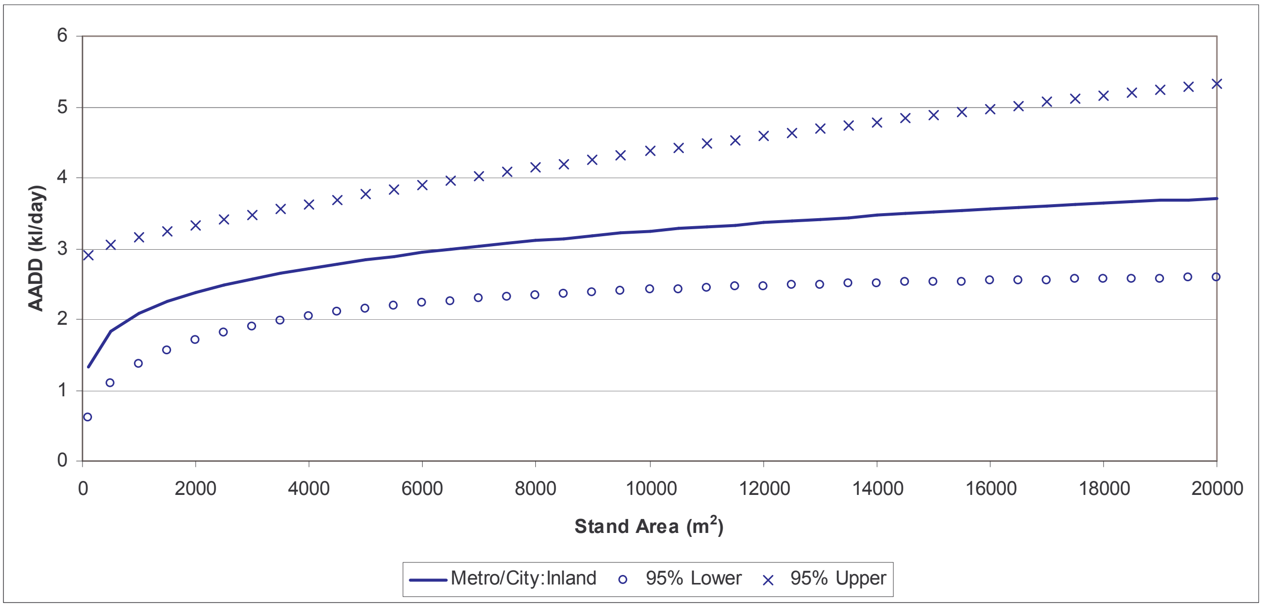


Figure 5-33: Single Variable Regression Results: AADD vs. Stand Area for Sports Users

From the models, it seemed that the development level might have an impact on the Business Commercial, Education, Government and Parks user categories in that more water is consumed in the metropolitan areas than in the rural areas for the same stand area and stand value. However, the results for farms and industrial were inconsistent with no clear indication whether development level had any impact. In fact, the results for farms indicated that farms in inland metropolitan areas did not follow the same trend as the other models at all, namely the model predicted the average AADD decreases with increasing stand area. The reason, may possibly be that for the farms category the metropolitan areas include more agricultural small holdings than what was the case for rural areas. Agricultural small holdings should rather be classified as domestic users. Their inclusion in the farms dataset would increase the unexplained variability in the results significantly.

Geographic location did not seem to have any consistent impact on any of the non-domestic user categories. Given this inconsistent outcome of the single variable analyses with the distinction made for development level and geographic location, it was considered prudent to also conduct single variable regression analyses on the non-domestic user categories without making any distinction on development level and geographic location at all. Furthermore, by making the distinction some of the sample sizes were too small to provide significant results (e.g. Parks and Sports).

Table 5.8 provides a summary of single variable regression analyses performed for each non-domestic user category for the independent variables of stand area and stand value.

Table 5-8: Summary of the Single Variable Regression Analyses for Non-Domestic Users Categories

Category	Independent Variable	Equation	N	Adjusted R^2
Business Commercial	Stand Area	$\ln(AADD) = 0.428\ln(StdArea) - 2.275 \pm 1.748\sqrt{3.8\times10^{-5} + \frac{(\ln StdArea - 7.247)^2}{32408}}$	26,391	0.22
	Stand Value	$\ln(AADD) = 0.385\ln(StdValue) - 4.214 \pm 1.74\sqrt{3.8\times10^{-5} + \frac{(\ln StdValue - 13.08)^2}{42146}}$	26,391	0.232
Education	Stand Area	$\ln(AADD) = 0.488\ln(StdArea) - 2.807 \pm 1.9\sqrt{6.8\times10^{-4} + \frac{(\ln StdArea - 8.659)^2}{3920}}$	1,471	0.404
	Stand Value	$\ln(AADD) = 0.468\ln(StdValue) - 4.883 \pm 2.03\sqrt{6.8\times10^{-4} + \frac{(\ln StdValue - 13.465)^2}{3307}}$	1,471	0.313
Farms	Stand Area	$\ln(AADD) = 0.07\ln(StdArea) + 0.098 \pm 1.503\sqrt{1.1\times10^{-4} + \frac{(\ln StdArea - 9.88)^2}{7926}}$ #	8,906	0.007
	Stand Value	$\ln(AADD) = 0.276\ln(StdValue) - 2.673 \pm 1.44\sqrt{1.1\times10^{-4} + \frac{(\ln StdValue - 12.52)^2}{6298}}$	8,934	0.091
Government and Institutional	Stand Area	$\ln(AADD) = 0.354\ln(StdArea) - 1.983 \pm 1.21\sqrt{1.1\times10^{-4} + \frac{(\ln StdArea - 6.57)^2}{13965}}$	9,062	0.335
	Stand Value	$\ln(AADD) = 0.358\ln(StdValue) - 3.847 \pm 1.23\sqrt{1.1\times10^{-4} + \frac{(\ln StdValue - 11.712)^2}{12805}}$	9,062	0.314
Industrial	Stand Area	$\ln(AADD) = 0.531\ln(StdArea) - 3.153 \pm 1.83\sqrt{1.2\times10^{-4} + \frac{(\ln StdArea - 7.819)^2}{11027}}$	8,138	0.306

Category	Independent Variable	Equation	N	Adjusted R^2
	Stand Value	$\ln(AADD) = 0.395\ln(StdValue) - 4.252 \pm 1.95\sqrt{1.2\times10^{-4} + \frac{(\ln StdValue - 13.303)^2}{13566}}$	8,138	0.208
Parks	Stand Area	$\ln(AADD) = 0.316\ln(StdArea) - 1.667 \pm 2.18\sqrt{2.6\times10^{-3} + \frac{(\ln StdArea - 8.691)^2}{780}}$	378	0.141
	Stand Value	$\ln(AADD) = 0.212\ln(StdValue) - 1.6 \pm 2.23\sqrt{2.2\times10^{-3} + \frac{(\ln StdValue - 12.217)^2}{973}}$	451	0.067
Sports	Stand Area	$\ln(AADD) = 0.103\ln(StdArea) + 0.03 \pm 2.1\sqrt{1.4\times10^{-2} + \frac{(\ln StdArea - 8.8)^2}{186}}$ #	72	0.01
	Stand Value	$\ln(AADD) = 0.122\ln(StdValue) - 0.657 \pm 2.1\sqrt{1.4\times10^{-2} + \frac{(\ln StdValue - 13.102)^2}{137}}$ #	72	0.01

 # Equation coefficients do not differ significantly from zero (excluded from further analysis)

Figures 5.34 and 5.35 plot the regression models for all the non-domestic user categories for the independent variables stand area and stand value respectively. Appendix C contains separate graphs for each user category with the 97.5 % confidence intervals for the average AADD.

5.2.3 Proposal for Revising the Existing Non- Domestic Demand Estimation Guideline

As discussed in Chapter 4 it is generally difficult to obtain information on the non-domestic use anticipated in a new development. Data such as the floor area or number of occupants is not easily collected. Sometimes, at most, it is possible to gather information on the type of non-domestic use. An ideal non-domestic guideline would therefore be one, where minimum 'input' data is required.

This study investigated the possibility of generating a guideline that only requires the anticipated number of users. In other words the independent variable for the guideline curve would be the number of users. The characteristics of the log normal distributions, that were fitted to the non-domestic data, were used to calculate the 95% tolerance limits for the means of the distributions. The tolerance limits give the 95% confidence that the average water demand for X number of users will fall within the two limits. For the purpose of demand estimation it is proposed to only use the upper 95% tolerance limit, because one would not want to use a design AADD value that is less than the mean AADD for a user category. The guideline curve would thus have an asymptote that is the mean of the log normal distribution. The format of the proposed equation is given below:

$$x = \bar{x} + Z\alpha S\sqrt{\frac{1}{N}}$$

Where:

y = ln(AADD)

x = mean of log –normal distribution of AADD

Zα = critical value of the normal distribution

S= standard error of the estimate of the log normal distribution fit

N= number of observations

This approach assumes that the non-domestic water demand follows a log-normal trend. This assumption is based on the good fits obtained with the frequency distributions discussed in Section 5.2.1.

Figure 5.34 to 5.39 portray the upper 95% tolerance limits for the means of the log-normal distributions for each of the non-domestic user categories. It is proposed that these curves are considered as a possible indication of what demand guidelines for non-domestic use could look like.

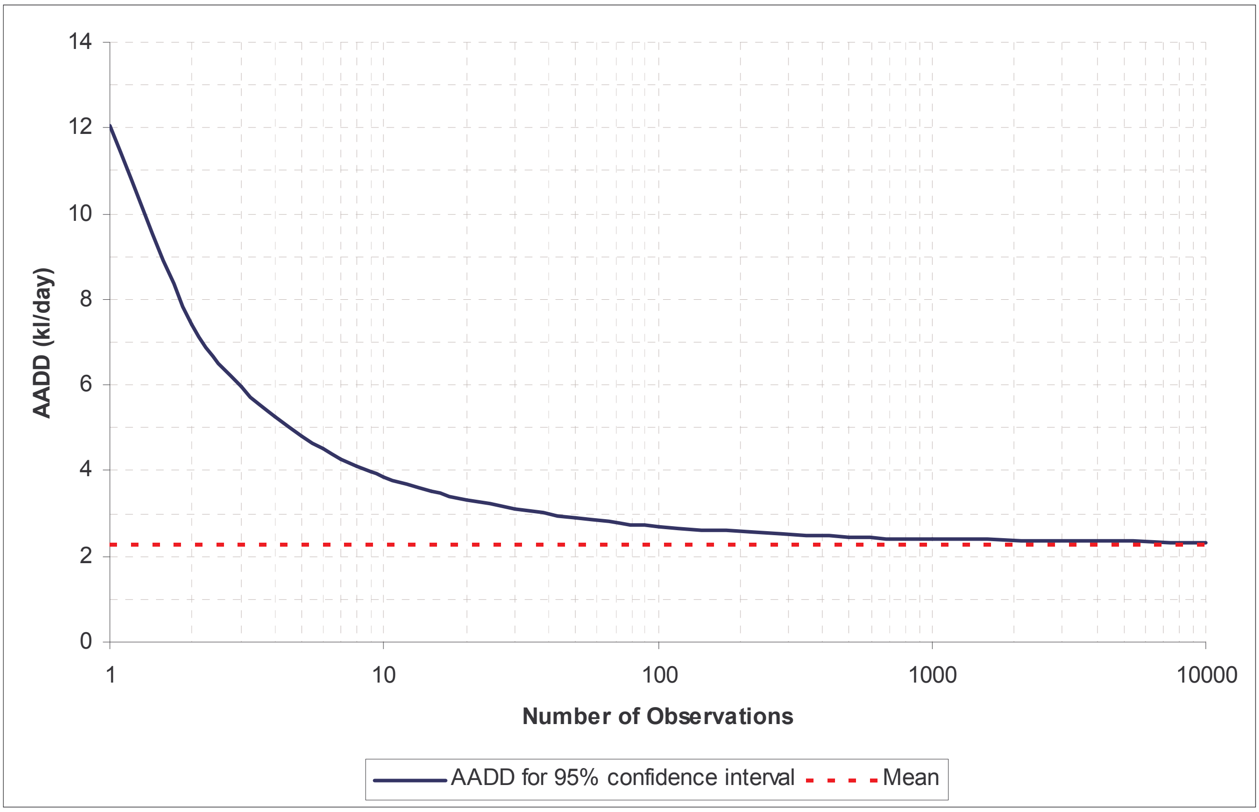


Figure 5-34: Possible Design Guideline for Business Commercial Water Demand

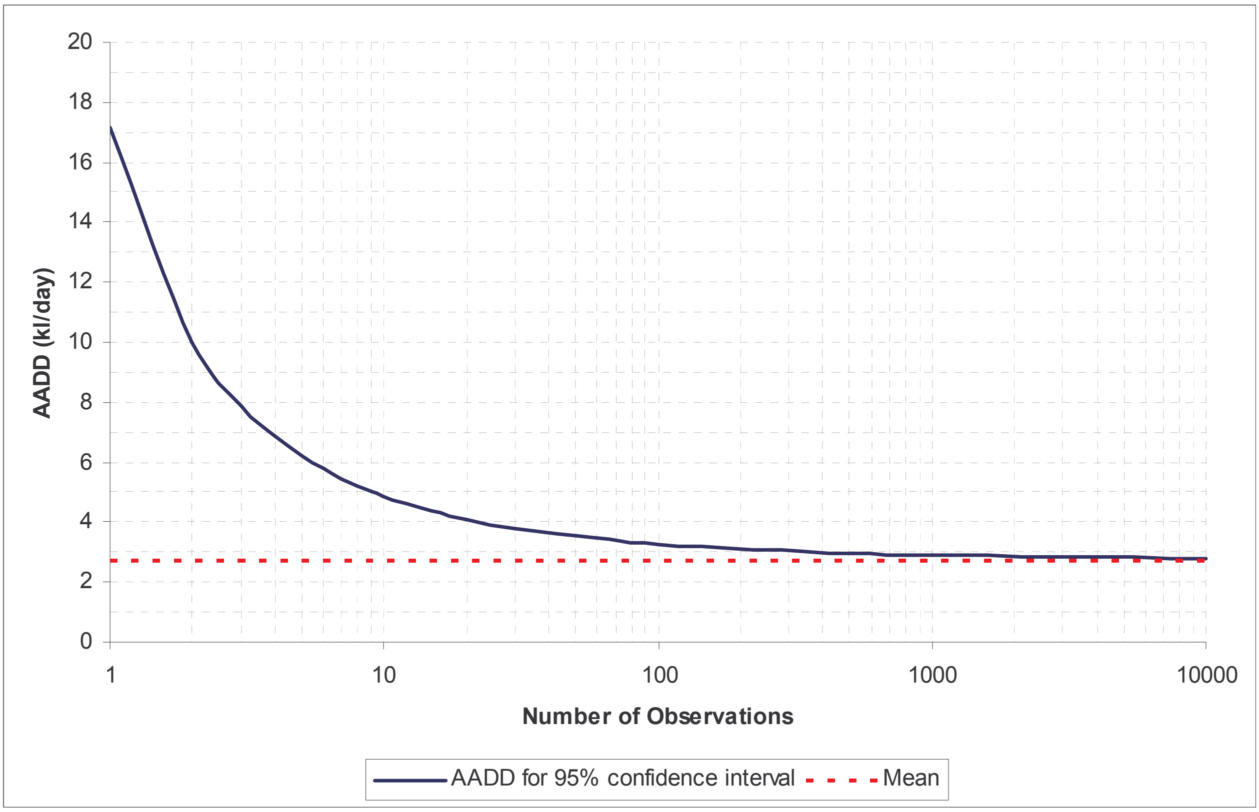


Figure 5-35: Possible Design Guideline for Industrial Water Demand

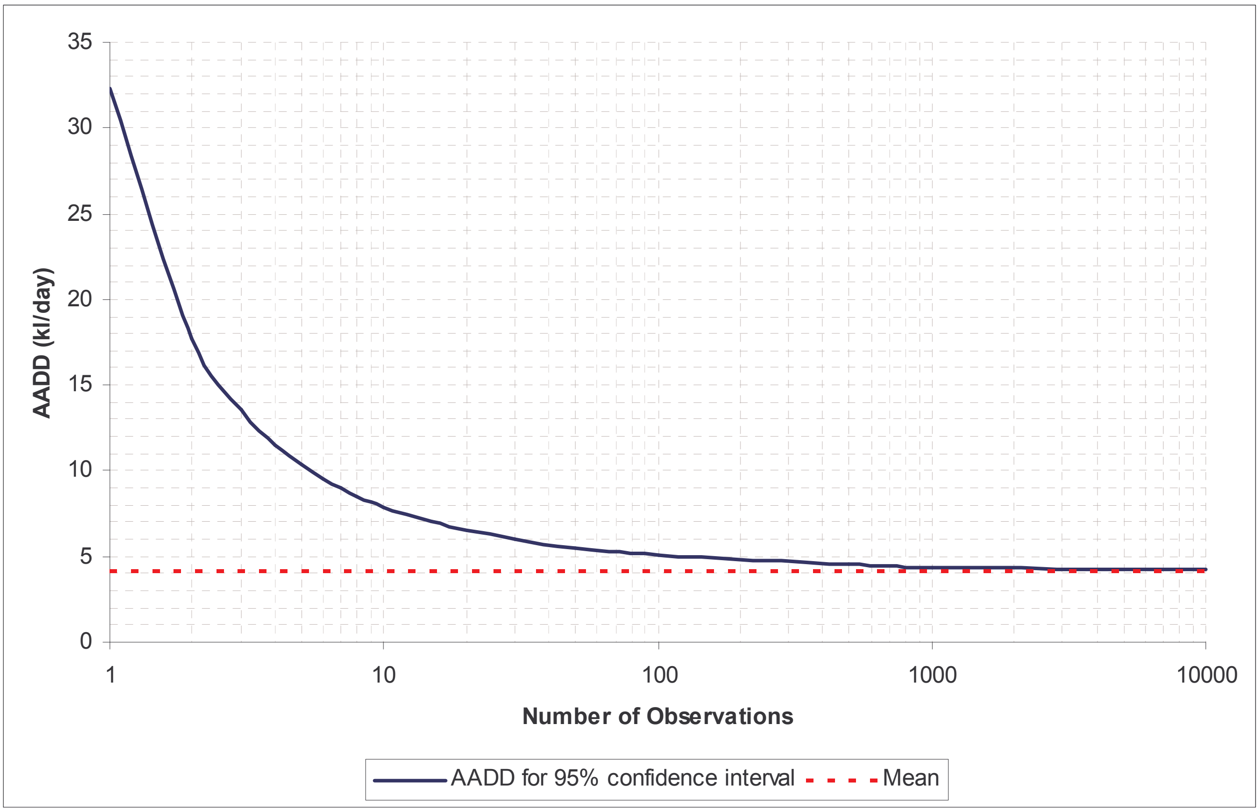


Figure 5-36: Possible Design Guideline for Education Water Demand

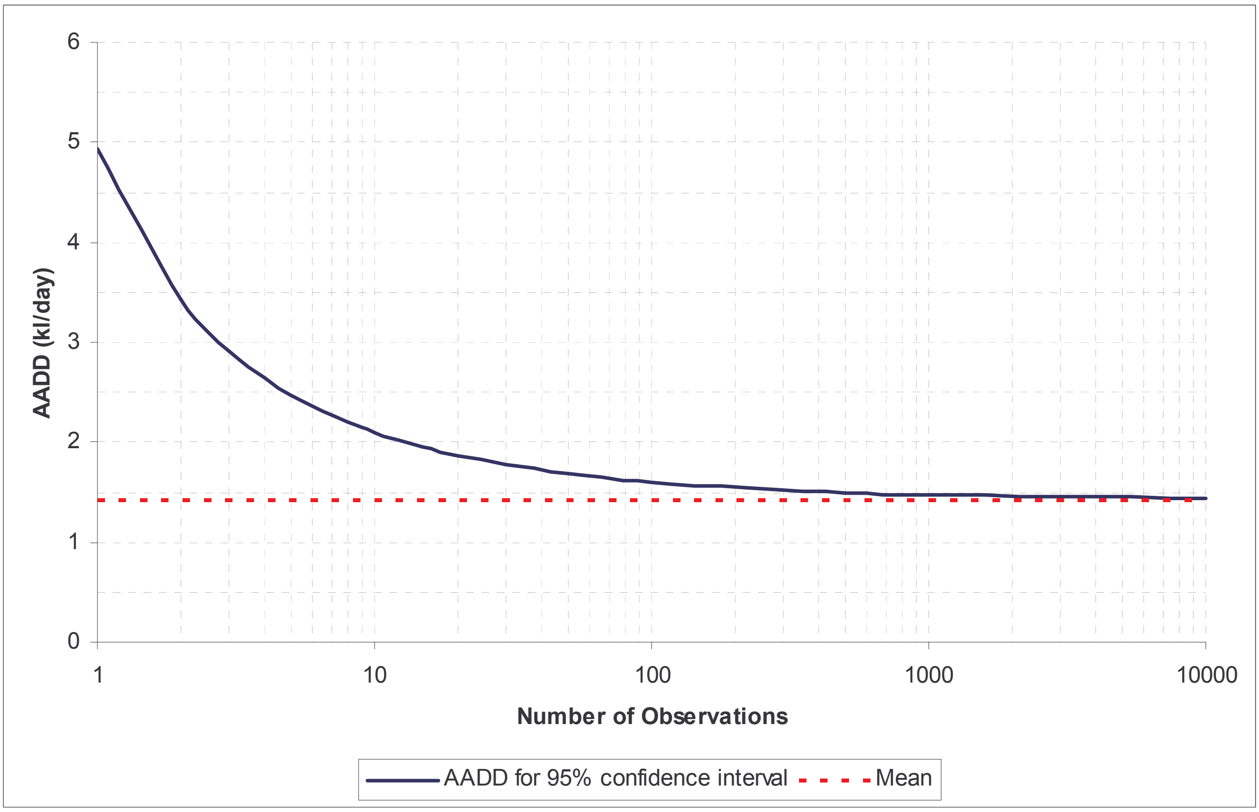


Figure 5-37: Possible Design Guideline for Government and Institutional Water Demand

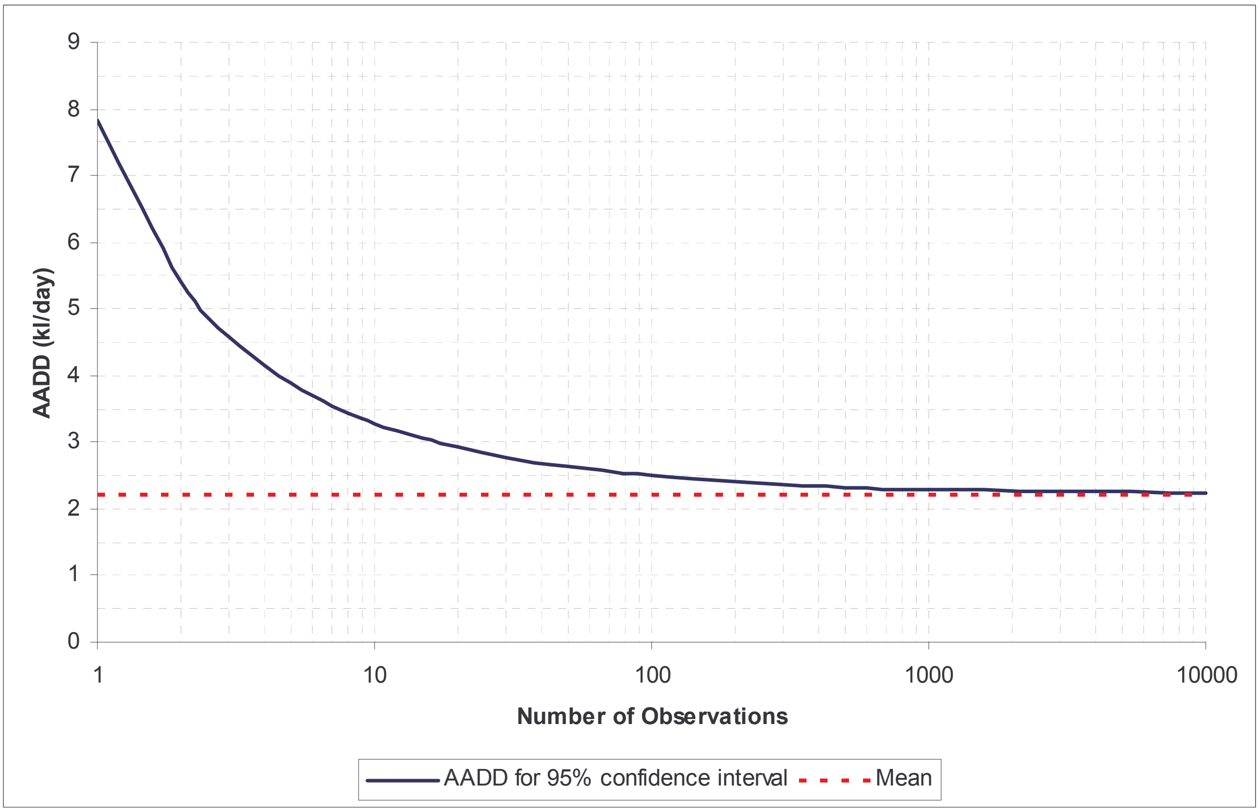


Figure 5-38: Possible Design Guideline for Farm Water Demand

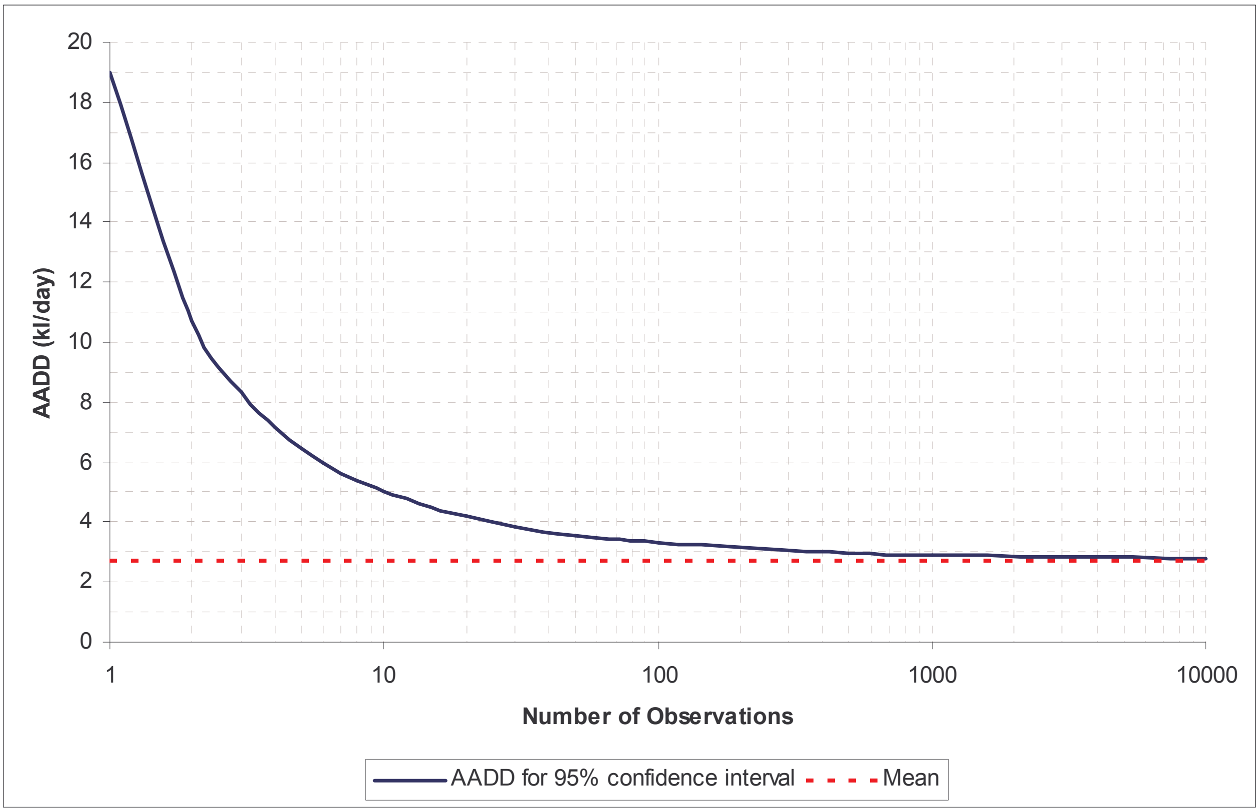


Figure 5-39: Possible Design Guideline for Park Water Demand

Figure 5.40 compares the guideline curves for all six non-domestic user categories considered above.

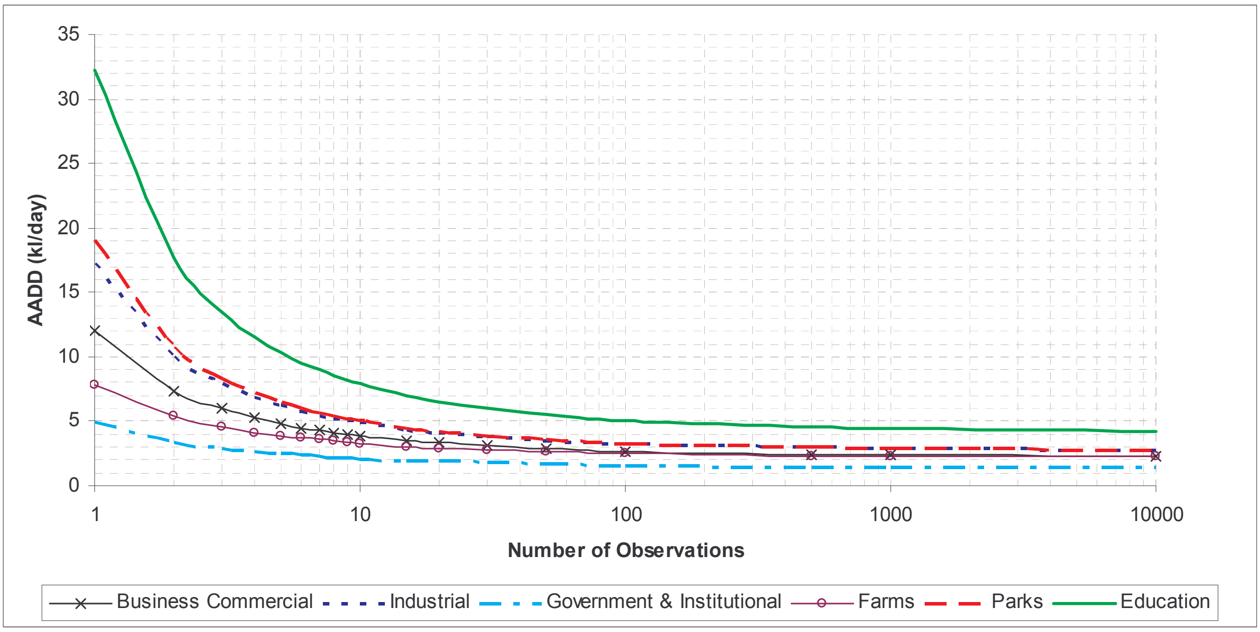


Figure 5-40: Comparison of Demand Estimation Curves for All Non-Domestic Categories Considered

6 CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

6.1 Conclusions

The main part of this study consisted of an analysis of more than a million individual water consumption records obtained from municipal treasury systems. Most of the records represented more than two years of consumption data. The main objective of this study was to evaluate the parameters that influence domestic and non-domestic water demands and to evaluate the existing South African water demand estimation guidelines. Climatic and socio-economic data was obtained and linked to the consumption data. Unfortunately, the socio-economic data was only available for political wards, which could include different suburbs with significantly different socio-economic characteristics.

Stepwise multiple-variable regressions were applied to the domestic and non-domestic consumption data together with the data on parameters possibly influencing water demand. This was done to determine the parameters that influence water demand most. In a large majority of cases, either stand area or stand value had the greatest significance.

The main conclusions of the study are as follows:

- 47 % of the average suburb domestic demands fell inside the design envelope proposed by the South African design guidelines.
- For domestic use, the regression models obtained with this study, predict average AADD values within the range of the South African design guideline envelope curves up to stand areas (areas) of approximately 1 700m^2. For larger stands, less water consumption is predicted.
- There is strong evidence that domestic water consumption increases with both increasing stand area and stand value (used as a proxy for household income).
- Inland domestic water demand is significantly higher that coastal demand.
- Non-domestic demand categories display good lognormal distribution trends.
- The most significant parameters influencing non-domestic water demand were stand area and stand value. The regression models obtained indicated that non-domestic demand will increase with increasing stand area and stand value.

- There seems to be no significant difference between inland non-domestic demand and coastal non-domestic demand.
- Development level namely city vs. small town has a significant impact on non-domestic demand.

This study investigated possible revisions to demand estimation guidelines for both domestic and non-domestic water use. For domestic water demand estimation, a single curve for AADD vs. stand area was developed as a possible revised guideline curve, using all the available domestic data (1 091 685 domestic stands). The new demand estimation curve has design envelopes around it that are based on the 25%, 75% and 95% confidence intervals for the number of data points. Figure 6.1 below presents the proposed revised guideline curve and compares it to the "Red Book" (CSIR, 2003) guideline curve.

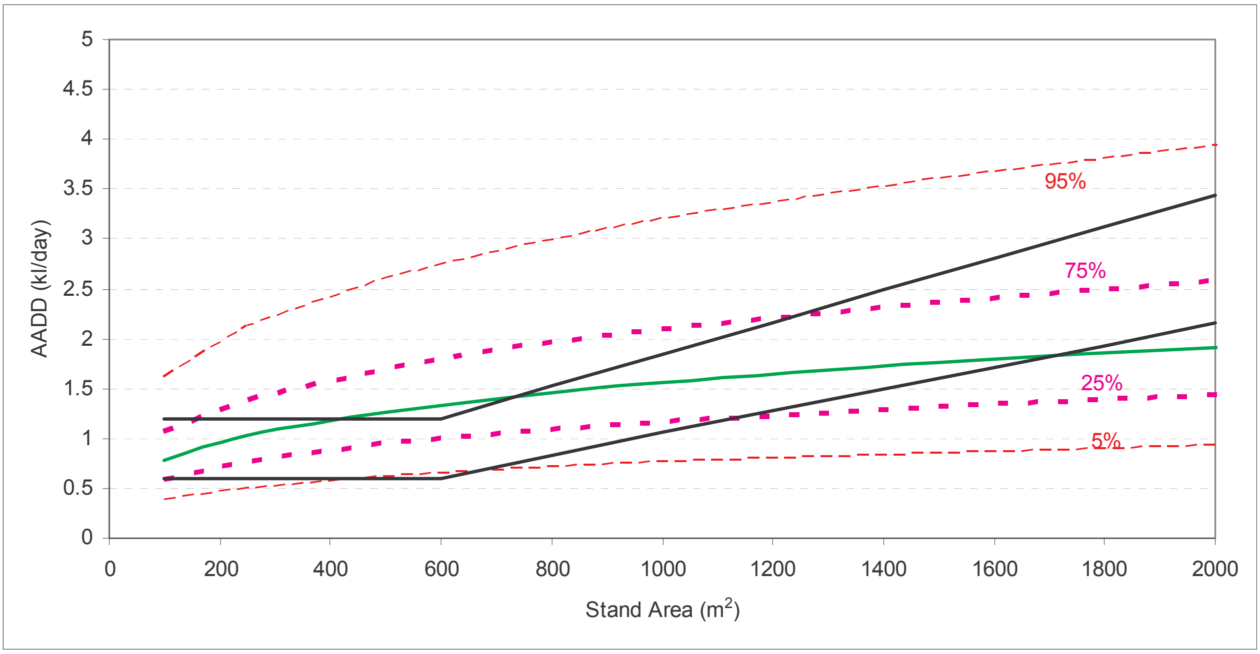


Figure 6-1: Proposed Revision to the Domestic Demand Estimation Guideline

For non-domestic demand, new preliminary estimation curves were developed for 6 of the 7 non-domestic user types studied. The curves are based on the log-normal distribution trends that were observed in the non-domestic data. In each case, the AADD is expressed as a function of the number of stands. The mean of the normal distribution of each category is an asymptote in each equation. The six equations are listed below:

Business Commercial: $$\ln AADD = 0.83 + (1.65)(1.01)\sqrt{\frac{1}{N}}$$

Industrial: $$\ln AADD = 1.003 + (1.65)(1.12)\sqrt{\frac{1}{N}}$$

Education: $$\ln AADD = 1.42 + (1.65)(1.25)\sqrt{\frac{1}{N}}$$

Government and Institutional: $$\ln AADD = 0.35 + (1.65)(0.76)\sqrt{\frac{1}{N}}$$

Farms: $$\ln AADD = 0.79 + (1.65)(0.77)\sqrt{\frac{1}{N}}$$

Parks: $$\ln AADD = 0.995 + (1.65)(1.18)\sqrt{\frac{1}{N}}$$

Figure 6.2 portrays the proposed estimation curves for all six non-domestic user categories considered above. It is not proposed that these curves are used for design purposes, but rather as an indication of what guideline curves could look like for various non-domestic user categories.

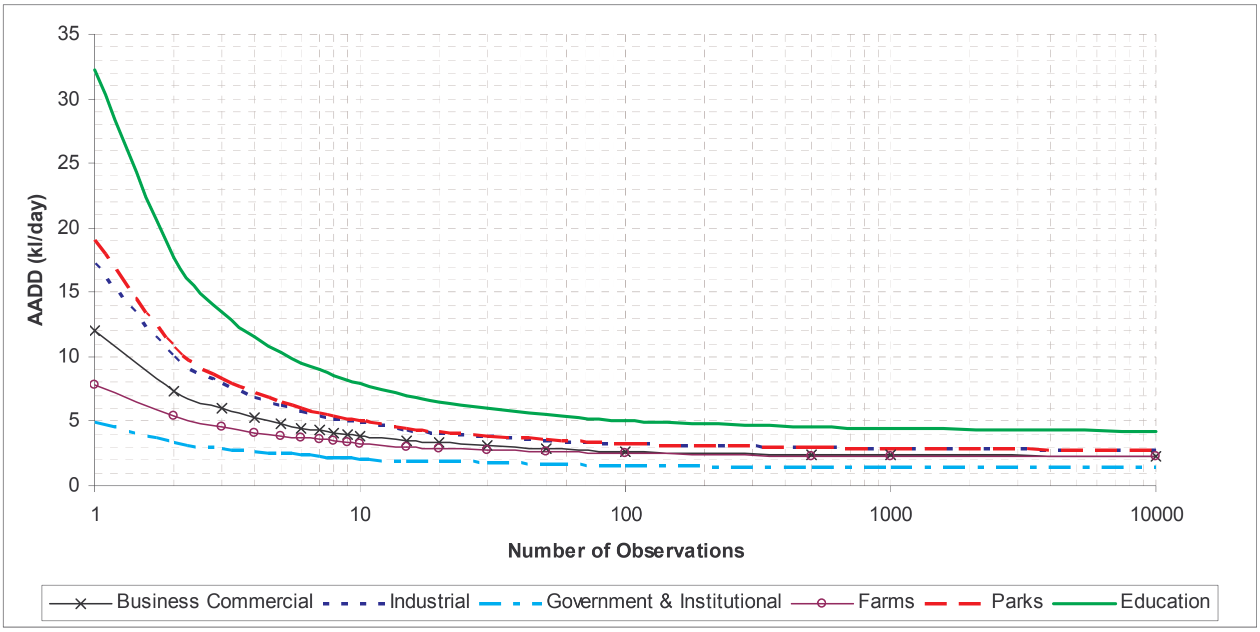


Figure 6-2: Possible Design Guideline Curves for Non-Domestic Categories Considered

6.2 Recommendations for Future Work

Future work is required on the following aspects:

i) Further research is required to understand the reasons for the difference between the proposed revision of the domestic demand estimation curve presented in this study and the existing South African demand estimation guidelines (CSIR, 2003).

ii) The socio-economic data that was studied with this project, was only available for political wards, which could include different suburbs with vastly different socio-economic characteristics. It would be very valuable if the socio-economic data could be obtained at suburb level, and the impact of the socio-economic factors in water demand re-assessed.

iii) This study relied heavily on the assigned land usage originating from the municipal treasury systems. Further verification and possible refinement of the land use codes in the data may improve the variability observed in the data with the regression analyses.

iv) More detailed study is required to establish the reason for the consistent difference between coastal and inland domestic water consumption observed.

v) From the literature survey undertaken as part of this study it is evident that water tariffs and water restrictions have a definite impact on both domestic and non-domestic water consumption. However, the available tariff information was limited with regards to accuracy and availability from the various municipalities studied. Consequently, this study did not investigate the impact of water tariffs and restrictions on water demand. Further work is necessary to determine the impact of water tariffs and restrictions on water demand in the short term and the long term.

vi) The literature review also indicated that available pressure within a distribution system has a significant impact on water demand. This study did not consider available pressure as a possible parameter, as pressure data was not readily available from the various municipalities. However, a further study is needed to attempt to quantify the impact of available pressure on water demand.

vii) Further investigation is necessary to verify and refine the proposed preliminary estimation curves for non-domestic user categories.

7 REFERENCES

BASSON, M S, VAN NIEKERK, P H, VAN ROOYEN, J A 1997. Overview of Water Resources Availability and Utilisation in South Africa. Report by Department of Water Affairs and Forestry, P RSA/00/0197, May.

CITY OF JOHANNESBURG 1989. Water Supply Guidelines for Township Development in Johannesburg. Water and Gas Department, April.

CSIR 1983. Guidelines for the provision of engineering services in residential townships (Blue Book). Compiled for the Department of Community Development by the CSIR (SA).

CSIR 1994. Guidelines for the provision of Engineering Services and Amenities in Residential Township Development (Red Book), Compiled for National Housing Advice Council.

CSIR 2003. Guidelines for human settlement planning and design. A report compiled under the patronage of the Department of Housing

DANDY D, NQUYEN T, DAVIES C 1997. Estimating Residential Water Demand in the Presence of Free Allowance. *Land Economics 73 (1)*

GARLIPP, K D C O .1979. Water Consumption Patterns in urban areas. M.Sc. Dissertation, University of Pretoria.

HOWE, C W, LINAWEAVER, F P 1967. The impact of Price on Residential Water Demand and its Relation to System Design and Price Structure. Water Resources Research, 3 (1): 13-32

HUSSELMAN, M L 2004. Estimation of Area and Income Elasticities of Water Demand in a Number of Cities and Towns in Gauteng. Magister Ingeneriae Dissertation. Rand Afrikaans University (now University of Johannesburg). August.

JACOBS, H E 2004. A Conceptual End-Use Model for Residential Water Demand and Return Flow, D.Ing. Thesis, Rand Afrikaans University

JACOBS,H E, GEUSTEYN, L C, LOUBSER, B F, VAN DER MERWE, B 2004. Estimating residential water demand in Southern Africa. Journal of the South African Institution of Civil Engineering, 46(4): 2-13, Paper 568

MIDGLEY DC, PITMAN WV, MIDDLETON BJ. 1994. Surface Water Resources of South Africa, WRC Report No 298/1/94,User Manual.. ISBN 1 86845 143 7 and all the subsequent volumes.

RAND WATER, 1999. Consumer Market Segmentation Study, Report EG03, Rand Water Marketing Department, Johannesburg

STEPHENSON, D AND TURNER, K 1996. Water demand patterns in Gauteng. IMIESA, 21 (1): 11-16

VAN VUUREN, S J, VAN BEEK, J C, 1997, Her-Evaluering van die Bestaande Riglyne vir Stedelike en Industriele Watervoorsiening gebaseer op Gemete Waterverbruike Fase 1: Pretoria Voorsieningsgebied, Report for the Water Research Commission by the Department of Civil Engineering, University of Pretoria, WRC Report No 705/1/97.

VAN ZYL, J E, HAARHOFF, J, HUSSELMAN, M L 2003. Potential application of end-use demand modelling in South Africa. Journal of the South African Institution of Civil Engineering, 45(2): 9-19.

VECK, G A, BILL, M R, 2000. Estimation of the Residential Price Elasticity of Demand for Water by means of a contingent valuation approach. Report for the Water Research Commission of South Africa. WRC Report No. 790/1/00.

WATER TRANSFER CONSULTANTS 2001. Central Area System Update: Determination of the Current and Forecasted Water Demands for the Central Area of Namibia. Report compiled December 2001.

WCE, ENVES, EPE and CSIR 2003. Economic feasibility study on the artificial recharge of the Windhoek Aquifer. Report compiled June 2003.