DERIVING PEAK FACTORS FOR RESIDENTIAL INDOOR WATER DEMAND BY MEANS OF A PROBABILITY BASED END-USE MODEL

by
Hester Maria Scheepers

Thesis presented in partial fulfilment of the requirements for the degree
Master of Science in Engineering at Stellenbosch University

Supervisor: Prof. H.E. Jacobs
Faculty of Engineering
Department of Civil Engineering
Division of Water and Environmental Engineering

December 2012

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: ....
H M Scheepers

Date: ....

ABSTRACT

The expected peak water demand in a water distribution system (WDS) is an important consideration for WDS design purposes. In South Africa the most common method of estimating peak demand is by multiplying the average demand by a dimensionless peak factor. A peak factor is the ratio between the maximum flow rate (which refers to the largest volume of flow to be received during a relatively short time period, say δt, expressed as the average volume per unit time), and the average flow rate over an extended time period.

The magnitude of the peak factor will vary, for a given daily water demand pattern, depending on the chosen value of δt. The design guidelines available give no clear indication of the time intervals most appropriate for different peak factor applications. It is therefore important to gain a better understanding regarding the effect of δt on the derived peak factor.

A probability based end-use model was constructed as part of this study to derive diurnal residential indoor water demand patterns on a temporal scale of one second. These stochastically derived water demand patterns were subsequently used to calculate peak factors for different values of δt, varying from one second to one hour.

The end-use model derived the water demand patterns by aggregating the synthesised end-use events of six residential indoor end-uses of water in terms of the water volume required, duration and the time of occurrence of each event. The probability distributions describing the end-use model parameters were derived from actual end-use measurements that had previously been collected in a noteworthy North-American end-use project (Mayer et al., 1999). The original comprehensive database, which included water measurements from both indoor and outdoor end-uses, was purchased for use in this project.

A single execution of the end-use model resulted in the synthesised diurnal water demand pattern for a single household. The estimated water demand pattern for simultaneous water demand by groups of households was obtained

by adding individual iterations of the end-use model, considering group sizes of between one and 2 000 households in the process. A total of 99 500 model executions were performed, which were statistically aggregated by applying the Monte Carlo method and forming 4 950 unique water demand scenarios representing 29 different household group sizes. For each of the 4 950 water demand scenarios, a set of peak factors was derived for eight selected δt values.

The end-use model presented in this study yielded realistic indoor water demand estimations when compared to publications from literature. In agreement with existing knowledge, as expected, an inverse relationship was evident between the magnitude of the peak factors and δt. The peak factors across all time intervals were also found to be inversely related to the number of households, which agreed with other publications from literature. As the number of households increased, the degree to which the peak factor was affected by the time intervals decreased.

This study explicitly demonstrated the effect of time intervals on peak factors. The results of this study could act as the basis for the derivation of a practical design guideline for estimating peak indoor flows in a WDS, and the work could be extended in future to include outdoor water demand and sensitivity to WDS pressure.

OPSOMMING

Die verwagte water spitsaanvraag is ‘n belangrike oorweging in die ontwerp van ‘n waterverspreidingsnetwerk. Die mees algemene metode in Suid Afrika om spitsaanvraag te bereken is deur die gemiddelde wateraanvraag te vermeningvuldig met ‘n dimensielose spitsfaktor. ‘n Spitsfaktor is die verhouding tussen die maksimum watervloei tempo (wat verwys na die grootste volume water wat ontvang sal word tydens ‘n relatiewe kort tydsinterval, δt, uitgedruk as die gemiddelde volume per tyd eenheid), en die gemiddelde watervloei tempo gedurende ‘n verlengde tydsinterval. Die grootte van die spitsfaktor sal varieer vir ‘n gegewe daaglikse vloeipatroon, afhangende van die verkose δt waarde. Die beskikbare ontwerpsriglyne is onduidelik oor watter tydsintervalle meer geskik is vir die verskillende spitsfaktor toepassings. Daarom is dit belangrik om ‘n beter begrip te verkry ten opsigte van die effek van δt op die verkrygde spitsfaktor.

‘n Waarskynliksheidsgebaseerde eindverbruik model is opgestel om deel te vorm van hierdie studie, om daaglikse residensiële binnenshuise wateraanvraag patrone af te lei op ‘n temporale skaal van een sekonde. Die stogasties afgeleide wateraanvraag patrone is daarna gebruik om die verskeie spitsfaktore te bereken vir verskillende waardes van δt, wat varieer vanaf een sekonde tot een uur.

Die eindverbruik model stel die daaglikse vloeipatroon van een huis saam deur die eindeverbruik gebeure van ses residensiële binnenshuise eindverbruike saam te voeg in terme van the vereiste water volume en die tyd van voorkoms van elke gebeurtenis. Die waarskynliksheids distribusie wat die eindverbruik model parameters omskryf is verkry van werklike gemete eindverbruik waardes, wat voorheen in ‘n beduidende Noord-Amerikaanse eindverbruik projek (Mayer et al. 1999) versamel is. Die oorspronklike en omvattende databasis, wat gemete waardes van binnenshuis en buite water verbruik ingesluit het, is aangekoop vir gebruik gedurende hierdie projek.

‘n Enkele uitvoering van die eindverbruik model stel gevolglik ‘n daaglikse wateraanvraag patroon saam vir ‘n elkele huishouding. Die wateraanvraag patroon vir gelyktydige water verbruik deur groepe huishoudings is verkry deur individuele iterasies van die eindverbruik model statisties bymekaar te tel met die Monte Carlo metode, terwyl groep groottes van tussen een en 2 000 huishoudings in die proses oorweeg is. ‘n Totaal van 99 500 model uitvoerings is gedoen, wat saamgevoeg is om 4 950 unieke watervraag scenarios voor te stel, wat verteenwoordigend is van 29 verskillende huishouding groep groottes. Vir elkeen van die 4 950 watervraag senarios, is ‘n stel spitsfaktore afgelei vir agt verkose δt waardes.

Die eindverbruik model aangebied in hierdie studie lewer ‘n realistiese binnenshuise wateraanvraag skatting, wanneer dit vergelyk word met verslae in die literatuur. Ooreenkomstig met bestaande kennis is ‘n sterk inverse verhouding sigbaar tussen die grootte van die spitsfaktore en δt. Dit is ook gevind dat die spitsfaktore oor al die tydsintervalle ‘n inverse verband toon tot die aantal huishoudings, wat ooreenstemmend is met ander publikasies in die literatuur. Soos die aantal huishoudings toeneem, het die mate waartoe die spitsfaktor geaffekteer is deur die tydsintervalle afgeneem.

Hierdie studie toon duidelik die effek van tydsintervalle op spitsfaktore. Die resultaat van hierdie studie kan dien as basis om praktiese ontwerpsriglyne te verkry in die skatting van binnenshuise spitsvloei in ‘n waterverspreidingsnetwerk, gegewe dat die werk in die toekoms uitgebrei kan word om ook buitenshuise waterverbruik in te sluit, asook sensitiwiteit tot druk in die waterverspreidingsnetwerk.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Prof. Heinz Jacobs for his excellent ideas and guidance throughout the study.

This project would not have been possible without the Residential end-uses of water study database supplied by Peter Mayer (Aquacraft).

My sincere gratitude is due to my family, especially my mom Christelle Scheepers for her constant support.

I would also like to extend a word of appreciation to Prof. Musandji Fuamba (École Polytechnique) for the opportunity and training provided in Canada.

Financial support for this work was granted by the University of Stellenbosch, the Water Research Commission (WRC), and the South African National Research Foundation (NFR).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

Symbols used in subscripts retained the same definitions as the symbols given below. In this study, reference was made to separate publications that described various parameters using similar symbols. To ensure unique parameters were presented in this study, the symbols used by some authors were altered. In such instances the citation was added in brackets below.

A	Area
$\dot{A}^2$	A-D statistic
avg	Average
$B(I, D^*, \tau)$	Block function (Blokker et al., 2010)
c	Capita
C	Number of connections (Zhang, 2005)
d	Day
D	K-S statistic
D^*	Pulse duration (Blokker et al., 2010)
E	Population (Diao et al., 2010; Zhang, 2005)
F^*	Frequency of use (Blokker et al., 2010)
$F(x)$	Cumulative distribution function
$f(x)$	Probability distribution function
g	Gravitational acceleration
gal	Gallons
H_j	Head at node j when the demand of that node is Q_j
H_j^{min}	Nodal head below which the outflow at the node is unsatisfactory or zero
h	Hour

I	Pulse intensity
i	Bin
i^*	All busy times per end-use from 1 to $F_{j^*k^*}$ (Blokker et al., 2010)
$inst$	Instantaneous
j	Integer describing the network node number
j^*	All users from 1 to N^*(Blokker et al., 2010)
k	Number of bins
k^*	All end-uses from 1 to M^* (Blokker et al., 2010)
K_j	Flow resistance coefficient
ℓ	Litre
min	Minute
N	Number of data points / number of X values
$\ddot{N}$	Number of homes in the neighbourhood (Zhang et al., 2005)
$\widetilde{N}$	Number of consumers (Martinez-Solano et al., 2008; Tricarico et al., 2007)
$\grave{\mathrm{N}}_{\overline{\mathrm{C}}}$	Number of residential connections served by a pipe (Lingireddy et al., 1998)
n_j	Exponent applicable to head dependant analysis
P	Probability
$\mathcal{P}$	Pressure
$p(x)$	Probability mass function
p'	Maximum percentage (Brière, 2007)
Q	Flow rate
Q_j	Demand at node j
S	Sample space
s	Second

T	Time parameter
t	Time
δt	Short time interval
V	Volume
v	Velocity
w	Integer used in end-use model macro code
X	Random variable
$\bar{X}$	Median
$\mathcal{X}^2$	Chi-squared statistic
x	Possible values of X
z	Height above datum
α	Shape parameter
β	Scale parameter
γ	Location parameter
$\mathcal{E}_i$	Expected number of data points in bin i
Γ	Gamma function
Γ_x	Incomplete Gamma function
η	Number of houses (Zhang, 2005)
λ	Mean arrival rate of water demands at a single family household
λ^*	Arrival rate during the period of high water consumption
$\xi_{\grave{\rho}}$	$\grave{\rho}^{th}$ percentile of the Gumbel distribution (Zhang et al., 2005)
Θ_q	Coefficient of variation of PRP indoor water demand pulse
ρ	Density of liquid
$\grave{\rho}$	Percentile (Zhang et al., 2005)
$\ddot{\rho}$	Daily average utilization factor for a single family household (Zhang et al., 2005)

τ	Time at which the tap is opened
σ	Standard deviation
σ^2	Variance
Ψ^*	Dimensionless peak hourly demand factor
$\Phi(w)$	Laplace-Gauss integral

ABBREVIATIONS AND ACRONYMS

AADD	Average annual daily demand
A-D	Anderson-Darling
AWWA	American Water Works Association
AWWARF	American Water Works Association Research Foundation
CDF	Cumulative distribution function
CSIR	Council for Scientific and Industrial Research
DDA	Demand driven analysis
ee	Equivalent erven
GOF	Goodness of fit
HDA	Head dependent analysis
ICI	Industrial, commercial, and institutional
ILI	Infrastructure leakage index
IWA	International Water Association
K-S	Kolmogorov-Smirnov
MB	Mega bytes
MWD	Metropolitan water district
NFR	National Research Foundation
PF	Peak factor
PDD	Peak demand diversity
PDF	Probability density function
PMF	Probability mass function
PPH	Person(s) per household
PRP	Poisson rectangular pulse
REUM	Residential end-use model

REUWS	Residential end-uses of water study
SIMDEUM	Simulation of water demand, an end-use model
WDS	Water distribution system
WRC	Water Research Commission

1. INTRODUCTION

1.1. Background

The flow rate in a water distribution system (WDS) varies constantly, driven by fluctuating water demand. Water demand can be broken down into end-uses, where an end-use is a point where water is extracted from a WDS. In a residential setting, examples of end-use include taps, toilets, showers, baths, washing machines, dishwashers et cetera. Each time an end-use event occurs, it causes a flow rate in the WDS. When many end-uses occur simultaneously (representing peak demand), this results in a relatively large flow rate. Peak water demand is an important consideration in WDS design and analysis, since it is a factor, for example, when determining the capacity of pipelines and other infrastructure.

Various peak water demand estimation methodologies are available from design guidelines and research reports. In South Africa the most commonly used method to estimate peak demand is by means of a dimensionless peak factor (PF). The PF method involves calculating peak demand by multiplying the average water demand by a PF. The ratio between the peak water flow rate (which refers to the largest volume of flow to be received during a relatively short time period, say δt, expressed as the average volume per unit time), and the average water flow rate over an extended time period, is defined as the PF.

The magnitude of a PF is dependent on the value of δt that is used during the computation of the PF. Due to flow rate variations throughout the day, the average peak flow rate determined over a ten second time interval may be higher, when compared to the average peak flow rate determined over a five minute time interval, for example. Figure 1.1 illustrates a possible variation of instantaneous flow rate for a hypothetical residential area with an average demand of 432 $k\ell/d$. The highest five minute time interval within a 24 hour record period is depicted, together with the averaged flow rates during the peak ten second and five minute time intervals, as well as the average flow rate over

24 hours, represented by respective horizontal lines. From the figure, it is clear that the ten second peak factor ($PF_{10s} = 0.020/0.005 = 4.5$) would be larger than the five minute peak factor ($PF_{5min} = 0.017/0.005 = 3.8$). If the δt selected were too long, then the PF might not be representative of the peak water demand desired for the particular WDS analysis.

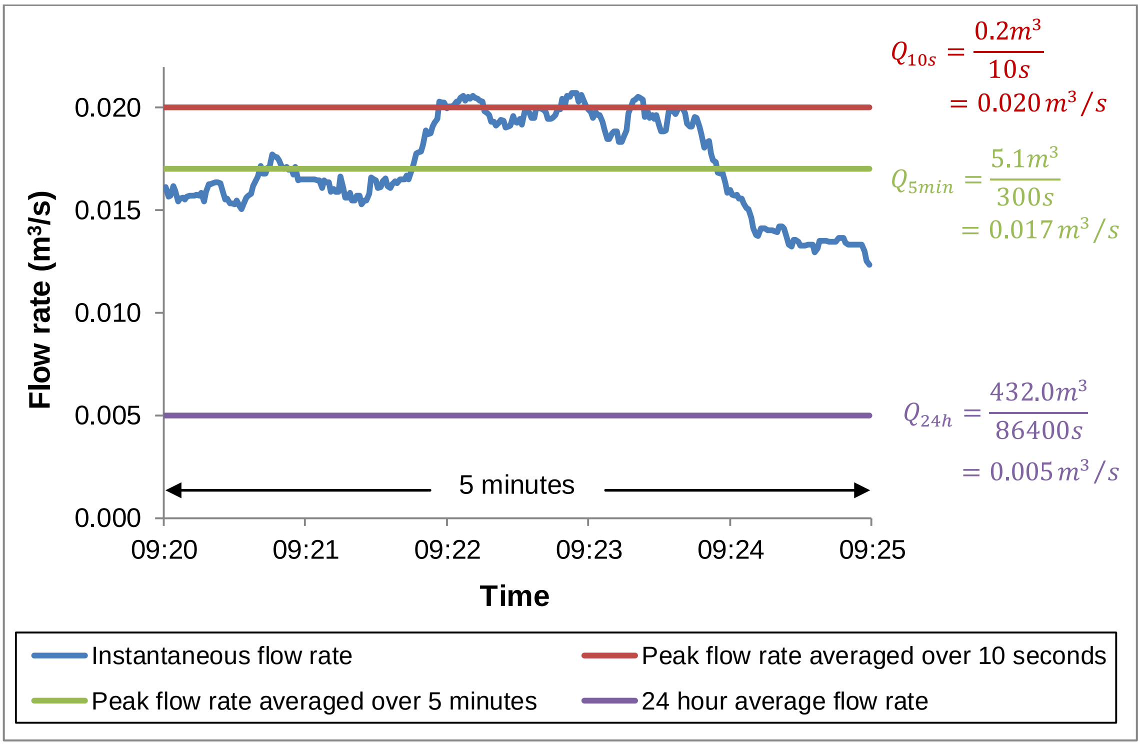


Figure 1.1: Typical diurnal flow rate variation of the highest 5 minute time interval

1.2. Terminology

Some studies use different terms to describe similar concepts. The terms defined below are used with their stated meaning in this thesis. The definitions are not comprehensive, but ensure consistency and clarity.

1.2.1. Water Demand

Billings and Jones (2008) defines water demand as the "total volume of water necessary or needed to supply customers within a certain period of time". The same definition is applied in this study.

1.2.2. Residential Water Demand

Residential water demand describes the water required per time unit by residential consumers for indoor and outdoor use. The term "residential" in this study refers to single family households. Domestic is another term used in literature to describe residential.

1.2.3. Residential Water Consumption

Residential water demand is not always metered or billed, although metering and billing is common in South Africa. Water consumption is the water flow rate that is actually utilised by consumers per time unit. Water consumption is derived from measured values obtained from a water meter or municipal treasury system. Monthly consumer water meter data has been used as the basis for various research projects locally over the past two decades (Jacobs and Fair, 2012).

1.2.4. Peak Factor

A PF is the ratio between the maximum water flow rate during a relatively short time period, say δt, and the average water flow rate during an extended observed period. The peak flow represents the period when maximum, or relatively high, flow rate occurs. In some cases the average annual daily demand (AADD) is used for the extended period; however, in this study, the extended period used as a basis for calculating the PF is taken as the average demand over one day.

1.2.5. End-use

An end-use of water is defined by Jacobs (2004) as a point (device, element, or fixture) where water is released from the pressurised water supply system to atmospheric pressure. This definition also applies to this study. The term micro-component is also used in literature to describe an end-use (Butler and Memon, 2006).

1.2.6. Diurnal Pattern

The cycle that repeats over a 24 hour period is termed the diurnal pattern.

1.3. Problem Statement

A commonly used South African guideline, the “Red Book” (Council for Scientific and Industrial Research [CSIR], 2003) relates instantaneous PFs to equivalent erven (ee), where 1 ee has an AADD of $1\ \mathrm{k}\ell/\mathrm{d}$. The CSIR (2003) lacks a definition describing the time interval, δt, that constitutes an instantaneous PF. Furthermore, the CSIR (2003) recommends applying an instantaneous PF regardless of the number of ee. This assumption is considered to be crude.

The diurnal water demand from a small number of consumers tends to be highly variable. It is expected that a relatively short time interval would be required to indentify peak events in cases with highly variable flow rates. Conversely, the aggregated water demand of many consumers tends to have a more regular diurnal pattern with less variability. Therefore, the peak event for a large number of consumers may possibly be represented adequately by using a longer time interval than that used for a small number of consumers.

A study by Booyens (2000) used measured water consumption data to investigate how the PFs changed, using different time intervals, for three residential areas consisting of 69, 444, and 794 stands, respectively. The study concluded that a time interval of 60 minutes could be used to determine the PF for residential areas that are greater than 100 ee, while a time interval of 15 minutes would be applicable to residential areas smaller than 100 ee. Booyens (2000) confirmed that the number of ee, or size of the study area, notably affected the PF.

1.4. Motivation

The concept of associating residential area sizes to PFs corresponding to particular time intervals, as Booyens (2000) suggested, would benefit by considering more than three residential area sizes. A greater number of different residential area sizes would enable a better understanding of the degree to which the PF changes with different time intervals and residential area sizes.

Limited research has been done to investigate the effect that δt have on PFs for different residential area sizes. A possible reason for this is that an empirical investigation would be very costly. Data loggers would need to record the water consumption of homogeneous residential areas of different sizes individually, with these smaller areas preferably nested within the larger areas. The logging frequency would also need to be very high to capture water flow rates over short time intervals of (say) one second.

An alternative to an empirical investigation would be to derive theoretical PFs by generating daily residential water demand profiles for individual households on a high resolution temporal scale. End-use models are based on a “bottom-up” approach, and may be a useful tool to build water demand profiles for this purpose.

The advantage of associating different time interval PFs with residential area sizes is that this may make it possible to design WDS infrastructure by choosing an applicable δt for the PF corresponding to a residential area size, instead of using an instantaneous PF in all cases, as suggested by CSIR (2003). Such a method should result in more efficient WDS design.

1.5. Research Objectives

The following research objectives were set for this study:

- To conduct a comprehensive literature review of previous work done on end-uses of water, water demand modelling (in particular end-use and stochastic models), as well as peak water demand estimation methodologies, with a focus on peak factors.
- To construct a computer based stochastic end-use model that estimates daily residential water demand for a single household on a temporal scale of one second.
- To populate the model parameters in the form of probability distributions based on recorded water consumption data, and to establish which standard distributions fit these best.
- To use stochastically generated diurnal water demand patterns from the end-use model to calculate PFs for differently sized areas by iteratively adding the water demand for individual households and using different time intervals in the PF calculation.

1.6. Delineation and Limitations

A typical urban water demand profile consists of water losses, industrial, commercial, institutional, and residential water demand. Residential water demand is therefore only one component of the total water demand that a WDS may need to cater for. This study focuses only on residential indoor water demand; the other components are beyond the scope of this study.

Residential water demand can be separated into indoor and outdoor water demand; together with leakage, which may occur both indoors and outdoors. Leakage is known to be site specific (Roberts, 2005), and the most notable leak instances flow continuously. Leaks would thus contribute to the base flow by increasing the water demand (ordinates of the demand pattern) at all abscissa without impacting the actual pattern.

Outdoor water demand is typically driven by seasonal changes and is highly dependant on climatic and geographical characteristics (Heinrich, 2007). This study considers only indoor demand which is non-seasonal, and excludes outdoor water demand and leakage. This focus on indoor consumption requires justification. A similar approach was adopted by some of the leading researchers in the field of end-use modelling (Blokker et al., 2008; Buchberger et al., 2008). In some urban metropoles such as Brisbane in Australia, permanent water conservation measures restrict outdoor use severely (Queensland Water Commission, 2012). Various levels of restriction on outdoor use have also been in place in the City of Cape Town (Jacobs et al., 2007). In contrast to outdoor use, water used indoors could be considered a basic necessity. According to White et al. (2004) outdoor water demand presents a general limitation to end-use analysis since consumer behaviour dominates outdoor demand, in contrast with the technical efficiency of equipment, which determines indoor demand.

Water flow rates in a WDS are dependant not only on the water demand, but also on the pressure in the system. If the WDS pressure were relatively low then a limited flow rate would be available, which might reduce the water consumption. There are benefits of describing peak flow rates as a function of pressure, but that is beyond the scope of this study. It is therefore assumed, for the residential end-use model developed in this study, that the pressure in the system is adequate to deliver the theoretically required peak water demand.

The probability distributions used to describe the parameters of the residential end-use model in this study were obtained from North American water measurements conducted for the Residential end-uses of water study (REUWS) by Mayer et al. (1999). The REUWS included water consumption measurements of both indoor and outdoor end-uses. The accuracy of any results in this study is therefore limited by the accuracy of the REUWS data used as input to the model. In addition, the water consumption characteristics of end-uses such as washing machines and toilets in the REUWS may be different to equivalent South African end-uses. The results of this study may, therefore, not be representative of all types of South African households, with particular limitations when comparing the results to local low income housing. To date, the REUWS is the largest database of end-uses available for this study and it has been used extensively since 1999 to conduct research into end-uses of water (Wilkes, 2005).

Microsoft Excel was used as part of this research project, to construct the residential end-use model, due to its availability and user-friendly programming style. The overall size of the resulting Microsoft Excel workbook was large, which affected the computation speed for a single execution of the model. This proved to be a limitation in that water demand patterns resulting from only 1 000 iterations could be analysed at a time, and time constraints restricted the number of executions of the end-use model that could be performed within a reasonable computing time.

1.7. Brief Chapter Overviews

This thesis comprises seven chapters and three appendices. Chapter 2 constitutes a literature review of previous work done on end-uses of water, water demand and peak water demand estimation methodologies. Chapter 3 provides an overview of relevant statistics and probability theory. The theory was applied in the study to describe discrete model input variables by known probability distribution functions. A background of the REUWS database is provided in Chapter 4.

Chapter 5 describes the methodology followed to construct and apply the end-use model; the PF calculation procedure is also provided. Chapter 6 presents a summary of the results of this study, including a comparison of how PFs for differently sized residential areas change with various time intervals. Chapter 7 concludes the findings of the study and recommendations of future work are also made.

Appendix A summarises the results of goodness of fit tests that were performed on the end-use data samples. Appendix B contains a comprehensive list of the daily event frequency and the event cycle count probability distributions for the end-use model parameters. Appendix C contains figures and tables depicting the resulting PFs for all household group sizes, and time intervals.

2. LITERATURE REVIEW

2.1. Overview of Water Distribution Systems

2.1.1. Water Service Provision

The fundamental purpose of a WDS is to provide customers with enough water to satisfy demand. This is achieved by means of three general processes. First, raw water is extracted from a source, such as a river or a dam, and transported to a treatment facility, which constitutes a bulk supply system. Secondly raw water is treated, and stored temporarily. The third process involves the water reticulation system, which transports the clean water to storage facilities. The water is then delivered to the customers. In this study the water reticulation system is referred to as the WDS.

Both the bulk supply system and WDS involve moving water through a network of linked pipes. Air valves at high points allow air to enter and exit, while drainage at the low points is facilitated by scour valves. The water is pumped at pumping stations where necessary, and stored in reservoirs and water towers.

One of the differences between a bulk supply system and a WDS is the water flow rate that the system has to facilitate. A bulk supply system consists of the main transmission lines without consumer connections. These pipes have relatively large capacities with fairly constant flow rates (Trifunović, 2006). A WDS, on the other hand, consists of smaller pipes, and directly serves the customers. The flows through these pipes are directly affected by the way customers use water over space and time. This leads to a much wider range of flow rates. The variation in flow rates should typically be incorporated in the design of a WDS.

2.1.2. Design Criteria

The design objectives of a WDS are to supply adequate volumes of water and to maintain the water quality achieved after the water treatment process (Trifunović, 2006). The engineering aspects involved in achieving the design objectives are, for example, choosing the most appropriate materials, sizes and placement of the different WDS components. In hydraulic design, this means ensuring that acceptable pressures and velocities are achieved in the pipes. A brief explanation of some of the design aspects is given below.

Flow rate is defined as the volume of fluid passing a point per second. This is expressed mathematically in (2.1).

$$Q = \frac{V}{t} \quad (2.1)$$

where:

Q = flow rate (m^3/s)

V = volume (m^3)

t = time (s).

Considering the conservation of mass, or the continuity equation, the flow rate can also be expressed as the product of the velocity of the fluid and the area of the pipe, as shown in (2.2).

$$Q = v \times A \quad (2.2)$$

where:

Q = flow rate (m^3/s)

v = velocity (m/s)

A = area (m^2).

The Bernoulli equation represents the continuity of energy and can be written as:

$$\frac{\mathcal{P}}{\rho g}+\frac{v^2}{2g}+z=constant \quad (2.3)$$

where:

$\mathcal{P}$ = pressure (N/m^2)

ρ = density of liquid (kg/m^3) (kg/m^3)

g = gravitational acceleration (m/s^2)

v = velocity (m/s)

z = height above datum (m).

From (2.3) it can be seen that as the velocity of the water increases, the pressure in the pipe decreases, and vice versa. Put another way, when the demand for water is at a maximum (peak flow rate), the pressure in the system is at a minimum. Relatively high and low pressures in a WDS have adverse repercussions on operation and maintenance. High pressures cause an increase in leakage and water losses, or pipe breaks. Negative pressures in a system can lead to pipe collapse, or may draw pollutants into the system. Customers also experience limited flow rates at low pressures, and some appliances fail to operate (Jacobs and Strijdom, 2009).

It follows from (2.2) that, relative to a fixed flow rate, a pipe with a small diameter will result in the water having a high velocity, while a large diameter pipe will result in the water having a low velocity. Exceedingly low velocities in pipelines cause sediment deposition which, in turn, leads to water quality degradation. Exceedingly high velocities increase the pipe head losses and are related to problems with water hammer.

To prevent such adverse effects and to ensure that a WDS operates satisfactorily, minimum and maximum pressures and velocities are prescribed for the design of a WDS. The pipe diameters are chosen such that the magnitude of the water velocity and subsequent pressure is within a prescribed desired range most of the time, even with varying flow rates. Since pressure and velocity are affected by the flow rate, a WDS is typically assessed against limiting demand conditions.

Limiting demand conditions can be described as the worst case water demand scenarios. These are used in design because of the assumption that if the system can operate at the limiting conditions, it will operate properly most of the time. Perelman and Osfeld (2006) investigated the hypotheses that if a system is operated at a load condition such as peak flow, then it will function properly at any other load condition. The authors concluded that for the purpose of designing pipes, steady state simulation runs for peak flow may be acceptable. Full extended period simulation is, however, important in order to check the behaviour of tanks, pumps and valves. Examples of limiting conditions that are often used are the fire flow rate, the storage capacity replenishment rate, and peak flow rates such as instantaneous peak flow rate $(Q_{inst})_{max}$, peak hour flow rate $(Q_h)_{max}$, and peak day flow rate $(Q_d)_{max}$. Burn et al. (2002) state that in cities with high living standards, the accepted norm for WDS design is based on peak flow. Hyun et al. (2006), as well as Johnson (1999), agree that peak day demands should be used to design bulk water supply pipelines.

Peak flow rates are differentiated according to the time interval (δt) over which the flow is measured. The American Water Works Association (AWWA) defines instantaneous peak flow rate as the rate of water measured at a particular moment in a day (AWWA, 1999). There is, however, no precise definition of which value of δt would sufficiently describe the instantaneous peak flow rate. Some studies include flows that are measured within a 10 second interval as instantaneous peak flow rates, as is the case in Mayer et al. (1999). Peak hour flow rate is defined as the consecutive 60 minutes of a day during which demand is at the highest. Peak day flow rate is similarly defined as the consecutive 24 hour period in a year during which demand is at the highest (AWWA, 1999).

Some water supply systems make provision for fire protection services. Fire fighting requires large volumes of water at very high flow rates. Since fire flow rates are often much larger than normal water demand, this is often the most limiting demand condition in a system.

Storage facilities designed for peak day flow rates are filled when the water demand is less than the average peak day demand, and emptied when water demand is greater than average peak day demand. The flows required to replenish the storage facility within a certain timeframe can sometimes be a limiting condition on pipelines (AWWA, 1989).

The sizes of pipes are determined by the volume and rate of flow expected in the system. AWWA (1989) recommends that pipes be designed based on the highest flow rate resulting from peak day flow rate plus fire flow rate, maximum storage-replenishment rate, or peak hour flow rate. Thereafter, it must be ensured that limits such as maximum velocities and head losses are adhered to. Burn et al. (2002) compared the reticulation pipeline costs for a cluster of 4 000 households, based on varying peak demand scenarios. The authors concluded that system cost savings of 25-45% could be achieved by lowering the peak demand for which the pipes were designed.

Pumps are required to fill storage facilities, and ensure that pressure is maintained in the system to allow the movement of water. The choice of pump size is dependent on many factors, such as the source capacity, storage availability, and peak demand (AWWA, 1999). According to AWWA (1989) pumps should be sized based on the maximum flow resulting from peak day flow rate, peak day flow rate plus fire flow rate, or peak hour flow rate.

Storage facilities enable pumps to operate at an average rate and not just during peak periods. Reservoirs and tanks are normally sized considering average, peak and fire flow rates, as well as emergency reserves in case of treatment plant or source failure (AWWA, 1999; CSIR, 2003). It is recommended that the limiting condition on system storage be the highest flow rate resulting from peak hour flow rate, or peak day flow rate plus fire flow rate (AWWA, 1989).

2.1.3. Intermittent Supply

The design of a WDS, as discussed in section 2.1.2, is based on the assumption of continuous water supply; in other words, the pipes remain full of pressurised water. In many developing communities and water stressed countries, the supply of water is not continuous, but rather intermittent. Intermittent water supply entails physically cutting off the water supply to customers for various periods, due to a lack of system capacity. When limited water is available, an intermittent system is one method of controlling water demand, and is usually a matter of necessity.

Intermittent systems have a number of serious problems such as low pressure, inequitable water distribution, water contamination and additional customer costs (Vairavamoorthy et al., 2007).

The demand at the nodes of the network is not driven by diurnal fluctuations based on consumer patterns. Instead the demand is dependent on the maximum amount of water that can be collected during the time of supply. The quantity of water collected is thus dependent on the pressure available at their point of abstraction. Therefore, when analysing the network, a demand driven approach should not be used, but rather a head dependent approach (Vairavamoorthy et al., 2007). The method of deriving peak flows as done in this research would thus be inappropriate for analysis of intermittent WDSs.

2.1.4. Demand Driven Analysis Versus Head Dependent Analysis

The operation of a WDS is often simulated and analysed using a computer model representing the network hydraulics. Algorithms of such software are usually based on demand driven analysis (DDA). DDA means that the demands allocated at the nodes of a network are assumed to be fully satisfied and remain constant. The resulting pipe flow and nodal pressures are, therefore, consistent with the demands calculated, and it is assumed that there is sufficient pressure in the system to deliver all of the demand (Tanyimboh et al., 2003). The DDA calculation procedure deals with pipe flows against the hydraulic gradients, and the pressure is calculated afterwards. Because the relation between pressure and demand is ignored, error is introduced in the model. Sufficient pressure is not always available in the system, and if the demand were to exceed the capacity of the system, then DDA would no longer be representative of the system performance.

The procedure whereby a WDS is analysed when taking pressure-related demand into account is known as head dependent analysis (HDA). If hydraulic calculations are done with DDA, some nodes may depict negative pressures, which is impractical. The HDA approach aims to determine, for each node, an outflow which is compatible with the outflows at the rest of the nodes in relation to the available pressure in the system (Tanyimboh et al., 1999).

Applying HDA causes a gradual reduction of the discharges at the nodes and the hydraulic gradient values. The typical relationship expressed by Chandapillai (1991) is given in (2.4):

$$H_j = H_j^{min} + K_j Q_j^{n_j} \qquad (2.4)$$

where:

Q_j = demand at node j

H_j = head at node j when the demand of that node is Q_j

K_j = flow resistance coefficient

n_j = exponent

H_j^{min} = nodal head below which the outflow at the node is unsatisfactory or zero.

Gupta and Bhave (1996) described the main methods for solving networks by means of the HDA approach.

HDA is advantageous because it can accurately determine the maximum amount of water that a system can provide for various minimum pressures. It can also identify the precise nodes with insufficient flow. According to Tanyimboh et al. (1999) this makes the results obtained from HDA superior to those of DDA.

Despite the benefits that HDA has portrayed, a comprehensive investigation of peak flow rate as a function of pressure is beyond the scope of this work. Engineers in practice remain proficient in the application of DDA and estimated peak flows based on PFs. This method is expeditious and the results obtained are considered acceptable in view of other uncertainties incorporated during WDS analysis.

2.2. Basic Concepts of Water Demand

2.2.1. Water Loss and Leaks

Water losses in municipal WDSs are a worldwide problem. The International Water Association (IWA) formed a Water Loss Task Force in 1996, to develop international best practices in the field of water loss management. The IWA Task Force published a “best practice” standard water balance, given in Figure 2.1.

<table>
<tr><td rowspan="9">System Input Volume (corrected for known errors)</td><td rowspan="4">Authorised Consumption</td><td rowspan="2">Billed Authorised Consumption</td><td>Billed Metered Consumption (including water exported)</td><td rowspan="2">Revenue Water</td></tr>
<tr><td>Billed Unmetered Consumption</td></tr>
<tr><td rowspan="2">Unbilled Authorised Consumption</td><td>Unbilled Metered Consumption</td><td rowspan="7">Non-Revenue Water</td></tr>
<tr><td>Unbilled Unmetered Consumption</td></tr>
<tr><td rowspan="5">Water Losses</td><td rowspan="2">Apparent Losses</td><td>Unauthorised Consumption</td></tr>
<tr><td>Customer Metering Inaccuracies</td></tr>
<tr><td rowspan="3">Real Losses</td><td>Leakage on Transmission and/or Distribution Mains</td></tr>
<tr><td>Leakage and overflows at Utility’s Storage Tanks</td></tr>
<tr><td>Leakage on Service Connections up to the point of Customer metering</td></tr>
</table>

Figure 2.1: IWA standard water balance (McKenzie and Lambert, 2004)

Performance measurement indicators and strategies to reduce water loss were also developed by the IWA Task Force. One such Performance Indicator for real losses is the infrastructure leakage index (ILI). The ILI was described by McKenzie and Lambert (2004) as the ratio of the the current annual real losses to the unavoidable annual real losses (the theoretical minimum leakage that can be achieved). An ILI of one therefore suggests the ideal leakage situation, with increasingly poor performance corresponding to higher ILIs.

McKenzie and Seago (2005) determined the ILI values for 30 water utilities in South Africa. The results for 27 of these utilities are presented in Figure 2.2. The average ILI for the South African utilities is about 6.3. This was compared with average ILI values for selected utilities in North America, Australia and England, which were 4.9, 2.9, and 2.6, respectively. McKenzie and Seago (2005) are of the opinion that ILI values below two would be unusual in South Africa, and that utilities that are in a reasonable condition would have an ILI value of around five.

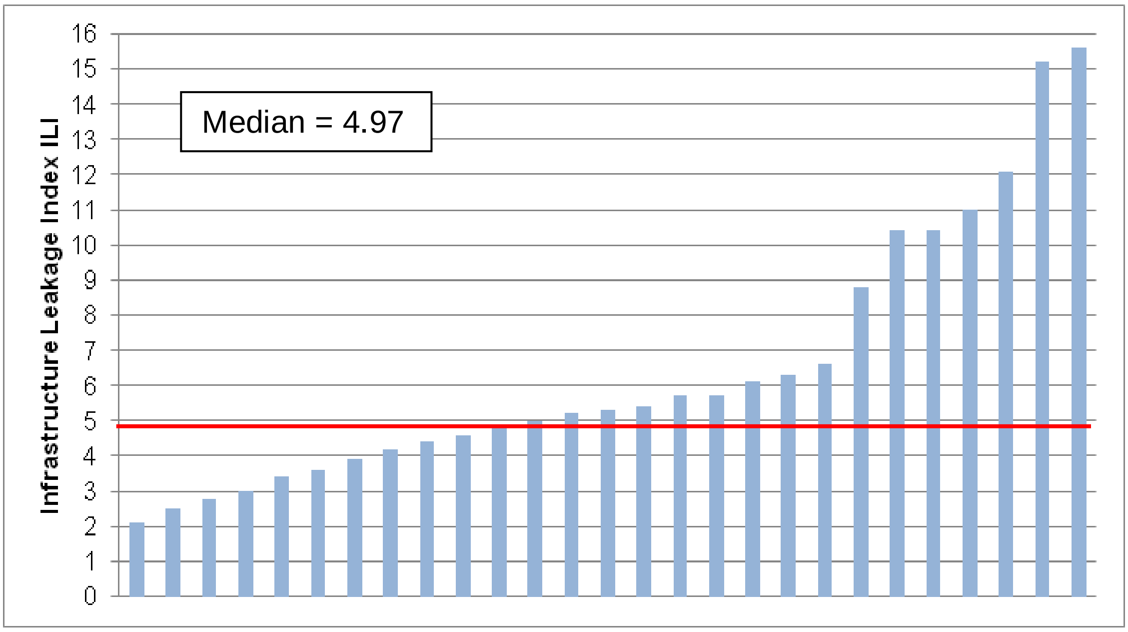


Figure 2.2: ILI results for South African WDS (McKenzie and Seago, 2005)

According to the IWA standard water balance, on-site leakage (on consumer's properties) is considered part of revenue water. In fact, where such losses are billed, municipalities may benefit from on-site leakage (Lugoma et al., 2012). However, in cases where consumers do not pay for their water, it is often in the municipality's interest to repair on-site leakages itself.

Lugoma et al. (2012) investigated on-site leakage in well-established Johannesburg suburbs. The study determined the on-site leakage by analysing the readings on relatively new municipal water meters. The average leakage of 182 properties was found to be approximately 25% of the measured consumption.

A number of studies investigated the leakage of individual households. Table 2.1 summarises the average leakage, as a percentage of total demand, that these studies observed.

Table 2.1: Examples of on-site leakage as percentage of total demand

Reference	Study area		Leakage (%) of total demand
DeOreo et al. (1996)	USA, 16 homes		2.3
Mayer et al. (1999)	USA/Canada, 1188 homes		5.5
DeOreo et al. (2001)	Pre-retrofit	USA, 37 homes	10.3$^{(1)}$
DeOreo et al. (2001)	Post-retrofit	USA, 37 homes	5.5$^{(1)}$
Loh and Coghlan (2003)	Australia, 120 homes		2.3
Roberts (2005)	Australia, 99 homes		5.7
Heinrich (2007)	New Zealand, 12 homes		3.7
Willis et al. (2009)	Australia, 151 homes		1.0

Note: $^{(1)}$Leakage as a percentage of indoor demand only

Leakage is very varied, even within homogeneous areas. Observations by DeOreo et al. (1996) and Heinrich (2007) were that the majority of leakage volume in their study area was contributed by only a few of the houses, and that the leakage in homes often arises from toilets. Britton et al. (2008) identified different types of leaks that originated from irrigation, hot water systems, meters, toilets, taps and pipes. Of the different types of leaks, 46% was attributed to toilet leaks.

Despite agreement in literature that toilets contribute notably to residential water leakage, it is difficult to estimate leakage, since it is site specific. Leakage therefore lends itself to being a component that can be added separately at the end of a water demand estimation procedure. For this reason water losses are excluded in the water demand computations for this study.

As previously illustrated by the IWA water balance, total water consumption consists of authorised consumption and water losses. Authorised consumption is divided into billed and unbilled consumption. An example of unbilled authorised consumption is the water required for fire fighting.

2.2.2. Fire Flow Requirements

Fire flow requirements are typically specified in design guidelines applicable to a specific region. In small networks fire flow is sometimes omitted due to budget constraints. It is computed separately from the metered water demand estimation, which is later applied as one of the limiting demand conditions. Myburgh (2012) conducted a detailed investigation into local fire flow requirements, but fire flow is not pertinent to the outcome of this study and therefore is not elaborated on further.

2.2.3. Water Demand Categories

Water demand characteristics are often used as a means to divide water customers into categories. Examples of categories often used are residential and non-residential water consumers. Non-residential water consumers can be further divided into industrial, commercial, and institutional (ICI) sectors. Typical ICI customers include shops, restaurants and offices that use water for toilets, cleaning and cooking, but also for production processes that may have relatively high volume water requirements. In addition, van Zyl et al. (2007) categorises farms, parks, educational, and sports users as non-residential water consumers.

CSIR (2003) provided non-domestic water demand estimation guidelines for both developing and developed areas in South Africa. According to van Zyl et al. (2007), non-domestic use is very difficult to estimate, and recommends that field measurements are used for estimation purposes.

In a study by van Zyl et al. (2007) forty eight municipal treasury databases were used to obtain water consumption data for non-domestic consumers. Climatic and socio-economic data was also linked to the consumption data. The non-domestic users were grouped into seven categories, namely industrial; business commercial; government and institutional; farms; parks; education; and sports. Frequency distributions were plotted using the natural logarithm of the AADD for all categories. The frequency distributions were described well by Log-Normal probability distributions in all cases.

Residential consumers may be single family households, or multiple family units such as apartments. This study focuses on single family households only.

2.2.4. Residential End-Uses

Residential water consumption emanates from water used by a consumer at various end-use points on a residential property. Typical end-uses found inside and outside the home are illustrated in Figure 2.3.

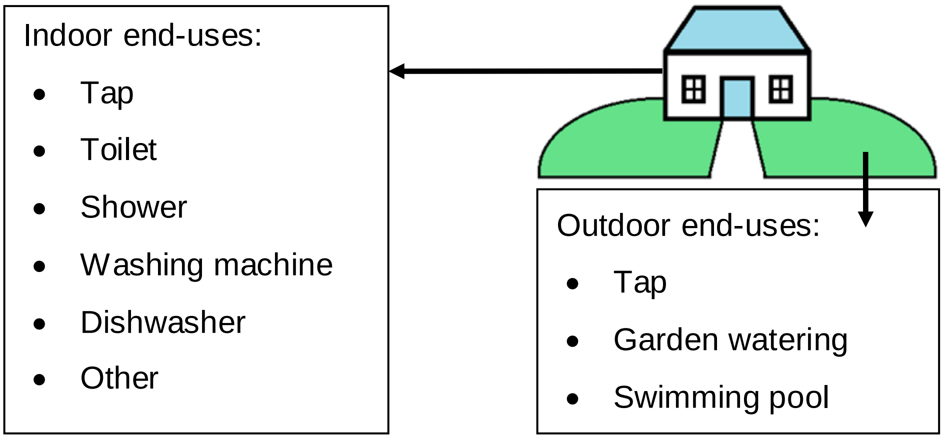


Figure 2.3: Examples of typical indoor and outdoor end-uses

A number of authors have reported field measurements of water used by the different end-uses. Different methods are available for measuring water consumption at the resolution of individual end-uses. A direct method was used by Edwards and Martin (1995) who measured the flow at each appliance separately by placing water meters at each end-use in the home. In that study an average of 14 water meters was placed in each of 100 sample households in the United Kingdom. Water volume was measured in 15 minute intervals over a one year period from October 1993 to September 1994. The publication by Edwards and Martin (1995) is relatively old, but remains impressive in terms of the scope and extent of the work.

An alternative method, called flow trace analysis, was used by DeOreo et al. (1996), Mayer et al. (1999), DeOreo et al. (2001), Loh and Coghlan (2003), Roberts (2005), Heinrich (2007), and Willis et al. (2009). Flow trace analysis is a process whereby a data logger is attached to a municipal water meter at a customer's residence. The data logger records the volume of water passing the water meter in a specified time interval, such as every 10 seconds. Software is then used to analyse the flows recorded by the data logger, disaggregate the flow and assign it to specific end-uses. Trace Wizard is an example of software designed for this purpose, which is described in detail in section 2.3.2.

The flow trace analysis concept is based on the premise that each end-use causes a unique flow pattern (or flow trace), which can be used to identify it by means of pattern recognition in a data time series. For example, when a tap is used the flow will be of short duration and relatively small flow rate. A toilet cistern filling after a flush will be within a particular volume range, and with a consistent flow rate. The flow trace corresponding to each end-use is initially defined in Trace Wizard. Thereafter Trace Wizard identifies flow traces within the flow data time series and assigns end-uses to every water consumption event. For each event, its statistics are calculated. These include the event's start time, stop time, duration, volume, peak flow rate, mode flow rate, and mode frequency.

A brief description of the most notable studies making use of flow trace analysis and other methods is given below:

- DeOreo et al. (1996) used data loggers to measure the flow rates from residential water meters in 10 second intervals in Boulder, Colorado, USA.

Sixteen single family households were each logged for a total of three weeks in the summer between June and September 1994. Flow trace analysis was used to identify signatures corresponding to individual flow events.

- Mayer et al. (1999) used flow trace analysis to obtain individual water consumption events from twelve study sites across the United States and Canada. Measurements were taken for two weeks in the summer and two weeks in the winter for about 1 200 single family households. In addition, 6 000 participants completed surveys detailing household level information, and water billing records were obtained for 12 000 households. This study by Mayer et al. (1999) became widely known as the REUWS.
- The flow trace analysis technique was used again by DeOreo et al. (2001) to disaggregate end-uses. The focus of the work was to determine the amount of water saved on each end-use after houses were retrofitted with high efficiency water-saving appliances. The investigation was carried out on 37 single family households in Seattle, USA. The pre- and post-retrofit measurements for the end-uses were compared to the REUWS by Mayer et al. (1999).
- Mayer et al. (2003) conducted a water conservation study that investigated the effect of water saving appliances. Flow trace analysis was used to compare the end-use water consumption before and after retrofitting the appliances in 33 single family households in East Bay Municipal Water District, USA. The impact that indoor water conservation measures had on both individual and aggregate water consumption patterns was investigated.
- Loh and Coghlan (2003) carried out a study in Perth, Australia, using low, middle and high income houses. A sample of 120 single family households was used to obtain water consumption measurements from November 1998 to June 2000. Flow trace analysis was used to disaggregate individual flow events from separate appliances. Household information from an additional 600 houses was obtained through questionnaires.
- The Yarra Valley in Melbourne, Australia, was used as the site for a water measurement study by Roberts (2005). Data loggers were installed in 100

homes, and measurements were taken for two weeks in February and two weeks in August to represent summer and winter usage, respectively. Data was collected at five second intervals, which enabled flow trace analysis to be performed.

- Willis et al. (2009) performed an end-use water consumption study on the Gold Coast near Queensland, Australia. A total of 151 households, consisting of both single reticulated (38) and dual reticulated (113) systems, was monitored for a two week period in the winter of 2008. Data loggers with a 10 second reading frequency enabled flow trace analysis to be performed.
- End-use flow measurements were taken from 12 residential households on the Kapiti Coast near Wellington, in New Zealand, by Heinrich (2007). Flow trace analysis was performed based on measurements taken at ten second intervals over two seasonal monitoring periods. The winter period extended from mid-July to mid-October 2006, while the summer measurements took place from mid-November 2006 to the end of February 2007.

2.2.5. Factors Affecting Water Demand

A comprehensive list of factors influencing peak water demand is provided by Day and Howe (2003). This study is concerned with how end-use events act as building blocks to construct a demand pattern, implying that it is more important here to better understand these end-use events and the nature of their occurrence, than to address factors that influence water demand on a larger spatial scale.

2.2.6. End-Use Frequency and Volume

The research in this thesis addresses the theoretical derivation of peak flows from a stochastic description of end-use events. For this reason it is considered important to present a review of end-use information from earlier studies. Tables 2.2 to 2.4 provide summaries of the useful information that was compiled in this regard.

Table 2.2: Examples of reported end-use volumes per event

Author Citation	Average volume of water per end-use event ($\ell/event$)					
	Toilet	Shower	Washing Machine	Tap	Dish-washer	Bath
DeOreo et al. (1996)	16.0	61.0	-	-	-	-
Mayer et al. (1999)	13.4	66.3	157.6	-	-	-
DeOreo et al. (2001) Pre-retrofit	13.7	-	155.0	-	-	-
DeOreo et al. (2001) Post-retrofit	5.2	-	92.0	-	-	-
Loh and Coghlan (2003)	10.0[(1)]	60.0[(2)]	150.3 / 57.8[(3)]	-	-	-
Mayer et al. (2003) Pre-retrofit	15	71.0	156.9	-	34.3	109.8
Mayer et al. (2003) Post-retrofit	6.4	59.1	30.7	-	-	105.2
Roberts (2005)	7.6	67.5[(4)]	143.0	1.3	23.9	123.0
Heinrich (2007)	6.2	82.0	134.0 / 50.0[(3)]	1.6	-	-

Note: [(1)]Single flush toilet sample.

[(2)]Normal flow rate shower sample.

[(3)]Average volume for automatic top loader and front loader respectively.

[(4)]Calculated as product of average duration and average flow rate.

Table 2.3: Examples of reported end-use volumes per capita

Author Citation	Average volume of water per capita per day for selected end-uses ($\ell/c \cdot d$)					
	Toilet	Shower	Washing Machine	Tap	Dish-washer	Bath
Edwards and Martin (1995)	47.9	5.8	30.5	36.3[(1)]	1.5	18.9
DeOreo et al. (1996)	26.3	17.4	24.8	14.7	3.0	2.3
Mayer et al. (1999)	71.3	44.7	57.8	42.0	3.9	4.6
DeOreo et al. (2001) Pre-retrofit	71.2	34.1	56.0	34.8	5.3	14.0
DeOreo et al. (2001) Post-retrofit	29.9	32.9	34.8	30.3	4.5	10.2

Author Citation	Average volume of water per capita per day for selected end-uses ($\ell/c \cdot d$)					
	Toilet	Shower	Washing Machine	Tap	Dish-washer	Bath
Loh and Coghlan (2003)	33.0	51.0[2]	42.0	24.0	-	-
Mayer et al. (2003) Pre-retrofit	76.7	46.2	53.6	40.5	3.9	11.6
Mayer et al. (2003) Post-retrofit	37.8	41.2	33.9	40.5	3.4	10.8
Roberts (2005)	30.0	49.0	40.0	27.0	3.0	3.0
Heinrich (2007)	33.4	67.8	40.9	23.5	2.4	4.3
Willis et al. (2009)	21.1	49.7	30.0	27.0	2.2	6.5

Note: [1]Combination of kitchen taps and bathroom taps.

[2]Combination of showers and baths.

Table 2.4: Examples of reported end-use frequencies

Author Citation	Average end-use event frequency ($frequency/c \cdot d$)				
	Toilet	Shower	Washing Machine	Bath	Dish-washer
DeOreo et al. (1996)	3.8	0.7	0.3	-	0.2
Mayer et al. (1999)	5.1	0.8[1]	0.4	-	0.1
DeOreo et al. (2001) Pre-retrofit	5.2	-	0.4	-	-
DeOreo et al. (2001) Post-retrofit	5.5	-	0.4	-	-
Mayer et al. (2003) Pre-retrofit	5.1	0.7	0.4	0.1	0.1
Mayer et al. (2003) Post-retrofit	5.7	0.7	0.6	0.1	-
Roberts (2005)	4.2	0.8	-	-	-
Heinrich (2007)	4.7		0.3		

Note: [1]Combination of showers and baths.

2.2.7. Temporal Variation in Demand

Variation in water demand can be attributed to spatial and temporal variations. Spatial variation is largely caused by differences in climatic variables. Temporal variation causes daily, weekly, and seasonal cyclic patterns, which are discussed in more detail below.

Examples of domestic water activities are toilet flushing, showering, hand washing, teeth brushing, laundry, cooking, drinking, et cetera. When any one of these activities is executed, a corresponding flow of water through the pipes is expected. It is unlikely that all the activities will occur simultaneously in one household. However, a combination of activities such as showering, teeth brushing and toilet flushing may typically coincide. In such a case, the instantaneous flow equals the sum of the flows for the various activities. By assessing instantaneous demands, a demand pattern can be built up for each house and, ultimately, for an entire distribution area (Trifunović, 2006).

For one home, or a small residential area, the exact time when water is used is unpredictable. However, people tend to have periodic activities which translate into their water using schedules. As the number of consumers increases, the demand pattern becomes more predictable, and clear daily water demand patterns emerge. For design purposes, Trifunović (2006) believes it is a valid assumption that a similar water demand cycle will be followed by residents over time.

Water demand generally tends to be more frequent in the mornings when people wake up, when they return home from work or school in the afternoons, and before they go to sleep in the evenings, than for other times of the day. Strong diurnal and weekly patterns, according to Race and Burnell (2004), reflect residential lifestyles. An example of a typical residential diurnal water demand pattern is shown in Figure 2.4. Seasonal differences are also presented in the figure.

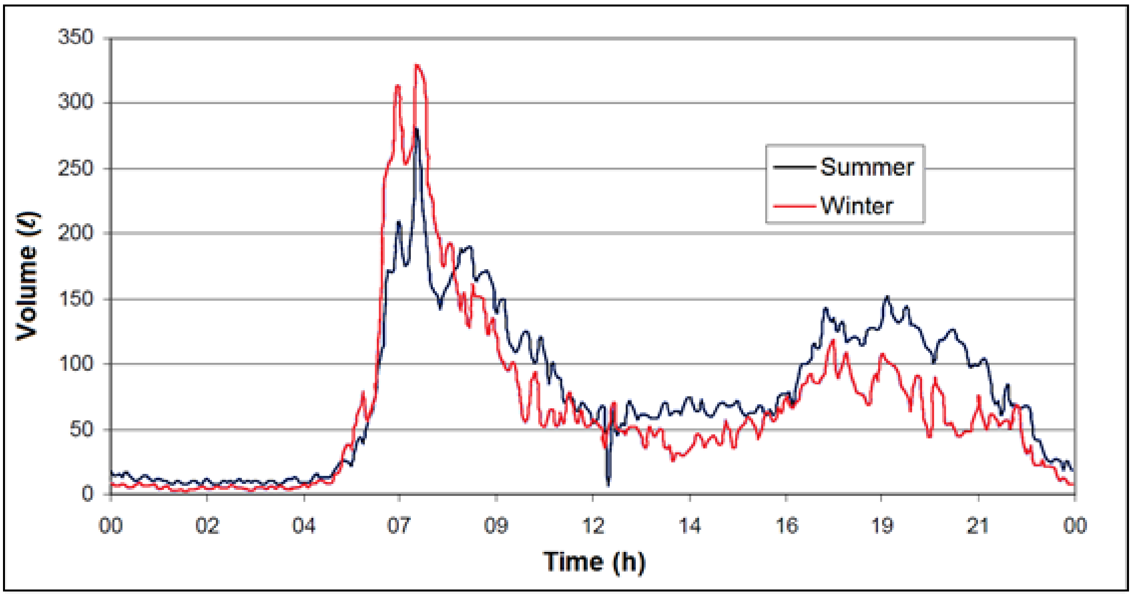


Figure 2.4: Typical diurnal water demand pattern adapted from Heinrich (2007)

Bowen et al. (1993) investigated the water consumption patterns of residential homes in five cities in the United States. A well-defined and consistent residential water consumption pattern was observed across all regions. The author described diurnal water demand by identifying four periods:

- The first period is night time from 23:00 to 05:00, and the lowest usage occurs during this period.
- In the morning from 05:00 to 12:00 there is a sharp rise in usage, with daily peak hourly usage normally occurring between 07:00 and 08:00.
- The usage then decreases, with continuous moderate usage from 12:00 to 17:00, and local minima were observed around 15:00.
- In the evening from 17:00 to 23:00 the usage increases again, a secondary peak is typically observed from 18:00 to 20:00.

Diurnal water demand patterns vary spatially. No two towns will necessarily have the same pattern. Bowen et al. (1993) noted that slight differences occurred across geographic regions. For example, the timings of the cycles, or peak values may differ, but the basic characteristics of the four periods remained the same. Mayer et al. (1999), for example, found that the diurnal pattern in their study exhibited the same four characteristics, but the moderate afternoon usage was defined from 11:00 to 18:00.

The daily water consumption pattern of individual household appliances was investigated by Mayer et al. (1999). Toilet use was the largest component of indoor use, and displayed a peak between 07:00 and 10:00, with a secondary peak between 17:00 and 23:00. Washing machine use peaked between 09:00 and 13:00, and remained moderately high until 21:00. Between 06:00 and 11:00 shower usage was relatively high, and a lower peak was evident between 18:00 and 23:00. Taps were used relatively consistently throughout the day, with a slight peak in the mornings and evenings.

Weekly demand patterns are influenced by working and non-working days. Usually Mondays to Fridays are working days, with very distinct diurnal cycles, as discussed previously. Festive holidays and sporting events have their own unique patterns, and these impact weekly cycles. On non-working days, such as Saturdays and Sundays, water demand is spread more evenly throughout the day, since people are home for a longer period of time. Higher peaks are, therefore, normally experienced on working days.

Loureiro et al. (2006) developed a water consumption characterisation program in Portugal and performed a demand analysis of the available data. Measured flow data was collected at ten to 15 minute time intervals for 20 metering districts which ranged in size between 2 000 and 12 000 connected properties. The water consumption data was compared with other variables such as socio-demographic data to derive daily consumption patterns. The daily water consumption pattern for different days of the week in the summer, in an average socio-economic area, is shown in Figure 2.5. It is clear that the pattern of water demand during the weekends is very different from that during workdays.

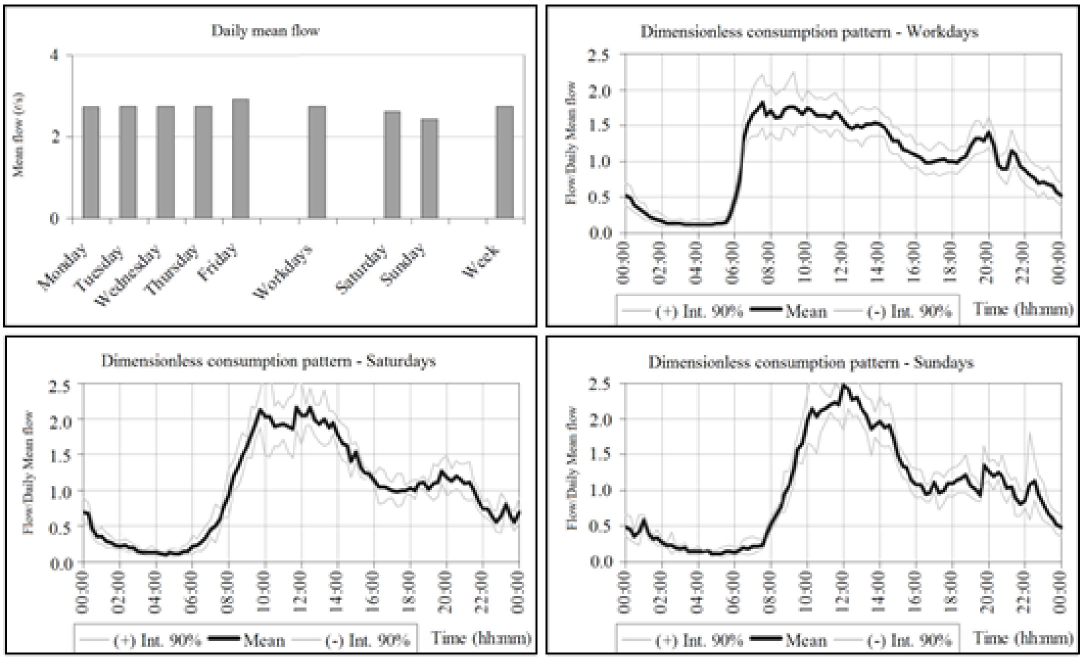


Figure 2.5: Weekday water demand variation adapted from Loureiro et al. (2006)

Variations in average water demand are also notable throughout the year. Higher temperatures in the summer months lead to increased outdoor usage, and cause distinct seasonal patterns. Indoor water demand is not affected as much by seasonal changes (Roberts, 2005). During hot seasons higher water consumption can also be observed in certain areas, due to a temporary increase in the number of consumers. This is typical of popular holiday destinations, which also exhibit unique peak demands.

According to Mayer et al. (1999) both indoor and outdoor water demand follows a diurnal pattern, but these peak at different times. For instance, outdoor use increases sharply from 05:00, while indoor use only increased after 07:00. Roberts (2005) divided water demand into seasonal and non-seasonal categories, where seasonal use was equivalent to outdoor use, and included indoor seasonal appliances such as evaporative air conditioners. Non-seasonal use is, therefore, not affected by annual cycles in water demand. Roberts (2005) considers the seasonal use empirical measurements of little consequence, since drought restrictions affect typical garden irrigation habits. This study only considers the non-seasonal component of demand, since outdoor water demand is excluded, as mentioned in section 1.6.

2.3. Models Available for Water Demand Analysis

2.3.1. Swift

SWIFT is a commercial software product developed by GLS consulting engineers (GLS Software, 2012). The software can access municipal treasury databases, where demographic and water consumption data on a large number of users can be obtained. SWIFT contains the data of every stand in the respective municipal treasury database. Such information includes the owner, consumer, address, land-use, zoning, consumption, tax tariffs, the value of the stand and any improvements. The data it contains relating to the meters includes the meter readings, the meter serial numbers, and the date of installation.

The user can view the information in a structured data table. Since SWIFT was designed with infrastructure managers as the users in mind, the data can be sorted, queried, and saved in reports. There are functions that enable the integrity of recent readings to be checked, by comparing them against historical meter records. Analyses can also be performed on any database by customising various settings. Jacobs and Fair (2012) presented a detailed account of Swift, and its impact on local research, for further reading.

2.3.2. Trace Wizard

Flow trace analysis is a means by which consistent flow patterns are isolated, identified and categorised. Software called Trace Wizard was specifically developed by Aquacraft Pty. (Ltd.) for this purpose. Trace Wizard was used by Mayer et al. (1999) who describes the process as follows:

Raw flow data obtained from water meters and loggers are disaggregated by Trace Wizard into individual water consumption events. For each event, its statistics are calculated. These are start time, stop time, duration, volume, peak flow rate, mode flow rate, and mode frequency.

Each study residence is then given a set of parameters (such as the volume, duration and peak flow rate of each end-use). This allows Trace Wizard to categorise each event based on its flow pattern, and assign it to a specific household end-use. Signature flows for each appliance can also be recorded when data loggers are first installed. The program uses these signature pulses to distinguish flow traces into various events such as a toilet event, leak event, tap event, et cetera.

Because of the unique parameters for each end-use, Trace Wizard can identify simultaneous events. A limitation of the measurement technology, however, is that there is no discrimination between taps such as bathroom tap, kitchen tap, or laundry tap. An example of the Trace Wizard output is shown in Figure 2.6. The separate end-uses are displayed in different colours.

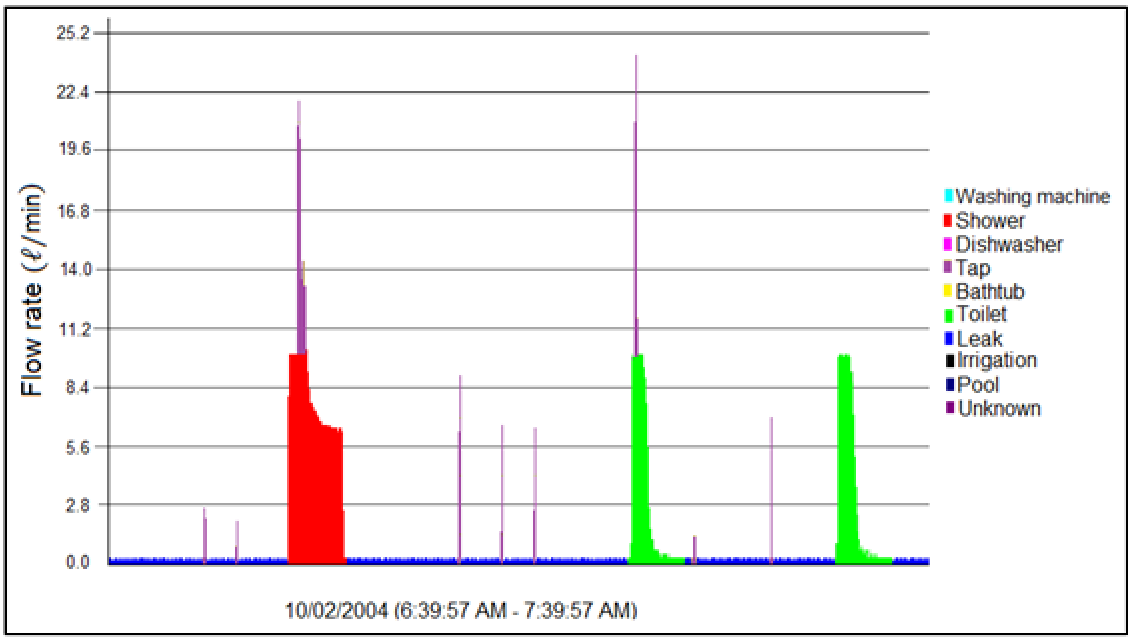


Figure 2.6: Example of Trace Wizard analysis result adapted from Roberts (2007)

2.3.3. REUM

Jacobs (2004) developed a first of its kind Residential End-Use Model (REUM). The model required the input information of 16 end-uses on a single residential stand. The end-uses included in the REUM are listed as: bath; bathroom basin;

dishwasher, kitchen sink; leaks; miscellaneous indoor; shower; washing machine; toilet flush (large); toilet flush (small); miscellaneous outdoor; pool filtering; pool evaporation; lawn; garden beds; and fruit trees or vegetables. The REUM estimated five different components of residential water demand and return flow. These were indoor water demand, outdoor water demand, hot water demand, wastewater flow volume and concentration of wastewater solutes.

Each of the indoor end-uses was modelled by four parameters which described the presence, volume of use, frequency of use, and quantity of that end-use. Each of the outdoor end-uses is modelled by five parameters which are the garden irrigation factor, the vegetated surface area, crop factor, monthly rainfall, and pan evaporation. An additional three parameters per end-use, model the hot water demand. The modelling of the component for waste water flow, and wastewater total dissolved solids concentration, each require one parameter for each end-use. A total of 111 parameters are therefore required to populate the REUM and model one month. The values for the parameters could be estimated by physically measuring the the values, through contingency evaluation, and by subjective evaluation based, on knowledge of the end-use and experience. Detailed analysis could be performed, as a result of the large number of input parameters.

The REUM was applied in a study by Jacobs et al. (2006). Questionnaires were completed by residents of 160 properties in Cape Town, and the responses were used as inputs to the REUM to estimate the demand. The modelled results were then compared to water meter information for the properties. Questionnaires that were distributed by hand received responses from 11 pilot study water consumers with a technical background and 117 low-income water consumers. The questionnaire was also available on the City of Cape Town’s website, which received 32 responses. The end-use results compared well with the measured data for one group, which indicated that some customers were able to provide a good estimate of their own water consumption.

2.3.4. Nonhomogeneous Poisson Rectangular Pulse Process

A stochastic model to estimate residential indoor water demands in a water distribution system was presented by Buchberger and Wu (1995). The approach used in the model has its premise in queuing theory. Using that analogy, customers are replaced with home occupants, and servers are represented by water fixtures and appliances. The arrival of customers (or frequency of water consumption) is approximated as following a nonhomogeneous Poisson process with a time dependent rate parameter. When servers are busy, the water demands occur as rectangular pulses, with each pulse having a random intensity and random duration. A single home often has 10 or more servers; however, in this model, all the servers are joined in one group. The water intensities and durations are described by a common probability distribution. Although this may decrease the resolution of the model, it also reduces the number of parameters required in the modelling process. The three parameters used are the average demand at a busy server; the variance of the demand at the busy server, and the time dependent utilisation factor for a typical single family household.

The validity of the nonhomogeneous Poisson rectangular pulse (PRP) process was tested by Buchberger and Wells (1996). Water flow was recorded at 1-second intervals for one year in four single family households, although only two residences were used to present the findings. Flow signals were processed so that each event was converted to an equivalent discrete rectangular pulse, which proved to be a satisfactory representation of water demand. At both residences the variance of the observed daily pulse count was too high to be modelled by a Poisson process. The study concluded that, although the results do not invalidate the model, it does require further refinement and investigation.

The PRP hypothesis was later verified by Buchberger and Schade (1997), using 30 days of water recordings in 18 single family households with 1-second intervals. The probability distribution of busy servers given by the PRP model showed a good fit to the hourly variation of observed values.

As part of a broader study to investigate water quality in dead end zones, Buchberger et al. (2003) tested the hypotheses that residential water demand is a time dependent Poisson process. The analysis was based on recordings of 21 homes in the city of Milford, Ohio, which were logged at 1-second frequency for 31 consecutive days from May 11 to June 10, 1997. Comparing the model predictions and observed values, the authors found good agreement to the number of busy homes on an hourly basis, server transitions, and busy server autocorrelation functions. Predicted mean flow rates in pipes also showed a good fit to observed values. Flow variances exhibited some discrepancies between predicted and observed values. Buchberger et al. (2003) further note that the PRP process should be applied separately to indoor demand and outdoor demand in order to estimate total demand. By taking the sum of coincident pulses during the peak time of the day, the maximum flow for that day may be obtained.

2.3.5. SIMDEUM

The PRP model, according to Blokker et al. (2010), is more of a descriptive model than a predictive one. Since the parameters of the model are derived from measurement results, and correlations to other data such as population size or installed appliances are not easily done, the PRP model does not lend itself to transferral to other networks. In an attempt to reduce the need for large logging projects, Blokker et al. (2010) developed a water demand model called Simulation of water Demand, an End-Use Model (SIMDEUM). The model is based on statistical information to simulate residential water demand patterns.

Similarly to the PRP model, SIMDEUM assumes that water demand occurs as rectangular pulses. However, the arrival time of the pulses over the day, the intensity and duration, are described by probability distributions for each end-use. The probability distribution parameters are obtained from surveys providing statistical information, not from measurements. The survey that was used to validate SIMDEUM was conducted in 2001 by Dutch water companies. About

3200 respondents answered questions on their household and fixtures, and filled in a diary for a week on their water consumption.

Blokker et al. (2010) incorporated eight end-uses in SIMDEUM. These were the toilet, shower, washing machine, dishwasher, kitchen tap, bathroom tap, bath and outside tap. Each end-use was assigned a penetration rate (number of households owning a specific type of appliance). Various subtypes constituting an end-use were also defined.

Household size, age, gender, and occupation were used to divide the users into groups. These groupings were related to the frequency of use, duration, and time of use (based on specific users' diurnal pattern) for each end-use. The diurnal patterns were constructed by assuming water demand is strongly related to when people are at home, awake and available to use water. Information on the availability of people was obtained from a time-budget survey.

Water demand was described by the following equations:

$$Q = \sum_{k^*=1}^{M^*} \sum_{j^*=1}^{N^*} \sum_{i^*=1}^{F^*{}_{j^*k^*}} B\left(I_{i^*j^*k^*}, D^*{}_{i^*j^*k^*}, \tau_{i^*j^*k^*}\right) \quad (2.5)$$

$$B\left(I_{i^*j^*k^*}, D^*{}_{i^*j^*k^*}, \tau_{i^*j^*k^*}\right) = \begin{Bmatrix} I_{i^*j^*k^*} & \tau_{i^*j^*k^*} < T < \tau_{i^*j^*k^*} + D^*{}_{i^*j^*k^*} \\ 0 & elsewhere \end{Bmatrix} \quad (2.6)$$

where:

k^* = all end-uses from 1 to M^*

j^* = all users from 1 to N^*

i^* = all busy times per end-use from 1 to $F_{j^*k^*}$

F^* = frequency of use

D^* = pulse duration

I = pulse intensity

τ = time at which the tap is opened

T = time parameter

$B(I, D^*, \tau)$ = block function.

Once the statistical information was put into the model, a single simulation represented a possible outcome for a single household on one day. A Monte Carlo simulation provided results for repeated simulation. The results of the model showed good agreement with measured water consumption data. According to Blokker et al. (2010), if the required statistical information were available, then the model could be applied to water networks at different locations.

2.4. Peak Water Demand Estimation Methodologies

2.4.1. Fixture Value Approach

The fixture value approach is a method used in North America to estimate peak flow. It is used in pipes, for sizing what is termed service lines. It entails calculating the potential peak demand by determining the probability that various water consuming fixtures are used simultaneously. The probability patterns of fixture use can be derived empirically as described by AWWA (2004).

A “fixture unit” method for estimating peak demand was developed by Roy Hunter in 1940. He produced the Hunter curve, which relates peak flow to the number of fixtures. Since Hunter used his own judgement regarding the probability function of fixtures flowing, it is necessary to use engineering judgement when applying the method (AWWA, 2004). The Hunter curve is based on a high probability that many of the fixtures are used at the same time. This has led to an overestimation of peak demands when the Hunter curve approach is used (AWWA, 2004).

In an attempt to refine the fixture unit method, Manual M22, published by the AWWA in 1975, incorporated demand curves derived from field measurements. A limitation of the new curves, according to AWWA, (2004), is that they were constructed on measurements from a small sample of customers in the United States and Canada. The 1975 M22 curves give much lower peak estimations than the Hunter curve. However, when applied to two case studies, the 1975 M22 curves provided a better representation of actual peak demand than the Hunter curves (AWWA, 2004).

AWWA (2004) recommends that the peak demand used in the engineering design of service lines is calculated using a modified fixture value approach, which is based on the method given in Manual M22. A fixture value is an estimate of peak instantaneous flow of a single fixture at a particular pressure. These are used, together with measured data, to develop probability curves.

The general procedure of the modified fixture value method entails first calculating the total fixture value. This is done by multiplying the fixture values of specific appliances by the number of appliances in use, and taking the sum thereof. The probable demand, corresponding to the combined fixture value, is determined from the probability curves. The fixture values were determined at a pressure of 413.7 kPa. Pressure adjustment factors are provided for appliances operating at alternative pressures. Probable demand is multiplied by the pressure adjustment factor to obtain total probable demand.

The method determines peak flow for irrigation (outdoor) demand and residential (indoor) demand separately. In cases where irrigation and residential demands occur at different times, the larger of the two is selected and, when they occur simultaneously, the sum is used. Furthermore, the demand may need to be increased in cases where fixture usage is uncertain, and continuous demands should be added to peak residential use.

2.4.2. Peak Demand Diversity Relation

It is possible to size pipes using fixed peak per capita demands for each residential connection served by that pipe. Lingireddy et al. (1998) were of the opinion that it is not necessarily the best method, because overall flow requirements may be overestimated, while the requirements for individual branch lines with few connections may be underestimated. This is especially relevant to rural households. Lingireddy et al. (1998) cite a study by

Williams (1968), which suggests that the maximum flow requirement for each pipe should be calculated using the peak demand diversity (PDD) relation. The PPD relation is given in (2.7).

$$Q_{max} = \dot{a}\sqrt{\grave{N}_{\bar{C}}} + \dot{b}\grave{N}_{\bar{C}} + \dot{c} \quad \text{.................................... (2.7)}$$

where:

Q_{max} = maximum flow rate

$\grave{N}_{\bar{C}}$ = number of residential connections served by the pipe

$\dot{a}$, $\dot{b}$, and $\dot{c}$ are constants obtained from field data.

The PPD relation takes into account that pressure drops during the delivery of instantaneous peak flow are dependent on the number of connections served by the pipe section. The higher the number of connections, the lower the peak flow requirement per connection. This is because the probability of all the users on the pipe section requiring maximum capacity simultaneously decreases. The opposite is true for pipes with a low number of connections. It was concluded by Lingireddy et al. (1998) that pipes should be sized using the PPD flow requirements for systems that are designed without incorporating fire flow requirements.

2.5. Peak Factors

2.5.1. Overview

One of the most common methods of determining peak water demand is by means of peak-to-average ratios, also known as peak factors, peak coefficients, or demand multipliers. The baseline, or average, demand is often represented by the AADD, and the baseline demand for PFs mentioned in the following sections is the AADD, unless stated otherwise. Once a baseline demand is obtained, peak flow is computed by multiplying the baseline demand by a PF, as shown by (2.8).

$$(Q)_{max} = (Q)_{avg} \times PF \quad \ldots \quad (2.8)$$

where:

Q_{max} = maximum flow rate (m^3/s, or any other unit of flow rate)

Q_{avg} = average flow rate (m^3/s, or the same unit of flow rate as for Q_{max})

PF = peak factor (dimensionless).

Hence, the PF is given by:

$$PF = Q_{max}/Q_{avg} \quad \ldots \quad (2.9)$$

2.5.2. Time Interval for Calculation of Peak Factor

PFs are highly related to the duration of peak flow of a WDS (Diao et al., 2010). The peak factor increases as the time interval over which flow is measured decreases (Johnson, 1999). This is because the average flow rate over the time interval is taken. Therefore PF_{inst} would be larger than PF_h, which is larger than PF_d. These PFs are, in turn, based on Q_{inst}, Q_h, and Q_d.

2.5.3. Number of Consumers

Diao et al. (2010) also state that the number of consumers has an impact on the PF. As the number of consumers increases, the magnitude of the peak factor decreases. This can be illustrated with the following example:

If over 24 hours there is a short moment when 100 people cause a flow of $(Q_{inst})_{max} = 0.104\ \ell/s$, and an average flow of $(Q_{inst})_{avg} = 0.013\ \ell/s$ is observed for the day, then the $PF_{inst} = 0.104/0.013 = 8$. This implies that, at one moment, water demand was approximately eight times greater than the average.

When more consumers are considered, say 10 000 people, water will not be used at exactly the same time by each of the consumers. The $(Q_{inst})_{max}$ may in such a case may be $0.600\ \ell/s$, and the $(Q_{inst})_{avg} = 0.200\ \ell/s$, resulting in

$PF_{inst} = 0.600/0.200 = 3$. Due to this clear relationship between PF and number of consumers, population size is often used as the independent variable in PF computations. In some cases the AADD is used as a surrogate for population size (Vorster et al., 1995). The CSIR (2003) used ee, which is related directly to AADD, as surrogate for population size.

According to Diao et al. (2010), service areas have a large effect on PFs. Since flow characteristics differ between locations, PFs are often determined from a field study of a particular area (Trifunović, 2006). Varying regions also have their own PF calculation methodologies, which include empirical equations in some cases.

2.5.4. Internationally Derived Peak Factors

Zhang (2005) referred to three examples of empirical PF and peak flow equations obtained from various US publications, namely the Central Iowa Committee (2004), Georgia minimum standards for public water systems (2000), and US bureau of reclamation design criteria (2002). The expressions from the above mentioned publications are listed in the same order by (2.10) to (2.12).

$$PF = \begin{Bmatrix} 9.0\ E \leq 0.22 \\ \frac{7}{E^{0.167}}\ E > 0.22 \end{Bmatrix} \qquad (2.10)$$

where:

E = population in thousands.

$$Q_{inst} = 43.40C^{0.54} \qquad (2.11)$$

where:

Q_{inst} = instantaneous flow (ℓ/min)

C = number of connections (for $C < 500$).

$$Q_{inst} = 18.19\eta^{0.5} + 3.41\eta + 22.36 \qquad (2.12)$$

where:

η = number of houses.

Diao et al. (2010) presented some German relations for peak factors that were derived by the German Technical and Scientific Association for Gas and Water (DVGW). These are presented by (2.13) to (2.16):

DVGW - Worksheet W 400-1 (2004):

$$PF_d = -0.1591 \cdot lnE + 3.5488 \quad (2.13)$$

$$\mathrm{PF}_h = -0.75 \cdot lnE + 11.679 \quad (2.14)$$

DVGW - Worksheet W410 (2007):

$$PF_d = 3.9 \cdot \mathrm{E}^{-0.0752} \quad (2.15)$$

$$PF_h = 18.1 \cdot \mathrm{E}^{-0.1682} \quad (2.16)$$

where:

PF_d = peak day factor

PF_h = peak hour factor

E = population.

Diao et al. (2010) further presented hourly PFs according to Australian design codes WCWA (1986) and WSAA (1999), as shown in Figure 2.7.

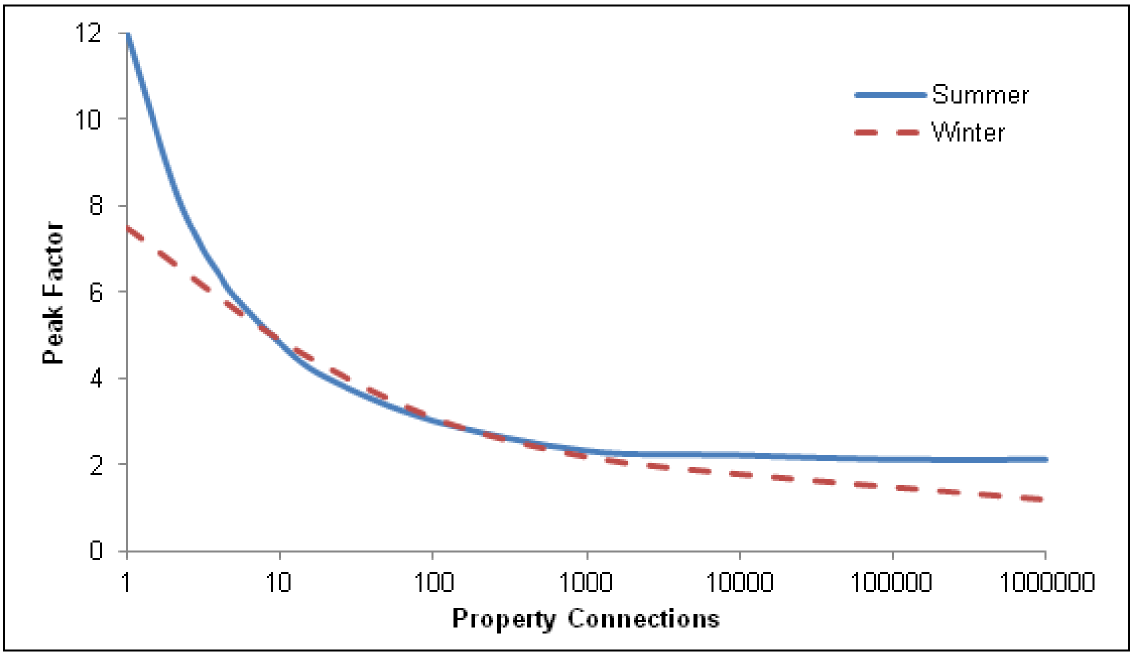


Figure 2.7: Australian PF_h estimation curve adapted from Diao et al. (2010)

Brière (2007) discussed two different PF methodologies. He stated that the Goodrich empirical formula is applicable to small residential municipalities to calculate the peak factor as a percentage.

The Goodrich formula is written as:

$$p' = 180t^{-0.10} \qquad (2.17)$$

where:

p' = maximum percentage (%)

t = period studied (days).

The formula is only applicable for t between 12 hours and 365 days. As an example of how the Goodrich formula is applied, if $t = 1d$, then $p' = 180(1)^{-0.10} = 180\%$, resulting in $PF_d = 1.80$.

Alternatively, Brière (2007) referred to PFs based on population size, as given by the Ontario Environment Ministry Guidelines for the Design of Water Storage Facilities, Water Distribution Systems, Sanitary Sewage Systems and Storm Sewers (May 1979), as presented Table 2.5.

Table 2.5: Peak factor for total water-consumption flow rates (Brière, 2007)

Population	PF_d	PF_h
Under 500	3	4.5
500 to 1 000	2.75	4.13
1 001 to 2 000	2.5	3.75
2 001 to 3 000	2.25	3.38
3 001 to 10 000	2	3
10 001 to 25 000	1.9	2.85
25 001 to 50 000	1.8	2.7
50 001 to 75 000	1.75	2.62
75 001 to 150 000	1.65	2.48
0ver 150 000	1.5	2.25

In Spain, Martinez-Solano et al. (2008) evaluated the PF by means of an expression that was obtained through statistical analysis of water consumption, as shown in (2.18):

$$\mathrm{PF} = \frac{17.12}{\sqrt{\widetilde{N}}} + 2.185 \quad (2.18)$$

where:

$\widetilde{N}$ = number of consumers.

The water consumption of a small town in Southern Italy was analysed by Tricarico et al. (2007). A statistical analysis was done on the sample, and the study showed that flow could stochastically be described by log-normal and Gumbel models. Using a deterministic approach, the authors developed a relationship to estimate maximum flow in relation to the number of users. The resultant equation was:

$$\mathrm{PF} = 11\widetilde{N}^{-0.2} \quad (2.19)$$

where:

$\widetilde{N}$ = number of consumers.

Tricarico et al. (2007) then studied the data using a probabilistic approach. This involved calculating PFs with confidence intervals of 90%, 95%, 98% and 99%. The PFs that were obtained, using both the deterministic and probabilistic approach, are given in Table 2.6.

Table 2.6: Peak factors (Tricarico et al., 2007)

Number of inhabitants		**100**	**250**	**750**	**1000**	**1250**
PFs resulting from deterministic approach		4.4	3.7	2.9	2.8	2.6
PFs resulting from probabilistic approach	90%	3.8	2.9	2.3	2.1	2.0
	95%	3.9	3.0	2.3	2.1	2.0
	98%	4.1	3.1	2.3	2.2	2.1
	99%	4.2	3.1	2.3	2.2	2.1

Zhang et al. (2005) developed a reliability based estimate of the PF by combining the results of the NRP model by Buchberger and Wells (1996) with principles from extreme value analysis. The expression had the form given in (2.20).

$$PF(N|p) = \Psi^* \left(1 + \xi_{\grave{\rho}} \sqrt{\frac{1 + \Theta_q^2}{\Psi^* \ddot{\rho} \ddot{N}}} \right) \quad (2.20)$$

Where:

$\Psi^* = \lambda^*/\lambda$	=	dimensionless peak hourly demand factor
λ	=	mean arrival rate of water demands at a single family household
λ^*	=	arrival rate during the period of high water use
$\xi_{\grave{\rho}}$	=	$\grave{\rho}^{\text{th}}$ percentile of the Gumbel distribution
$\ddot{\rho} = \lambda\breve{\tau}$	=	daily average utilization factor for a single family household
Θ_q	=	coefficient of variation of PRP indoor water demand pulse
$\ddot{N}$	=	number of homes in the neighbourhood
$\grave{\rho}$	=	percentile.

After calculating the PFs for a number of population sizes with a 99th percentile using (2.20), the expression proved to follow a similar trend to other empirical equations. Zhang et al. (2005) noted that most of the PFs calculated empirically by other authors are greater than the 99th percentile results using (2.20), which implies that the conventional methods of estimating PFs are conservative.

Hyun et al. (2006) applied four different methods of determining peak factors and evaluated the effect the methods had on the design capacity of pipelines. A summary of the methods used, as well as the resultant effects, is presented in Figure 2.8.

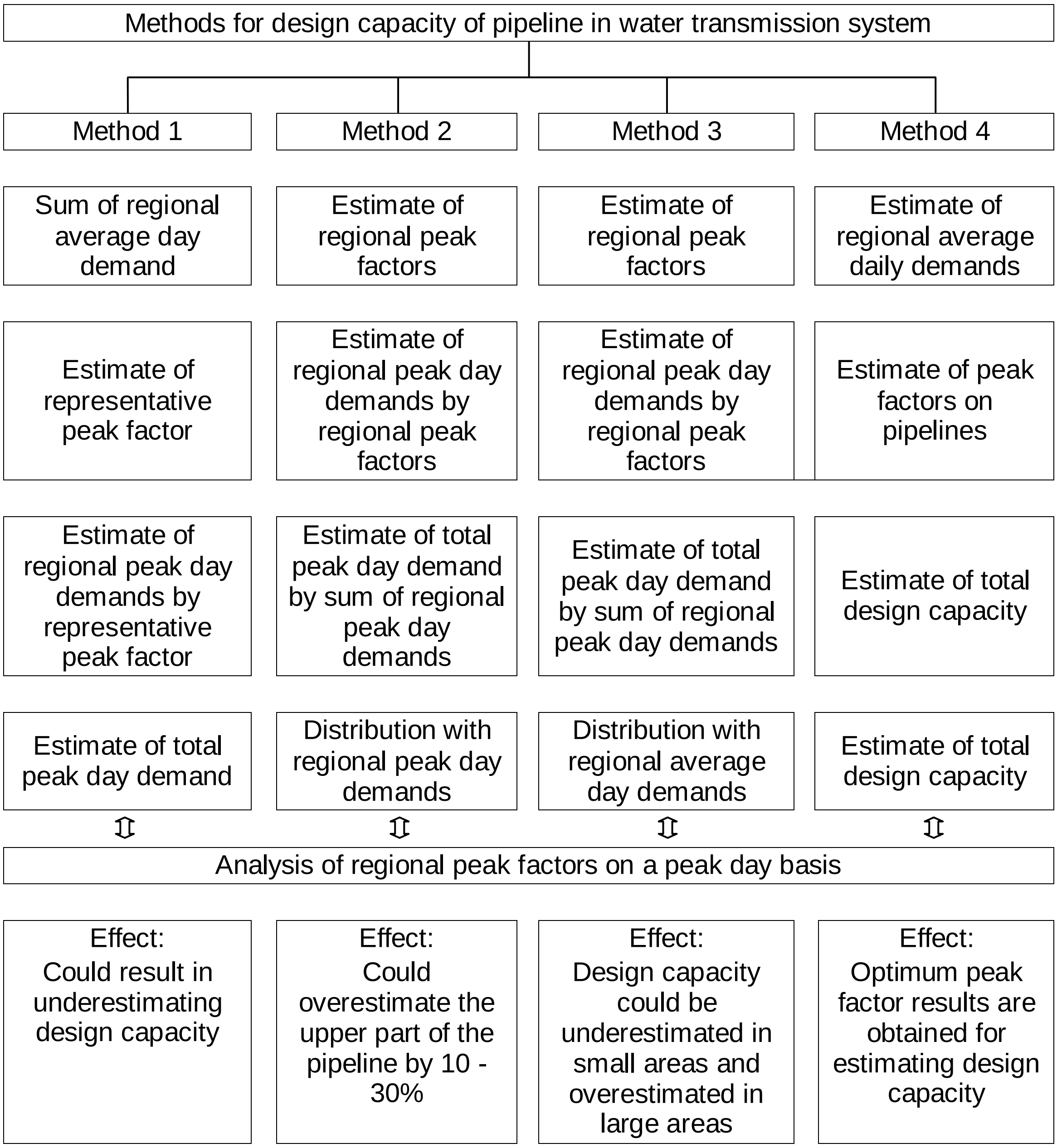


Figure 2.8: Pipeline design capacity calculation methods (Hyun et al., 2006)

2.5.5. South African Peak Factors

PFs are the preferred method of calculating peak flow in South Africa. In some cases consultants developed their own in-house PFs. Two consulting engineering firms compiled a master plan for an East Rand WDS in South Africa, and an overview was subsequently published (Vorster et al., 1995). The residential PFs used by the authors as part of the study are given in Table 2.7.

Table 2.7: Peak factors to be applied to AADD (Vorster et al., 1995)

Predominant land use in area under consideration	AADD for area $(M\ell/d)$	PF_d	PF_h
Low density residential	<1.0 1.0-5.0 5.0-20.0 >20.0	2.30 2.20 2.00 1.80	5.50 4.50 3.90 3.30
Medium density residential	<1.0 1.0-5.0 5.0-20.0 >20.0	2.30 2.00 1.80 1.70	4.60 4.00 3.30 2.90

A design guideline with the PFs used by many engineering practitioners in South Africa was published by the CSIR in various formats between 1983 (CSIR, 1983) and 2003 (CSIR, 2003). Figure 2.9 presents the PF diagram taken from the CSIR (2003), and is commonly used to determine the PFs for developed areas. To obtain the PF, the type of development has to be converted into ee (where 1 ee = $1\ \mathrm{k}\ell/\mathrm{d}$). The instantaneous peak flow is then calculated by multiplying the PF by the AADD. No definition is given by the CSIR (2003) of the time interval that is used to represent instantaneous demand. It thus remains unclear wether this derived peak flow would be the maximum as averaged over a second, a minute, an hour and so on.

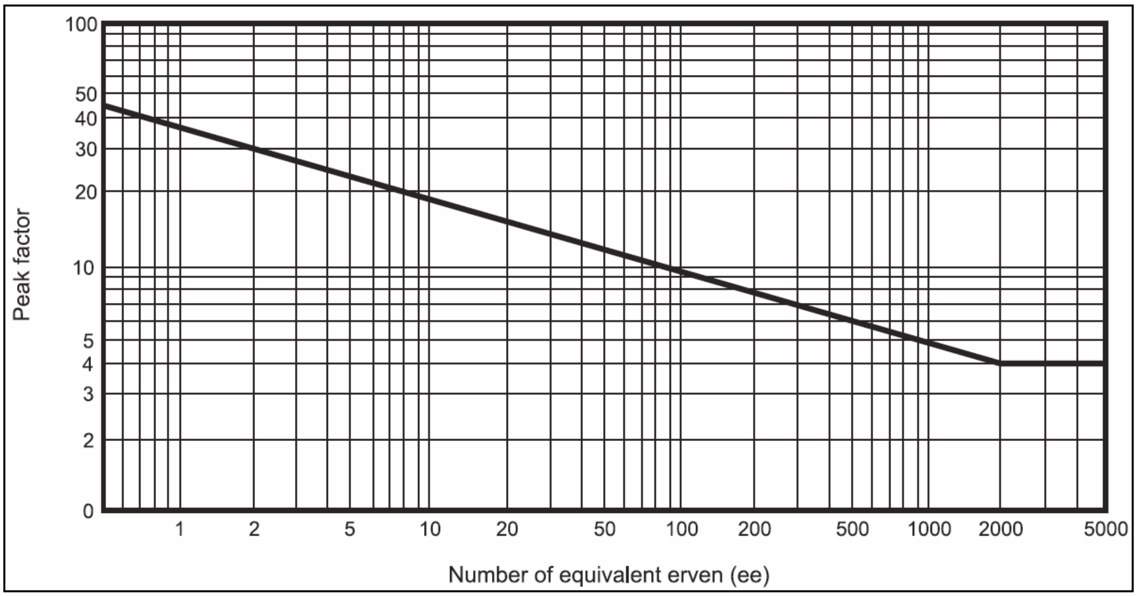


Figure 2.9: Peak factor estimation curve for developed areas (CSIR, 2003)

The figure presented by the CSIR (2003) remains identical to that presented in the original version of the publication (CSIR, 1983). Booyens (2000) mentioned that the data used in Figure 2.9 was obtained from questionnaires completed by designers and consultants in order to gather PFs used in practice at the time of compiling the initial document (CSIR, 1983).

Since 1983, several authors making use of water meters and electronic data loggers have contested the validity of Figure 2.9, stating that the PFs given by CSIR (2003) were too conservative. The findings of some of these studies are briefly discussed below.

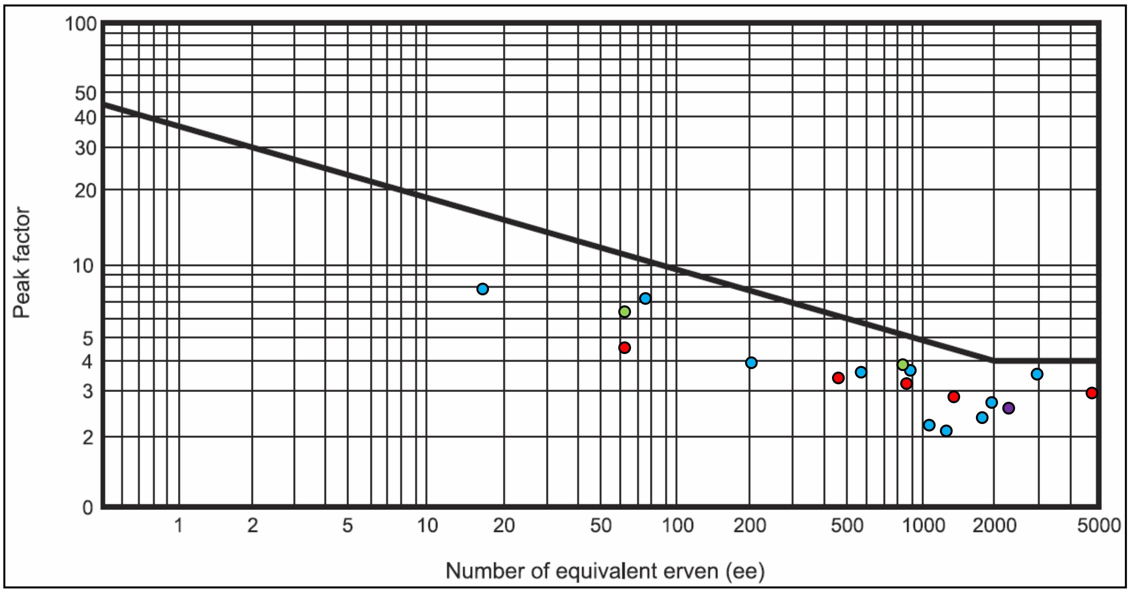
Hare (1989) isolated three different residential areas in Port Elizabeth, South Africa. A single meter and data logger monitored each of the respective areas. Flow was logged at 10 minute intervals during the summer of 1987 - 1988. Several problems were experienced, which the author lists as battery failure, blocked water meters, lack of accuracy in meters, and late delivery of equipment. The logging of only two of the areas resulted in decent results which could be plotted on the peak factor curve, as shown in Figure 2.10. Hare (1989) deemed the project incomplete and therefore inconclusive.

Water consumption data for the city of Pretoria, South Africa, was analysed by van Vuuren and van Beek (1997). The measuring period ranged between 1982 and 1994, for a total of 151 months. Residential water consumption data was isolated for analyses purposes. Problems included the fact that the data provided only an estimation of the actual water consumption per month, because the monthly readings did not correspond to calendar months, or 30 day periods. To determine peak consumption, hourly data was required, which could be obtained only from the Carinastraat reservoir. The hourly data was for the period between July 1995 and October 1995. A peak factor of approximately 2.75 was observed. This is lower than the PF of 4.0 recommended by the CSIR (1983), as shown in Figure 2.10.

Turner et al. (1997) measured the water consumption in 15 minute intervals for 14 areas in Gauteng over a 20 month period. A distinction was made between flow patterns on weekdays, Saturdays and Sundays. The average peak factor was used to plot probability intervals of +99%, +95%, -95%, and -99%. The probability intervals were determined by calculating the variance and standard deviation for each time interval. The 15 minute peak factor was plotted against the equivalent erven. The results for 13 areas are shown in Figure 2.10. Each of the measured peak factors was lower than the CSIR (1983) guideline. The authors, however, proposed that instead of following the suggested peak factors precisely, a range of peak factors should be considered. This would allow a utility to choose the acceptable reliability.

Peak factors were calculated from measured results in South Africa by Booyens (2000). Three data loggers and two telemetry systems were used to record the flow for 5 zones in the Boksburg municipality. The zones contained approximately 3094 stands (4585 ee); 863 stands (1352 ee); 794 stands (828 ee); 444 stands (446 ee); and 69 stands (62 ee), respectively. The study area proved to have a homogeneous water demand pattern, and comprised of mainly residential properties. Booyens (2000) used the data to calculate PFs using different time intervals, and used a 15 minute time interval to calculate probabilistic peak factors, as well as different return periods. A comparison with the CSIR (2003) guideline is presented in Figure 2.10.

Johnson (1999) noted that it is important to associate the PF with the time intervals at which peak flows are measured because, as the time interval duration increases, the PF decreases. In the case of all the abovementioned comparative studies, the PFs were calculated using longer time intervals than the instantaneous PF that the CSIR (2003) is claiming to depict. It is, therefore, reasonable to expect that the larger time interval PFs exhibit lower PF magnitudes.



- Hare, (1989) - PF_{10min}
- Turner et al., (1997) - PF_{15min}
- van Vuuren and van Beek (1997) - PF_{1h}
- Booyens (2000) - PF_{15min}
- CSIR (2003) - PF_{inst}

Figure 2.10: South African measured peak factors comparison

When peak factors are high, then essentially it means that large pipeline capacities are maintained to be used only for short intervals of peak flow, lowering the degree of utilisation of the pipelines. Johnson (1999) defines the degree of utilisation as the reciprocal of the PF:

$$\begin{matrix} Degree\ of\ utilisation \\ (specifying\ the\ period) \end{matrix} = \frac{100}{peak\ factor\ or\ function}\ (per\ cent)\ \ldots\ldots\ldots\ldots \qquad (2.21)$$

Johnson (1999) used probability theory to determine the recurrence interval of peak events and the related degree of utilisation. The maximum 15 minute flow and average daily flow from a reservoir over a period of 120 months was used to calculate 15 minute peak factors. The degree of utilisation probability graph derived by Johnson (1999) is shown in Figure 2.11.

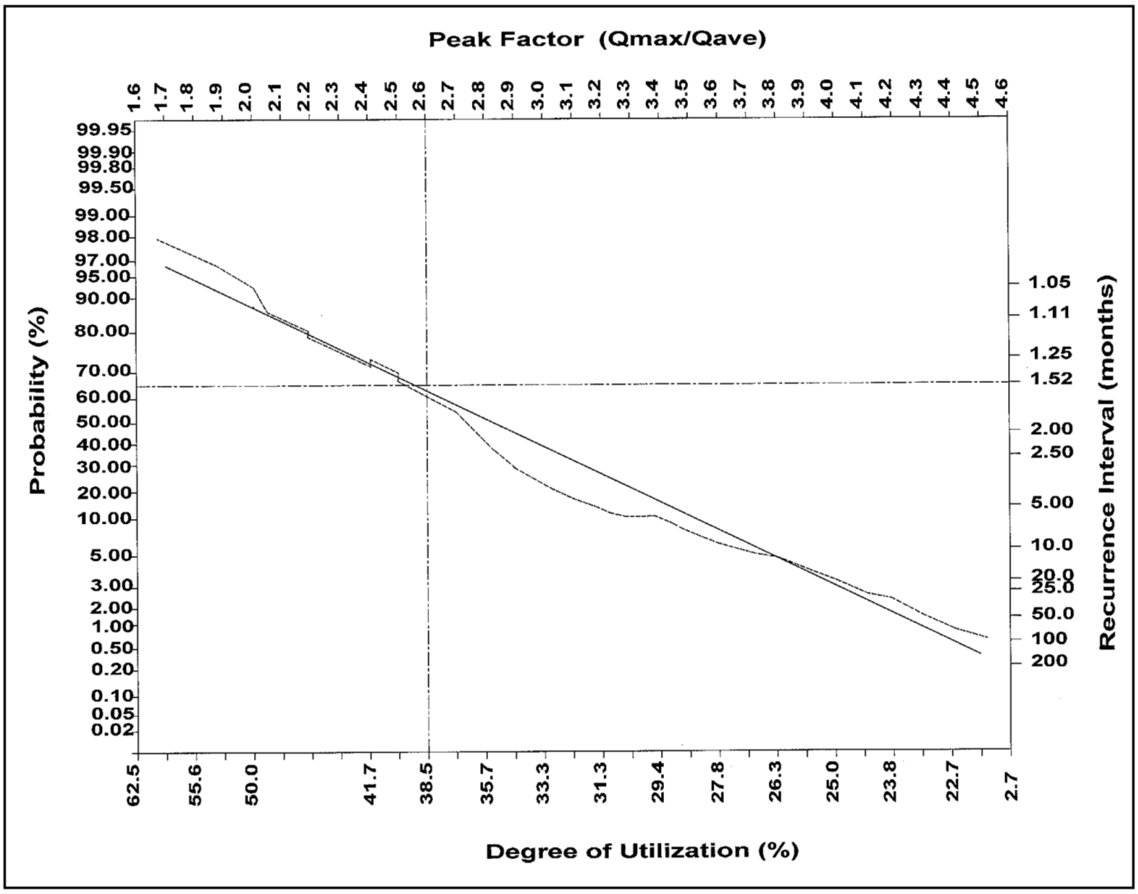


Figure 2.11: Degree of utilisation - probability graph (Johnson, 1999)

It was demonstrated by van Zyl (1996) how peak factors could be determined for supply areas that have different sizes, but similar characteristics, by analysing a typical demand pattern. According to van Zyl (1996) an average diurnal flow pattern should be based on water consumption records over a long time. The water consumption at a particular time of the day, divided by the average water consumption for the entire day, results in a flow pattern which is represented by peak factors. An assumed flow pattern is illustrated by van Zyl (1996) in Figure 2.12.

It was pointed out by van Zyl (1996) that residential water consumption is not continuous, but is due to a combination of discrete water withdrawals from utilities (end-uses). A water demand pattern is, therefore, indicative of how many end-uses are active (open) at any given moment.

According to van Zyl (1996) a residential water demand pattern could be interpreted as a probability function relating the times that end-uses are active throughout the day. The probability distribution of active end-uses corresponding to the assumed flow pattern is illustrated in Figure 2.13, where the total area under the graph is equal to unity.

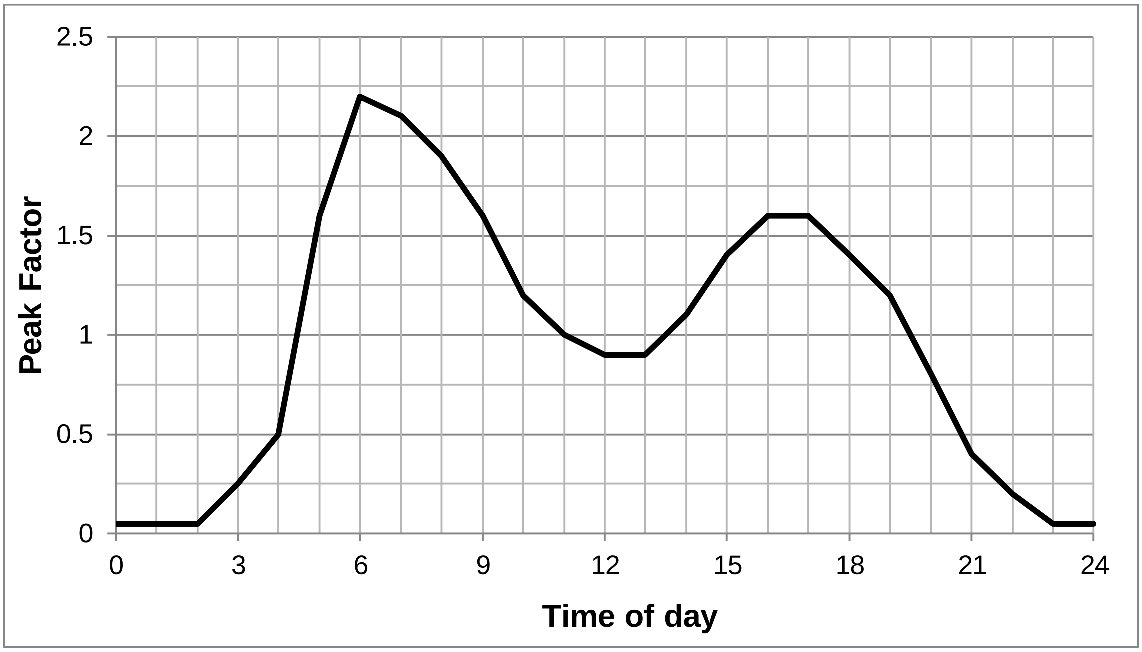


Figure 2.12: Assumed diurnal flow pattern adapted from van Zyl (1996)

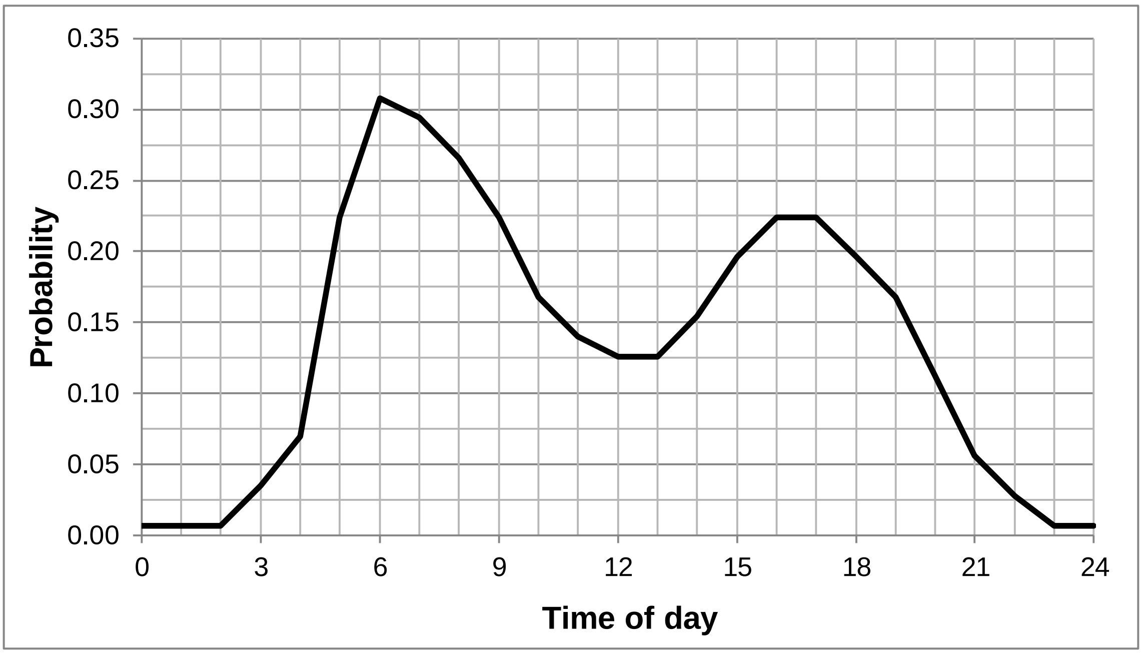


Figure 2.13: Probability pattern of active end-uses adapted from van Zyl (1996)

A computer program was developed by van Zyl (1996) that simulated active end-uses throughout the day by making use of a probability pattern such as the one as shown in Figure 2.13. For each simulation, the computer program assigned a random number between zero and one for every minute of a day. At each minute, the random number was compared to the probability value of the probability pattern. If the random number was less than or equal to the probability value, then an end-use was considered active. If the random number was greater than the probability value, an end-use was considered inactive. It was assumed that a constant flow rate occurred each time an end-use was active. A single simulation represented the end-use activity resulting from a single user (consumer). Simultaneous simulations, therefore, denoted the end-use activity due to more consumers.

An example of a simulation by van Zyl (1996) for a single end-use is presented in Figure 2.14, where an active end-use is represented by a solid black line. At the the times of day when an end-use was active relatively frequently, a greater density of black lines was observed. Figure 2.15 shows the results of simultaneous simulations of 10, 100, 1 000, and 10 000 users, respectively. It is clear that as the number of consumers increases, the number of active end-uses throughout the day tends to follow a similar pattern to the probability pattern in Figure 2.13.

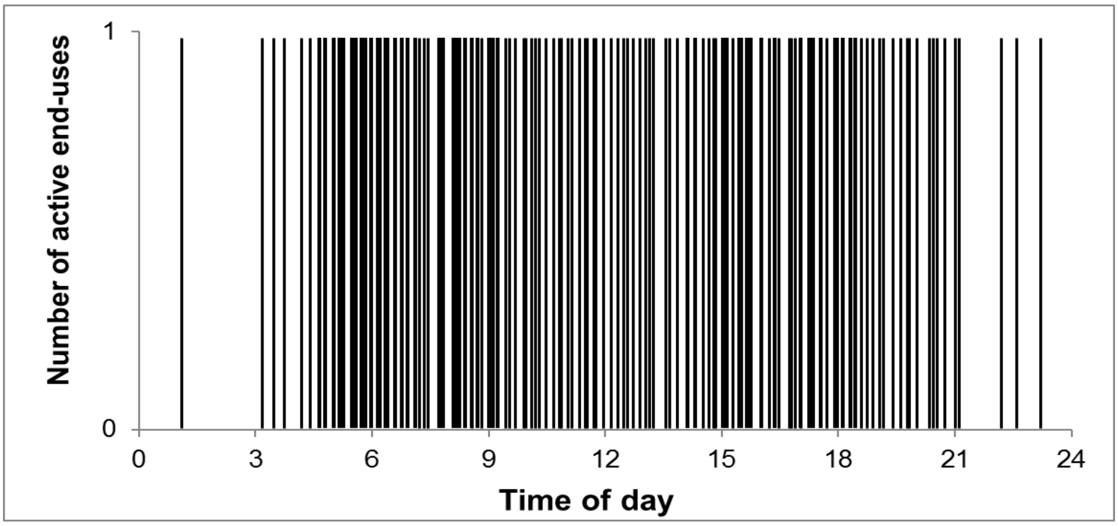


Figure 2.14: End-use activity results for 1 simulation adapted from van Zyl (1996)

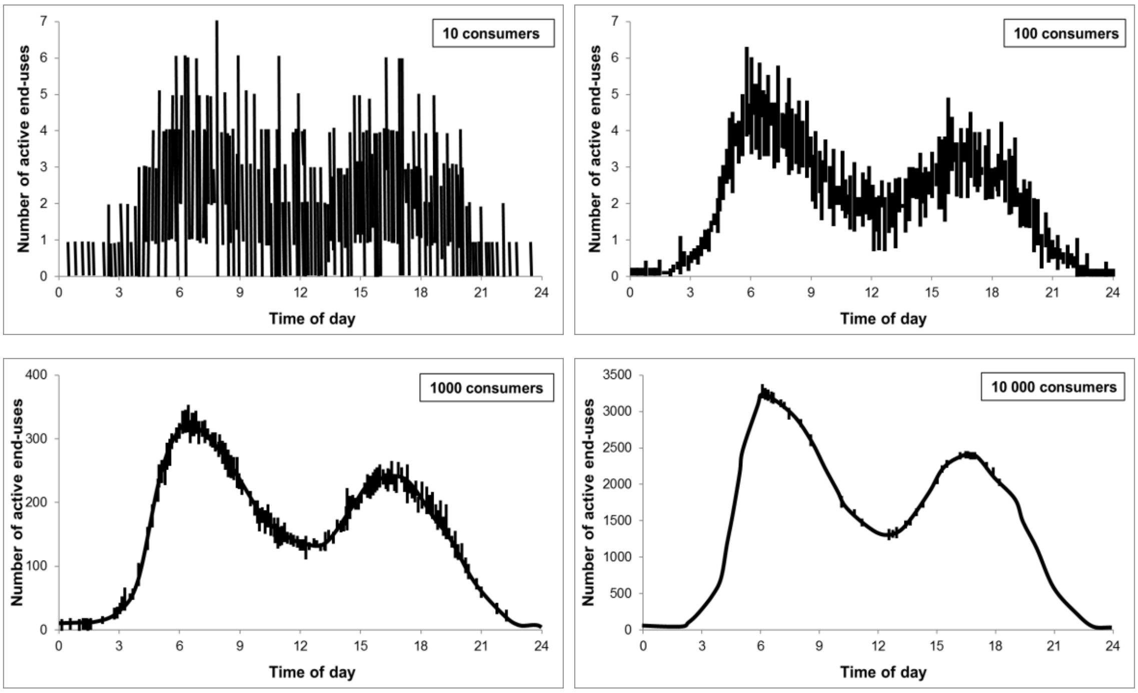


Figure 2.15: Simulation results for many consumers adapted from van Zyl (1996)

It was proposed by van Zyl (1996) that the simulations of active end-uses could be used to obtain peak factors by dividing the maximum number of active end-uses by the average number of active end-uses for the day. According to van Zyl (1996), with further investigation and calibration, the proposed method could be used to determine design peak factors for residential areas of similar type, but different size. The approach by van Zyl (1996) has not yet led to design peak factors, however, similar concepts were applied in the end-use model developed in this thesis, such as using probability patterns to establish diurnal end-use activity, and the representation of area sizes by means of simultaneous simulations.

3. STATISTICS AND PROBABILITY THEORY

3.1. Introduction

The estimation of residential water demand and peak flows contains an element of uncertainty, which is present because there are many factors that affect water demand. One of the simplest ways in which to resolve uncertainty is by substituting each uncertain quantity by its average, median, or critical value. A deterministic approach can then be used. Loucks and van Beek (2005), however, warn that when important parameters are highly variable, then replacing those uncertain quantities by the average values can affect the outcome severely. Since the factors determining water demand varies greatly from one neighbourhood to another, and even from household to household, this approach of substituting average values is not ideal.

The probability theory and stochastic processes that were used to incorporate the random factors in water demand are discussed in the following sections. Theoretical probability distribution functions were used in this study to statistically represent elements of the recorded end-use data. An overview of the theoretical probability distribution functions which best represented the respective end-use elements utilised in this study is also provided. It is important to note that although definitions of statistics used throughout the study are provided, this is not an exhaustive overview of the topic.

3.2. Random Variables

In statistics, a subset of a population is termed a sample. In a probabilistic experiment, the set of possible chance outcomes is called the sample space, denoted by S. A random variable is the function that associates a value with each outcome in in the sample space (Forbes et al., 2011).

Let X denote a random variable and x a possible value of the random variable X. A distinction can be made between discrete random variables and continuous random variables.

A discrete random variable is defined by Devore (2004) as a random variable whose values makes up a finite set of values, or can be listed in an infinite sequence. For example, if X = the number of times that a coin toss will land heads up, then X is a discrete random variable. Possible values of X in that case are $x = \{1, 2, 3, 4, \ldots\}$. It is not possible, for example, for a coin to fall 2.3 times heads up.

A random variable is said to be continuous if its set of possible numbers can be an entire interval on the number line (Devore, 2004). For example, if X = the pH of a randomly selected compound, then X is a continuous random variable because the pH can be any possible value between 0 and 14.

In this study, the following parameters were identified as being discrete random variables:

- Household size, measured in units of people per household (PPH)
- Frequency of event per day
- Number of cycles per event
- Starting hour of event.

The following parameters were identified as being continuous random variables:

- Flow rate of event
- Volume of event.

3.3. Measures of Central Tendency

Central tendency is defined by Gravetter and Wallnau (2000) as a statistical measure that best represents the entire distribution by a single value. No method will produce a representative value for a distribution in every situation. Three methods that are often used are the mean, the median and the mode.

3.3.1. The Mean

The mean is also known as the arithmetic average of the set. It is computed by taking the sum of the values in the distribution, and dividing it by the number of individual values. If $\bar{X}$ is the mean, ΣX is the sum of all the values of X, and N is the number of X values, then the formula for calculating the mean is given as:

$$\bar{X} = \frac{\Sigma X}{N} \quad \text{(3.1)}$$

A number of properties of the mean can be noted. Firstly, the mean will change if any single value changes. Its value is, consequently, very sensitive to outliers. Moreover, it minimises the sum of squared deviations around it. Advantages of the mean, as explained by Howell (2002), are that it can be manipulated algebraically, and an estimation of the population mean is generally better achieved by the sample mean than by either the median or mode.

3.3.2. The Median

The median is the value that divides the distribution exactly in half when the data is ranked in numerical order. It is also equal to the 50th percentile (Gravetter and Wallnau, 2000). The goal of the median is to identify the precise midpoint of a distribution. The method of computation depends on whether the sample has an even or odd number of observations. An advantage of the median is that it is unaffected by extreme outliers (Howell, 2002).

According to Devore (2004), if N is odd, the median can be calculated from the formula:

$$Median = \left(\frac{N+1}{2}\right)^{th} ordered\ value \quad \text{(3.2)}$$

whereas if N is even, the median can be calculated from the formula:

$$Median = average\ of\ \left(\frac{N}{2}\right)^{th} and\ \left(\frac{N}{2}+1\right)^{th} ordered\ values \quad \text{(3.3)}$$

3.3.3. The Mode

The mode is the value in a distribution that occurs with the greatest frequency. Gravetter and Wallnau (2000) consider the mode useful because it can be used for any scale of measurement. Furthermore, it often provides the most sensible measure of central tendency, because it is the most typical case of the sample. It is also possible to have more than one mode. A distribution with two modes is called bimodal, and a distribution with more than two modes is called multimodal.

3.4. Measures of Variability

Measures of central tendency give only partial information about a distribution. There may be cases where two samples have the same mean and median, but the individual values of the one sample are spread further from one another than the other sample. It is therefore valuable to determine the variability within samples.

3.4.1. The Range

The range is a measure of distance. It is defined as the difference between the highest and the lowest values in a distribution (Devore, 2004). The range serves as an obvious way to describe the spread of the data. A disadvantage of the range as a means to describe variability is that takes into account only the two extreme values.

3.4.2. Percentiles

A percentile divides the distribution into hundredths, in terms of the number of samples. A percentile is the position in a distribution below which the specified percentage of X is situated. For example, if 90% of the observations lie below a certain value, then that value is the 90^{th} percentile. The percentiles often employed include 25th, 50th and 75th.

3.4.3. The Variance

Measures of variability involve the deviations from the mean (Devore, 2004). The deviations from the mean are found by subtracting the mean, $\bar{X}$, from each of the observations. When the observation is larger than the mean, the deviation will be positive, and when the observation is smaller than the mean, the deviation will be negative. Gravetter and Wallnau (2000) state that, in order to get rid of the positive and negative sign, each deviation is squared.

If all of the deviations are small, then the observations are close to the mean and there is little variability. If, however, many of the deviations are large, then the observations are far from the mean and the variability is great. The variance can then be computed by calculating the mean of the squared deviations.

According Devore (2004), the sample variance (σ^2) is given as:

$$\sigma^2 = \frac{\Sigma(X_i - \bar{X})^2}{N-1} \quad \ldots\ldots \quad (3.4)$$

3.4.4. The Standard Deviation

The square root of the variance is known as the standard deviation (σ) and is the positive square root of the variance, given as:

$$\sigma = \sqrt{\sigma^2} \quad \ldots\ldots \quad (3.5)$$

According to Howell (2002), both the variance and the standard deviation are very sensitive to extreme values.

3.5. Probability Distributions

3.5.1. Frequency Distribution

The frequency of a particular x value is the number of times that value occurs in a dataset. The relative frequency of a value is the fraction of the total number of times that a particular value occurs. When the relative frequency is multiplied by 100, a percentage is obtained. According to Ang and Tang (1984), the relative frequency could be used as a means of estimating the probability of events.

Tabulating the frequencies or relative frequencies of each x value creates a frequency distribution, which can be graphically displayed by means of a histogram. In a histogram, the height of the bars usually represents the frequency, while the width of the bars is equal to the interval size chosen for the data (Gravetter and Wallnau, 2000). When relative frequencies are used to construct a histogram, then the sum of the areas of the rectangles equals one.

For discrete random variables, the centre of the rectangle is positioned on the x value, and the width is taken as the distance between successive x values. Figure 3.1 shows a histogram of the daily starting hours for shower events using relative frequencies.

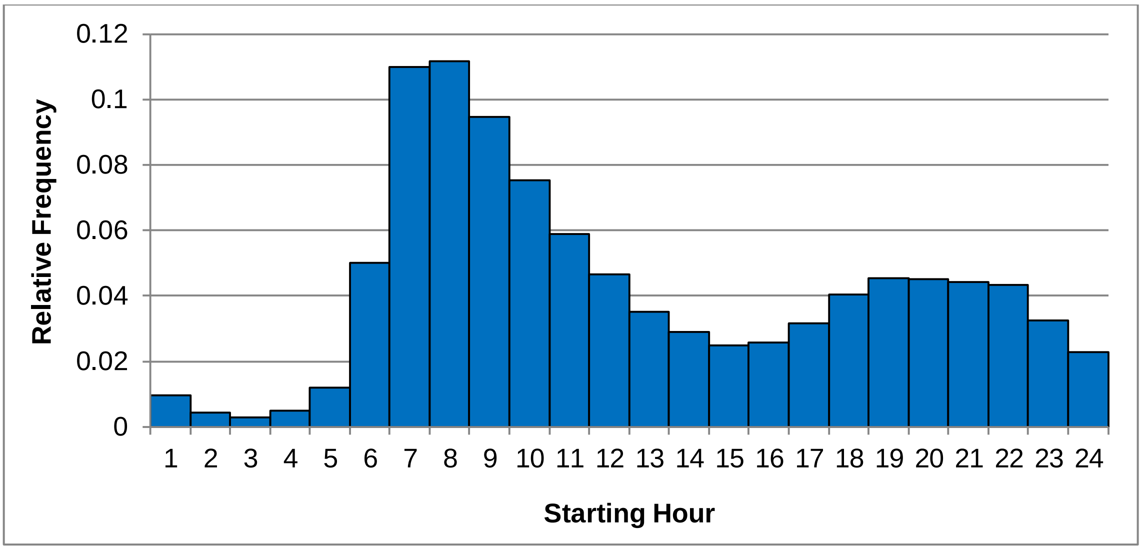
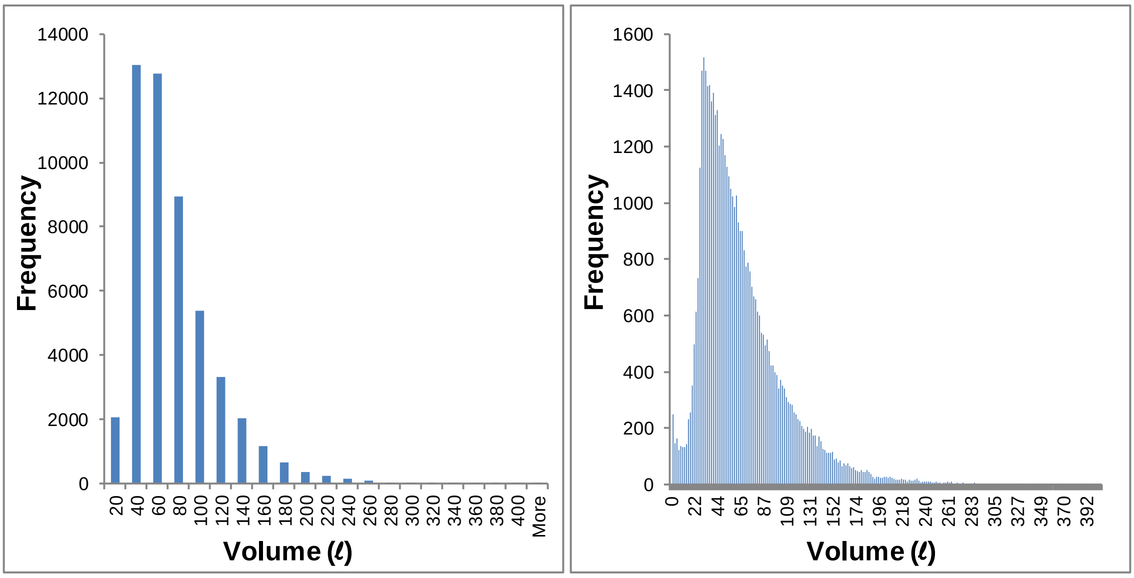


Figure 3.1: Histogram of daily starting hours for shower events

For continuous random variables, the measurement axis is divided into classes, so that each measurement is contained in only one class. As the classes are made smaller, the rectangles become narrower, until the histogram approaches a smooth curve, which is called a density curve. Figure 3.2 shows a histogram of the shower event volumes, with wide interval classes (a), and narrow intervals (b).



a) Large class intervals b) Small class intervals

Figure 3.2: Histogram interval comparison of shower event volumes

3.5.2. Probability Mass and Density Functions

When a frequency distribution is plotted, it is characterised by a certain pattern of variation. The frequency distribution can be described by a continuous mathematical function f, which is assumed to be defined over the entire real line.

The probability that a continuous random variable X takes on a value between the interval $[a, b]$ is given by the probability density function (PDF), $f(x)$, as follows:

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx \quad (3.6)$$

The graph of $f(x)$ is also referred to as the density curve. The probability that a value is between a and b is graphically represented by the area under the density curve, as shown in Figure 3.3.

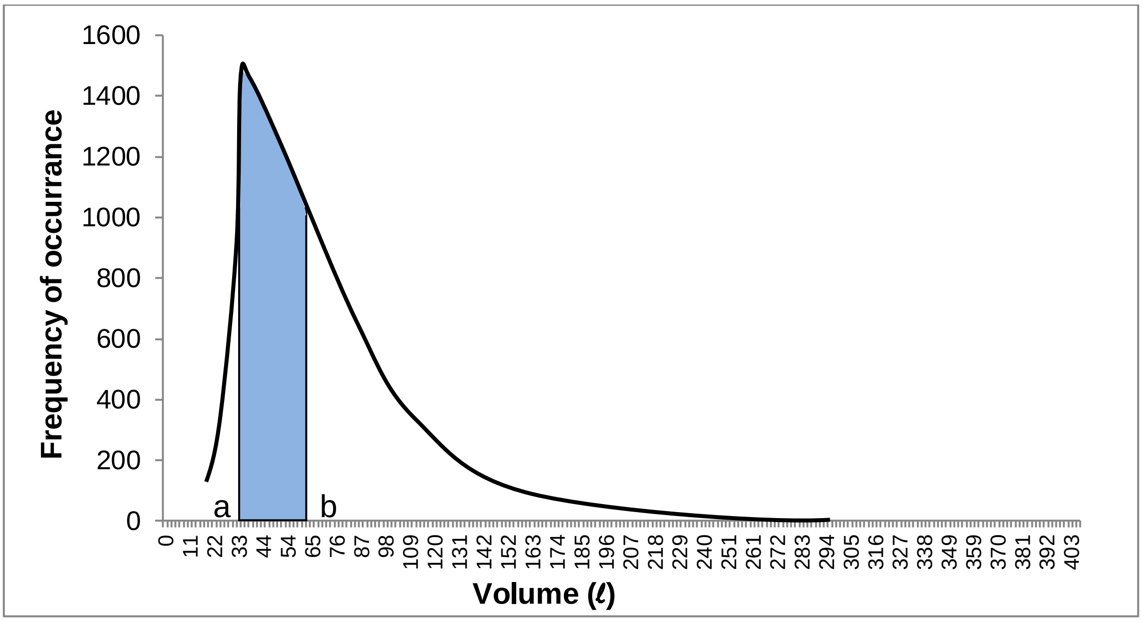


Figure 3.3: Density curve for shower event volume

Devore (2004) defines the probability mass function (PMF), $p(x)$, of a discrete random variable as follows:

For every number $x, p(x) = P(X = x)$ . . . . (3.7)

3.5.3. Cumulative Distribution Function

The cumulative distribution function (CDF) for a continuous random variable is obtained by integrating the PDF between the limits -∞ and x, and gives the probability $P(X \leq x)$. Devore (2004) defines the CDF as $F(x)$ by:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt \quad (3.8)$$

Figure 3.4 shows a PDF and the associated CDF. For any random variable X, the CDF $F(x)$ equals the probability that X is less than or equal to x (Louckes and van Beek, 2005).

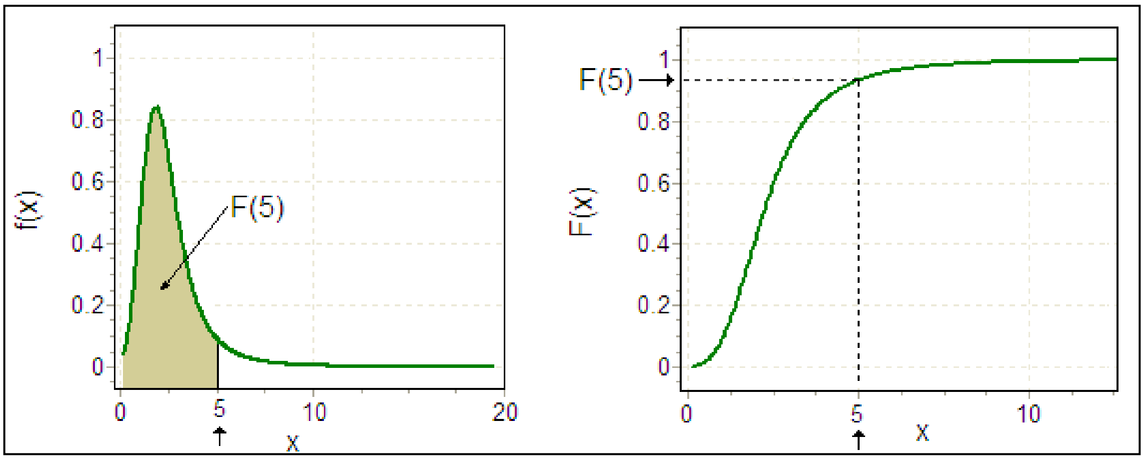


Figure 3.4: A PDF and associated CDF for a continuous random variable

Figure 3.5 shows a PMF and the associated CDF for a discrete random variable. The CDF for a discrete random variable is the sum of the probabilities t, that are less than or equal to x. Devore (2004) defines the CDF $F(x)$ by:

$$F(x) = P(X \leq x) = \sum_{\mathrm{b:b \leq x}} \mathrm{p(b)} \quad (3.9)$$

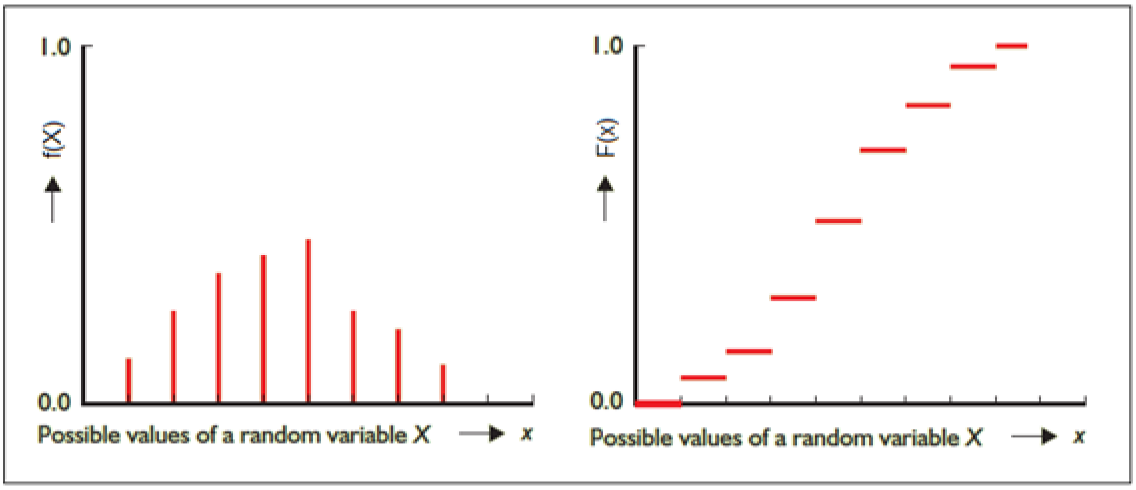


Figure 3.5: A PMF and associated CDF for discrete random variables

3.6. Parameters of Continuous Variables

It is often difficult to describe a data set's probability distribution function mathematically. There are several theoretical distributions however, for which the mathematical properties and parameters have been well studied and explained. If the frequency distribution of a dataset has a similar form to the known theoretical distribution, then the properties of the theoretical distribution can be applied to the data, allowing for a certain margin of error. Within a family of distributions, a large variety of forms is possible. The form is described by means of shape, location and scale parameters. Some distribution families do not contain all of the previously mentioned parameters.

3.6.1. Shape Parameter

Distributions that contain a shape parameter are very useful, because this allows a distribution the flexibility to take on a variety of different shapes. This in turn enables the distribution to model a variety of data sets. The shape depends on the value of the shape parameter, α. Figure 3.6 shows the Weibull distribution, with scale parameter β =1 and shape parameters α = 0.5, 1, and 3, respectively.

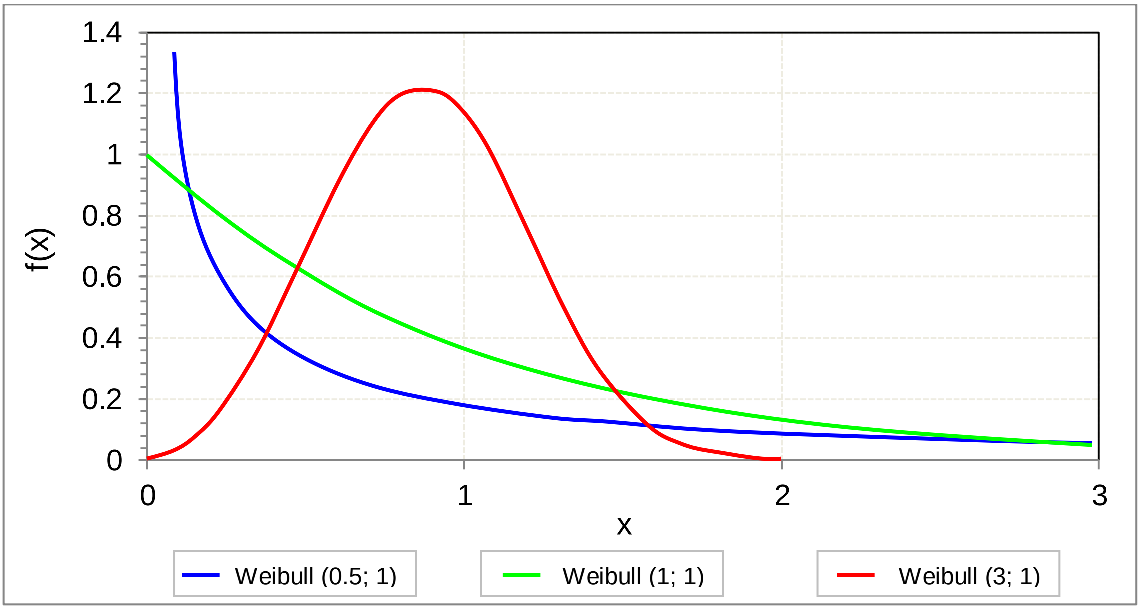


Figure 3.6: Effect of a shape parameter on the Weibull distribution

3.6.2. Scale Parameter

The scale parameter, β, has the effect of stretching or compressing the graph. The PDF will be stretched out along the x-axis if the scale parameter is greater than one. The stretching increases as the value increases. A scale parameter less than one has the effect of compressing the PDF. As the scale parameter approaches zero, the PDF makes a sharper spike. Scale parameters cannot have negative values. A location parameter of zero and scale parameters of $\beta = 0.5, 1, \text{and } 2$ are used in Figure 3.7 to show the effect of a scale parameter on a standard Normal distribution, where β and γ are the scale and location parameters, respectively, and are presented as $(\beta;\gamma)$.

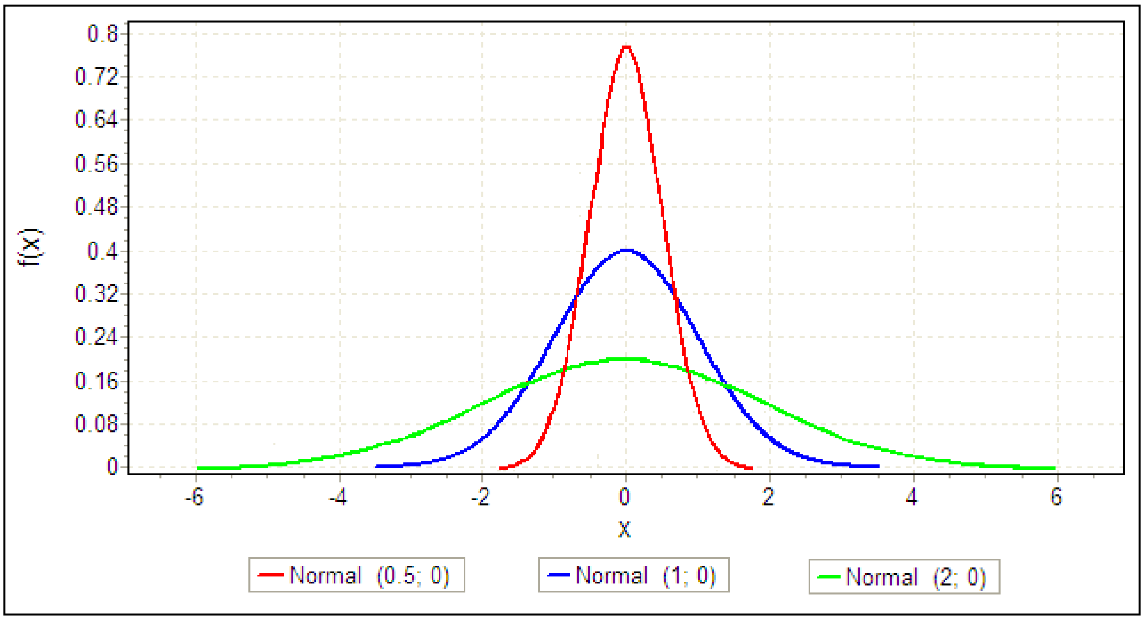


Figure 3.7: Effect of a scale parameter on the standard Normal distribution

3.6.3. Location Parameter

The location parameter, γ, has the effect of shifting the graph left or right on the horizontal axis relative to the standard distribution. If the standard Normal distribution were to be considered as an example, then a location parameter of three would translate the graph three units to the right. A location parameter of negative three would shift the graph 3 units to the left on the horizontal axis.

Figure 3.8 demonstrates the effect of a location parameter on the standard Normal distribution, where β and γ are the scale and location parameters respectively and are presented as $(\beta; \gamma)$.

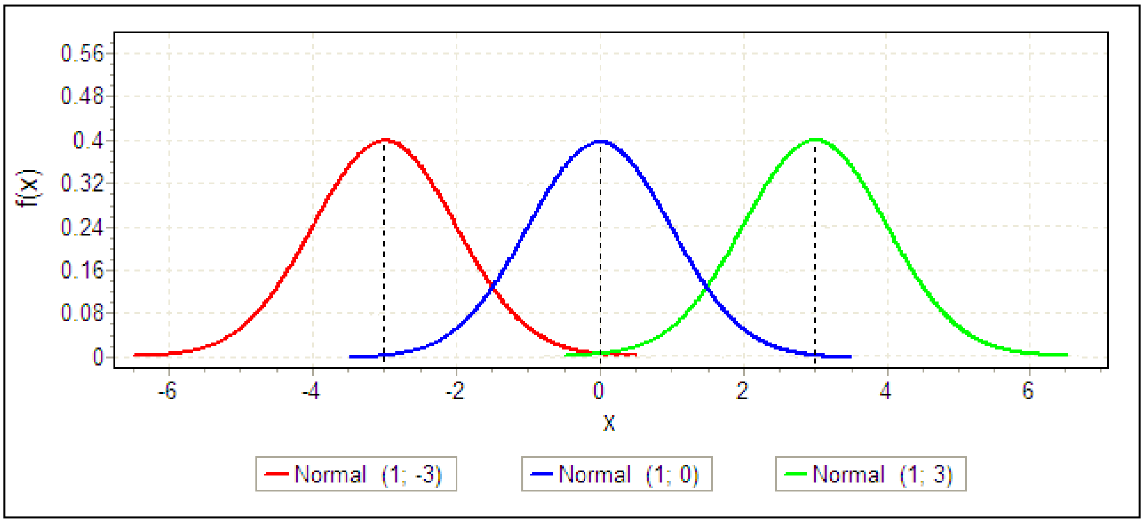


Figure 3.8: Effect of a location parameter on the standard Normal distribution

3.7. Goodness of Fit Tests

The degree to which a data set follows a given theoretical distribution is known as the goodness of fit (GOF). The compatibility between random data and a theoretical distribution can be measured with a GOF test. The Kolmogorov-Smirnov (K-S) test, the Anderson-Darling (A-D) test, and the Chi-Squared test are three examples of such tests.

3.7.1. Kolmogorov-Smirnov Test

In the Kolmogorov-Smirnov (K-S) test, the K-S statistic (D) is based on the largest vertical difference between the data's cumulative distribution and that of a specified theoretical cumulative distribution. The critical value of D is found in tables, enabling one to ascertain whether the difference between the distributions is larger than what could be expected (Johnson, 1994).

The K-S statistic is given as:

$$D = sup[|F_N(x) - F_X(x)|] \quad (3.10)$$

where:

N = total number of data points

$F_X(x)$ = the fitted CDF

$F_N(x) = \frac{N_x}{N}$

N_x = the number of $X_i's$ less than x.

The values of cumulative distributions vary from zero to one, which means that at the two extremes of a distribution, the K-S statistic will tend to be small. This has the implication that more weight is given to the centre of the distribution than the tails. The distribution must also be fully specified in terms of its location, scale, and shape parameters for the test to be valid.

3.7.2. Anderson-Darling Test

The A-D test is a modification of the K-S test. A weighting factor is multiplied to the difference between the two comparative distributions. When both $F(X_i)$ and $F(X_{N-i+1})$ approach either 0 or 1, then the weighting factor is larger at the two tails. In this way, the A-D statistic gives more weight to the tails, when compared to the K-S test.

The A-D statistic is given as:

$$\dot{A}^2 = -N - \frac{1}{N}\sum_{i=1}^{N}(2i-1)\cdot[lnF_X(x_i) + ln(1 - F_X(x_{N-i+1}))] \quad (3.11)$$

where:

N = total number of data points

$F_X(x_i)$ = value of the theoretical cumulative distribution at the largest observation x_i

3.7.3. Chi-Squared Test

The Chi-squared test requires the data to be divided into a number of bins. The bins can be in terms of either equal probability or equal width. The chi-squared statistic ($\mathcal{X}^2$) is affected by the method of binning. The fitted parameters are then used to compare the number of data points in each bin with the number of data points expected in each bin. A limitation of the test is that it is not valid for small samples. Each bin requires at least five data points for the approximation to be applicable.

The chi-squared statistics is given as:

$$\mathcal{X}^2 = \sum_{i=1}^{k} \frac{(N_i - \mathcal{E}_i)^2}{\mathcal{E}_i} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3.12)$$

where:

N_i = number of data points in bin i

$\mathcal{E}_i$ = expected number of data points in bin i

k = number of bins.

3.8. Theoretical Probability Distributions

This study made use of a number of different theoretical probability distributions to describe the continuous random variables applied in the end-use model. The Erlang, Gamma, Log-Logistic, Log Normal, Rayleigh, and Weibull distributions were applied in this study, and their mathematical descriptions are therefore briefly noted in the following sections. The equations given below were obtained from the @Risk user guide (Palisade Corporation, 2010).

3.8.1. Erlang Distribution

Parameters	α	integral shape parameter $m > 0$
	β	continuous scale parameter $\beta > 0$
Domain	$0 \leq x < +\infty$ continuous	
PDF	$f(x) = \frac{1}{\beta(\alpha-1)!}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-x/\beta}$ .... (3.13)	
CDF	$F(x) = \frac{\Gamma_{x/\beta}(\alpha)}{\Gamma(\alpha)} = 1 - e^{-x/\beta}\sum_{i=0}^{\alpha-1}\frac{(x/\beta)^i}{i!}$ .... (3.14)	

3.8.2. Gamma Distribution

Parameters	α	continuous shape parameter $\alpha > 0$
	β	continuous scale parameter $\beta > 0$
Domain	$0 \leq x < +\infty$ continuous	
PDF	$f(x) = \frac{1}{\beta\Gamma(\alpha)}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-x/\beta}$ .... (3.15)	
CDF	$F(x) = \frac{\Gamma_{x/\beta}(\alpha)}{\Gamma(\alpha)}$ .... (3.16) where: Γ = the Gamma Function Γ_x = the Incomplete Gamma Function	

3.8.3. Log-Logistic Distribution

Parameters	γ	continuous location parameter
	β	continuous scale parameter $\beta > 0$
	α	continuous shape parameter $\alpha > 0$
Domain	$\gamma \leq x < +\infty$ continuous	
PDF	$f(x) = \frac{\alpha r^{\alpha-1}}{\beta(1+r^{\alpha})^2}$ .... (3.17)	

CDF	$F(x) = \frac{1}{1+\left(\frac{1}{\mathrm{r}}\right)^{\alpha}}$ . . . . (3.18) where: $r \equiv \frac{x-\gamma}{\beta}$

3.8.4. Log Normal Distribution

Parameters	μ	continuous alternative parameter $\mu > 0$
	β	continuous scale parameter $\sigma > 0$
Domain	$0 \leq x < +\infty$ continuous	
PDF	$f(x) = \frac{1}{x\sqrt{2\pi}\beta'} e^{-\frac{1}{2}\left[\frac{lnx-\beta'}{\alpha'}\right]^2}$ . . . . (3.19)	
CDF	$F(x) = \Phi\left(\frac{lnx-\beta'}{\alpha'}\right)$ . . . . (3.20) where: $\mu' \equiv ln\left[\frac{\mu^2}{\sqrt{\beta^2+\mu^2}}\right]$ $\beta' \equiv \sqrt{ln\left[1+\left(\frac{\beta}{\mu}\right)^2\right]}$ and $\Phi(w)$ is also called the Laplace-Gauss Integral.	

3.8.5. Rayleigh Distribution

Parameters	β	continuous scale parameter $\beta > 0$
Domain	$0 \leq x < +\infty$ continuous	
PDF	$f(x) = \frac{x}{\mathrm{\beta}} e^{-\frac{1}{2}\left[\frac{x}{\beta}\right]^2}$ . . . . (3.21)	
CDF	$F(x) = 1 - \mathrm{e}^{-\frac{1}{2}\left[\frac{x}{\beta}\right]^2}$ . . . . (3.22)	

3.8.6. Weibull Distribution

Parameters	α	continuous shape parameter $\alpha > 0$
	β	continuous scale parameter $\beta > 0$
Domain	$0 \leq x < +\infty$	continuous
PDF	$f(x) = \dfrac{\alpha x^{\alpha-1}}{\beta^{\alpha}} e^{-(x/\beta)^{\alpha}}$ .... (3.23)	
CDF	$F(x) = 1 - e^{-(x/\beta)^{\alpha}}$ .... (3.24)	

4. REUWS DATABASE BACKGROUND

The parameters of the best-fit theoretical probability distribution functions used in the end-use model were obtained from measured indoor water consumption data. It was considered important to include a large sample of households in this analysis to ensure accurate results. Detailed large scale end-use measurement projects are usually very costly and time consuming. This study therefore made use of data collected by a previous project. At the time of conducting this research study, the REUWS by Mayer et al. (1999) was the largest known collection of recorded end-use data readily available. The REUWS contained indoor and outdoor end-use data and survey information previously collected by Aquacraft, Inc. of Boulder, Colorado, in the USA, and its subcontractors. The REUWS was funded by the American Water Works Association Research Foundation (AWWARF) and participating water utilities and was one of the largest end-use studies to date. A brief explanation of the data set follows in the sections below. For the full details relating to the methodology used to construct the REUWS database, see Mayer et al. (1999).

4.1. Study Sites

One of the purposes of the REUWS was to collect water consumption data from varied locations in North America. Twelve study sites in fourteen cities across the United States and Canada were therefore included in the project. At the request of the utilities, some cities combined to form one study site in order to share costs and include a wider range of water consumers. The study sites were representative of each of their locations, but not necessarily representative of all North American cities. The geographical locations of the study sites are shown in Figure 4.1. The utilities and supporting agencies which participated are listed in Table 4.1.

Figure 4.1: Sites used in the Residential end-uses of water study (Google Earth)

4.2. Study Group Selection

Mayer et al., (1999) explained that in each utility a representative sample of 1 000 single family households was selected, to whom a questionnaire survey was mailed. The account number, service address, account status, date of account initiation, meter reading dates, meter readings and consumption data for a twelve month period was collected for each of the mail survey targets. The sample of 1 000 homes in each study site was referred to as the “Q1000” database.

The Q1000 database went through various quality assurance and control tests, one of which was to test whether the sample was statistically representative of the population. Statistically significant differences occurred in only one site, namely Tempe, Arizona. Corrective action was later performed in that case during the study group selection process.

The survey that was mailed to the Q1000 consumers database included questions relating to water-using appliances and fixtures, water using habits, household and landscape characteristics and demographic information.

To ensure customer anonymity, unique key-codes were used in subsequent databases to identify customer responses. Each key-code consisted of five numbers. The first two digits represented the study site, and the following three digits designated the residential customer. The key-code assignments that were used in the REUWS are presented in Table 4.1.

Table 4.1: REUWS cities and key-code assignments (Mayer et al., 1999)

Key-code	City/Utility
10000 - 10999	Boulder, Colorado
11000 - 11999	Denver, Colorado
12000 - 12999	Eugene, Oregon
13000 - 13999	Seattle, Washington (includes 4 water purveyors in the Seattle area)
14000 - 14999	San Diego, California
15000 - 15999	Tampa, Florida
16000 - 16999	Phoenix, Oregon
17000 - 17999	Tempe and Scottsdale, Arizona
18000 - 18999	Regional Municipality of Waterloo, Ontario (includes the cities of Waterloo and Cambridge)
19000 - 19999	Walnut Valley Water District, California (part of the Metropolitan Water District (MWD))
20000 - 20999	Las Virgenes Valley District, California (part of MWD, includes Calabasas and surrounding communities)
21000 - 21999	Lompoc, California

Statistical tests were performed as part of the work, to establish whether significant water consumption differences existed between survey respondents and survey non-respondents. Corrective action was taken where needed; an example of corrective action was the removal of outliers, where such removal was justified, by Mayer et al. (1999).

A target of approximately 100 homes in each study site was chosen in which to install data-loggers. These households formed a sub-sample of the mail survey respondents. The households in the logging sample went through another test to ensure that statistically representative houses were selected before the group was approved. Consent letters were sent to the data logging sample explaining the project. In total approximately 40 households chose not to participate, and those were replaced by other households. The total logging sample in the REUWS ultimately consisted of 1188 households.

4.3. End-Use Data Collection

The water consumption data was collected by means of a portable data logger attached to the water meter (which measured both the indoor and outdoor water consumption) at each of the 1188 selected houses. The data logger recorded the average volume of water passing through the water meter every ten seconds. In the REUWS a total of 100 loggers were used at any one time, with ten additional loggers available as backup. This meant that the loggers had to be rotated amongst the 100 homes in each of the 12 study sites. The target collection period was two weeks in the summer and two weeks in the winter for each house. The schedule of when data was collected is shown in Table 4.2.

Table 4.2: Data collection schedule (Mayer et al., 1999)

Site	City	Data collection period	
		1	2
1	Boulder, Colorado	21 May - 7 June, 1996	3 Sep - 19 Sep, 1996
2	Denver, Colorado	5 June - 21 June, 1996	27 May - 13 June, 1997
3	Eugene, Oregon	24 June - 11 July, 1996	1 Dec - 20 Dec, 1996
4	Seattle, Washington	16 July - 2 Aug, 1996	7 Jan - 24 Jan, 1997
5	San Diego, California	6 Aug - 26 Aug, 1996	3 Feb - 21 Feb, 1997
6	Tampa, Florida	1 Oct - 18 Oct, 1996	3 Mar - 21 Mar, 1997
7	Phoenix, Arizona	6 May - 23 May, 1997	4 Nov - 21 Nov, 1997
8a, 8b	Scottsdale and Tempe, Arizona	29 Oct - 15 Nov, 1997	2 Dec - 19 Dec, 1997
9a, 9b	Waterloo and Cambridge, Ontario	24 June - 11 July, 1997	7 Oct - 24 Oct, 1997
10	Walnut Valley, California	22 July - 8 Aug, 1997	6 Jan - 23 Jan, 1998
11	Las Virgenes, California	12 Aug - 29 Aug, 1997	27 Jan - 13 Feb, 1998
12	Lompoc, California	9 Sep - 26 Sep, 1997	24 Feb - 13 Mar, 1998

4.4. End-Use Data Analysis

The flow rates (or flow traces) recorded by the data loggers were analysed using Trace Wizard software. The recorded flow data was disaggregated into water consumption events, as explained earlier in this thesis. While determining the separate events, the start time, stop time, duration, volume, peak flow rate, mode flow rate, and mode frequency was calculated for each event. Thereafter water consumption events were categorized and assigned to a specific end-use in the household. Trace Wizard was employed by implementing user defined parameters for each household. The parameters consisted of ranges of possible values for volume, flow rate, and duration, which were unique to a particular end-use. An analyst on their team repeated the routine and fine-tuned the parameters to build a parameter file that correctly identified as many end-uses as possible, based on expert-input (Mayer et al., 1999).

A number of the end-use event data and characteristics obtained from the REUWS were used to derive probability distributions in the end-use model developed in this study. Section 5 provides the full details regarding the data that was extracted from the REUWS database and that was ultimately utilised in the end-use model.

5. PROBABILISTIC END-USE MODEL

5.1. Overview

One of the objectives of this study was to develop a computer based stochastic end-use model which would generate residential flow profiles on a high resolution temporal scale. The compilation of diurnal flow rate profiles was a prerequisite for achieving a further objective of the study, which was to calculate peak factors for differently sized areas, using different time intervals.

The model which was constructed to reach the above mentioned objectives is discussed in this chapter. In the research design section the selected technique (statistical simulation model) is reviewed. The choice of software, as well as an explanation of the model concept and structure, is then provided. Within the methodology section, the data used as input to the model is fully described, and thereafter a detailed description of the characteristics of each parameter in the model is presented.

5.2. Research Design

A model was sought that could be applied as part of this research into peak flow. Ideally, this would consist of a mathematical model that could be used for repetitive calculations. A number of water demand estimation models were discussed in section 2.3. The SIMDEUM was very successful in constructing diurnal water patterns by describing the the arrival time, the intensity and duration of rectangular water pulses, with probability distributions for each end-use. A similar statistical simulation model was therefore selected as that to be developed in this study. The previously measured end-use data from REUWS was used to obtain the descriptive probability distributions. The entire REUWS database, including all the raw data compiled during that study, was purchased from the original authors as part of this research project. The REUWS data was considered the most appropriate data available for the purpose of this study.

5.3. Software

5.3.1. End-Use Model Software Choice

A number of software options that could be used to develop a statistical model were available. Microsoft Excel is one of the Microsoft Suite of Applications. It is a widely known and utilised software application with a spreadsheet interface. It has many tools and features that make it possible to analyse data and perform complex calculations. Other possible software applications require the use of programming languages such as C++, Delphi, Fortran or Matlab. Some advantages and disadvantages of the software options are presented in Table 5.1.

Table 5.1: Comparison of software options to construct an end-use model

Software	Microsoft Excel	C++ / Delphi / Fortran / Matlab
Advantages	Microsoft Excel is a readily available and commonly used package	Computation speed of model can be enhanced
	The calculation equations within the model are visible, enabling easy understanding of the structure	With well written code, the model can be compact and efficient, with a executable program
	Little knowledge of programming language is required	Ideal for iterative calculations
	Future work or improvement of the model is not limited by knowledge of a specific programming language	Changes in the model structure are easily incorporated with additional code
Disadvantages	The model could become clumsy, as a number of calculation steps may be required to perform single processes	Software may have to be purchased
	Workbooks containing large volumes of data may limit the computation speed of the model	Calculations are obscured by programming code
	It is not ideal for iterative calculations	Comprehensive knowledge of a programming language is essential
	Small changes in the model structure may require a lot of rework	Future work or improvement of the model is limited by knowledge of specific programming language

Microsoft Excel was chosen as the preferred software application in this study, mainly because it was readily available, suitable for the desired purpose and the disadvantages were not considered to be limiting in terms of this research project’s outcomes.

A stochastic end-use model generating daily residential flow rate profiles has not previously been developed in South Africa. The priority of this study was therefore not to acquire knowledge of a particular programming language, but rather to establish whether the model structure presented in section 5.4 would be successful in achieving the objectives of this study. Since the model structure could be fully tested and constructed in Microsoft Excel, this was chosen as the preferred software application. Should the model prove successful, future work could include improving the model by converting it to a more efficient software application with the use of a programming language. The software packages that were used in this study are discussed in more detail below.

5.3.2. Microsoft Access

Microsoft Access is one of the Microsoft Office Suite of applications. It is an effective software application for the purpose of creating and managing large databases. User-friendly features enable information to be manipulated and viewed easily. Macros can be used to create or connect tables, queries, filters, forms, and reports. Microsoft Access was used in this research project to extract data from the REUWS database and to perform queries on the dataset.

5.3.3. Microsoft Excel

Microsoft Excel is also one of the Microsoft Suite of applications. It is a widely known and used software application with a spreadsheet interface. It has many powerful tools and features that makes it possible to analyse, share and manage data. Microsoft Excel was chosen to store the results for this project because most of the people who need it have access to the application, and the management of the data is facilitated by the creation of a worksheet for each different suburb, while maintaining a minimum number of workbooks. Microsoft Excel was used in this research project to develop the end-use model.

5.3.4. @Risk

The @Risk software is a risk analysis and simulation Add-in for Microsoft Excel. It contains many functions which allow different distribution types to be specified for cell values. The @Risk software has simulation capabilities with supported techniques such as Monte Carlo and Latin Hypercube sampling. The software also performs GOF tests by making use of the K-S test, the A-D test, and the Chi-Squared test to compare the theoretical distributions with the given data. Each GOF test ranks the distributions based on the fit. The available distributions in @Risk are as follows:

- Beta
- Beta General
- Beta-Subjective
- Binomial
- Chi-Square
- Cumulative
- Discrete
- Discrete Uniform
- Error Function
- Erlang
- Exponential
- Extreme Value
- Gamma
- General
- Geometric
- Histogram
- Hyper geometric
- Inverse Gaussian
- IntUniform
- Logistic
- Log-Logistic
- Lognormal
- Lognormal2
- Negative Binomial
- Normal
- Pareto
- Pareto2
- Pearson V
- Pearson VI
- PERT
- Poisson
- Rayleigh
- Student's t
- Triangular
- Trigen
- Uniform
- Weibull

The @Risk software was used in this research project to perform GOF tests on selected elements in the REUWS data in order to rank the distributions according to the best fit. Based on the results of the rankings, theoretical distributions were selected to describe some of the end-use model parameters.

5.4. Model Structure

The end-use model developed in this study determined the flow rates caused by end-use water events at a resolution of one second. Instantaneous flow rates that occur within a single water consumption event generally vary in magnitude. In this study, however, it was assumed that a constant flow rate occurred for the entire duration of a single end-use event. The water events could therefore be equated to rectangular water pulses. Buchberger and Wells (1996) showed that rectangular water pulses described indoor residential water demand successfully in their PRP model.

The concept of rectangular pulses is illustrated in Figure 5.1. The height and length of the rectangle represented the flow rate and the duration of the event, respectively. The area of the rectangle represented the volume of the event. When water events overlapped, the flow rate was calculated as the sum of the flow rates for the overlapping period. In summary, each time an end-use was activated, it caused a rectangular pulse, and when individual water demand events from different end-uses were added together, the total water demand profile for a single household was obtained.

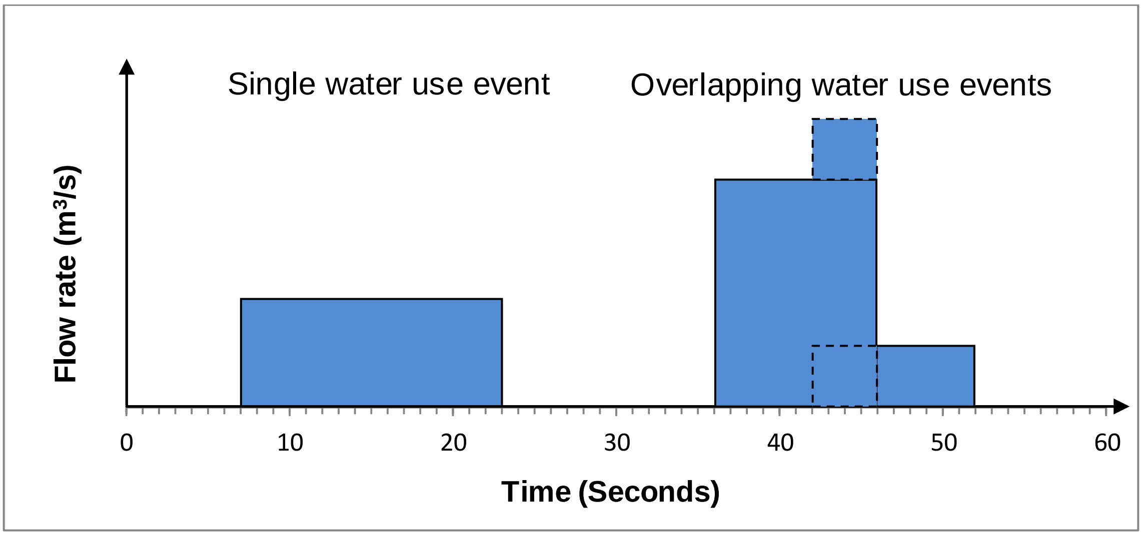


Figure 5.1: Rectangular water pulses

In the end-use model, the elements required for the rectangular pulses were obtained from end-use specific probability distribution functions for the flow rates and volumes. The time of the day that end-use events occurred was also determined from end-use specific probability distributions.

Water demand is strongly related to the number of people residing in a home (household size), who are available to use water. The number of times that a specific end-use was activated during a single day (the frequency) was therefore related to household size. Six household size categories were included in the model, ranging from 1 PPH to 6 PPH.

Once the model had selected a household size, a corresponding probability distribution was applied, to establish the number of events that occurred on the simulated day. Starting hours were assigned to every event, based on starting hour probability distributions. The individual event starting times were subsequently obtained using random minutes and seconds within each selected starting hour.

End-use specific probability distributions were used to obtain event flow rates and volumes, which were assigned to every event starting time. The duration and ending times for the water events were then computed with the available information. In addition, end-uses with cyclic water demand patterns, such as dishwashers and washing machines, involved prescribing the number of cycles per event and the duration between cycles with probability distributions. The parameters of all the probability distributions were derived from water consumption measured data from the REUWS.

With the above mentioned end-use event values in place, the flow rates occurring throughout the day from different end-uses were summed. The flow rate observed in one second intervals for a single house of particular household size was subsequently available. A Monte Carlo simulation method was applied in the model to generate many unique water demand scenarios. A Monte Carlo method entails the repeated calculation of the model, each time using randomly selected input parameters for the probability distributions.

Figure 5.2 shows a simplified schematic of the end-use model structure. The number of cycles per event, and duration between cycles parameters were only applicable to the dishwasher and washing machine end-uses, the dashed lines illustrates this in the figure.

The mathematical description of the end-use model is very similar to the SIMDEUM developed by Blokker et al. (2010). Key differences between the two models were that the end-use model in this study derived the input data from measured water consumption, while the SIMDEUM used statistical information on consumers and end-uses based on consumer surveys as input. The SIMDEUM determined the actual household size, and simulated the water demand for each occupant separately; these were then summed. The current end-use model simulated households as a whole, with a possible range of one to six PPH. The SIMDEUM also considered end-use sub-types such as single- and dual-flush toilets, while the end-use model in this study incorporated the effects of sub-types in the probability distributions derived from the previously measured data.

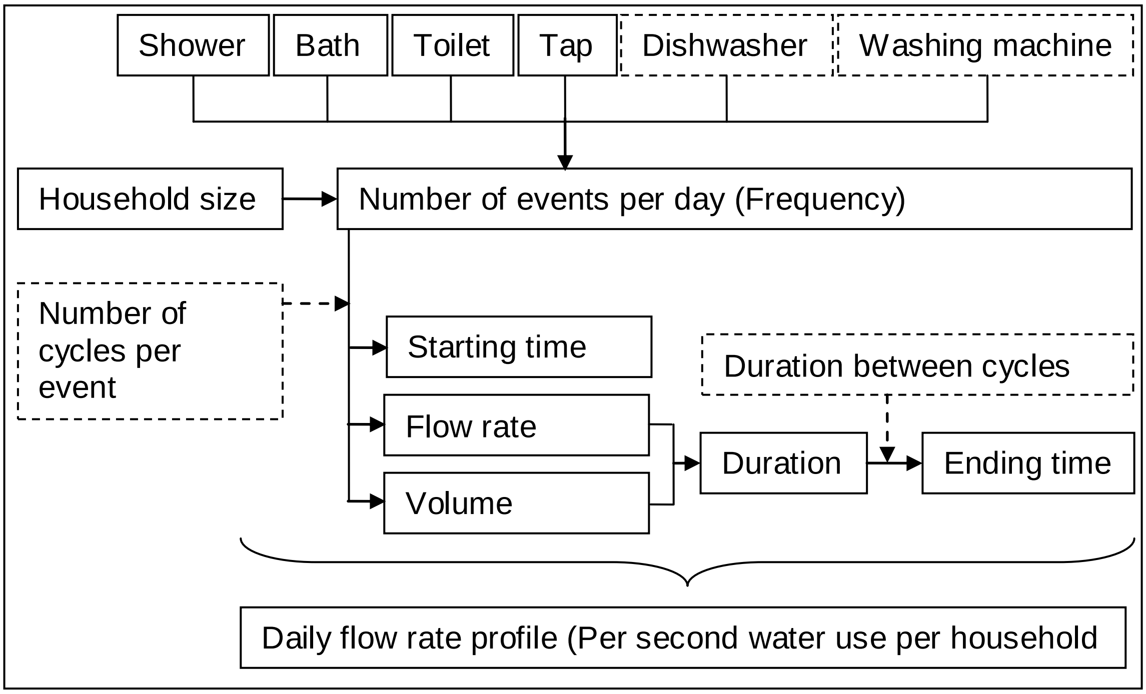


Figure 5.2: Simplified schematic of end-use model structure

5.5. Data Preparation

5.5.1. REUWS Database

The comprehensive REUWS database was purchased as part of this research and was presented on a CD as a 230 Mega Byte (MB) Microsoft Access 2003 database. Microsoft Access 2010 was available for this study, so the database was converted into the later format as part of this research project, for further manipulation. The database contained 20 tables of data which were linked with the key-code field, as presented in Table 4.1. A description of some of the tables within the REUWS database described by Mayer et al. (1999) is presented in Table 5.2.

Table 5.2: Overview of the REUWS database (Mayer et al., 1999)

Table name	Description
Daily use	Summed water consumption volume by end-use for each logged day from each city
Survey responses	Each coded survey response from 12 cities. Also includes the key-code field to link water consumption and survey data
Q1000	There are twelve Q1000 tables in the database, one for each of the 12 participating cities. These tables contain historic billing records for a random sample of 1000 single family accounts in the service area. The fields in each table vary, but the units of water consumption are kgal (thousand gallons)
Irrigated area data	Measured irrigated area and irrigation application rate from 1130 of the 1188 single family households in the study
Weather data tables	Weather stations and daily weather tables contain climate data from weather stations near each study home. These tables are related to each other by the station ID field
Logging data	Each individual water event recorded during the two-year study is included in this table, including toilets, showers, washing machines, taps, irrigation, etc. Logging data is related to survey responses via the key-code field

Within the logging data table, various fields were available. A description of some of the fields is provided in Table 5.3.

Table 5.3: Description of the REUWS logging data table fields (Mayer et al., 1999)

Field name	Description
Key code	The unique identifier for each household in the study
Use type	The type or category of water consumption. For example, toilet, shower, washing machine, etc.
Date	The date the water consumption event occurred. For events that start at 11:59 p.m. and extend into the next day, the date is the start date
Start	The start time of the water consumption event
Duration	The duration (in seconds) of the water consumption event
End	The end time of the water consumption event
Volume	The volume (in gallons) of the water consumption event
Peak	The peak flow rate (averaged over 10 seconds, and presented as gallons per minute) observed during the course of the water consumption event
Mode	The mode flow rate (averaged over 10 seconds, and presented as gallons per minute) observed during the course of the water consumption event
Mode No.	The number of occurrences of the mode flow rate during the water consumption event

In the Use type table field, the labels of end-uses identified by Trace Wizard were:

- Bath
- Clothes washer
- Clotheswasher1
- Cooler
- Dishwasher
- Dishwasher1
- Faucet
- Hot tub
- Humidifier
- Leak
- Irrigation
- Shower
- Swimming pool
- Toilet
- Toilet@
- Treatment
- Unknown

The terms “faucet” and “clothes washer” were used in the REUWS; however the terms “tap” and “washing machine” are used to refer to the respective end-uses in this study.

The tap use type did not discriminate between the taps used in a bathroom, and taps in the kitchen.

The dishwasher and washing machine events occurred as sequential cycles, with a time laps between cycles. The first cycle in each multi cycle event was named washing machine1, and dishwasher1, respectively. This enabled the number of multi-cycle events per day to be counted easily.

The toilet@ label was given to toilet flush events which appeared to be partial or double flushes. Mayer et al. (1999) stated that the toilet@ use types did not reflect accurate flush volumes, but should be incorporated in daily count applications.

To reduce the number of records, Mayer et al. (1999) summed the leakage events daily and gave the total as the leak value. Where flow trace analysis could not confidently identify events, such events were placed in the unknown category prior to the completion of the REUWS.

5.5.2. Table Selection

Only indoor water demand was considered in this study, which eliminated the need for some of the tables available in the REUWS database. The tables with weather data and properties' irrigated areas were thus omitted. Daily totals or historic billing records were not required for the development and application of the end-use model, therefore the daily use table and the Q1000 tables were also excluded.

End-use data and customer characteristics were the essential information required to extract from the REUWS database. The logging data table and the survey responses table were subsequently used. The survey responses were considered more useful when linked to the customers’ logging data. However, the survey responses table contained all response data, including that of households that did not participate in the end-use logging portion of the study. A single table was therefore created, which included all the logging data as well as the corresponding survey responses. Such a merge was possible because of the common key-code field identifying the households. The additional survey responses were omitted.

The two tables were combined by creating a query in Microsoft Access. The logging data table and the survey responses table were selected, and all the columns contained in the tables were merged. The query was converted into a table named "Logging data and survey responses". This table contained 1 959 120 records, with 1187 unique key-codes (or households). The column titles in the logging data and survey responses table were as listed in Table 5.4. The column numbers are represented as they appeared, in order from left to right. Q1, for example, represents the answers to Question one of the survey. Column numbers 14 to 96 and 103 to 114 were not utilised, and are thus shown as condensed lines in Table 5.4.

Table 5.4: Logging data and survey responses table columns

Column Number	Column Title
1	Logging data key code
2	Use type
3	Date
4	Start
5	Duration (s)
6	End
7	Peak (gal/min)
8	Volume (gal)
9	Mode (gal/min)
10	Mode No.
11	Survey responses key code
12	Stat ID
13	Stat ID 2
14 - 96	Q1a - Q29
97	Q30adults
98	Q30teen
99	Q30child
100	Q31adults
101	Q31teen
102	Q31child
103 - 114	Q32 - Comments

5.5.3. End-Use Selection

The REUWS identified 14 different end-uses. However, not all of the end-uses were considered essential for inclusion in the end-use model developed here. In order to identify the relevant end-uses, the percentage of households in which the events occurred was investigated. Table 5.5 shows a summary of end-use occurrence in households.

Table 5.5: End-use occurrence in households extracted from REUWS data

End-use type	Number of records	Number of households	Proportion of households (%)
Tap	1 150 872	1 187	99.9
Toilet + Toilet@	348 345	1 186	99.8
Washing machine1 + Washing machine	120 756	1 160	97.6
Dishwasher1 + Dishwasher	33 832	906	76.3
Shower	50 286	1 172	98.6
Bath	4 105	556	46.8
Irrigation	69 245	1 117	94.0
Swimming pool	5 147	111	9.3
Leak	27 587	1 184	99.7
Unknown	27 881	1 020	85.6
Cooler	102 063	64	5.4
Hot tub	896	38	3.2
Humidifier	3 861	11	0.9
Treatment	14 244	180	15.1

Leakage and outdoor data were not included in this study, which meant the end-uses called leak, irrigation, and swimming pool were automatically excluded. The Unknown data would not be useful in this research, and was, subsequently, excluded. The cooler, hot tub, humidifier, and treatment were end-uses that were present in fewer than 20% of the study group households. They were therefore considered uncommon in both the majority of households from REUWS and local households, and were excluded in the current study. The six remaining end-uses which were utilized in the end-use model were the tap, toilet, washing machine, dishwasher, shower and bath. A total of 12.1% of the records was excluded.

In order to isolate the data for each of the six applicable end-uses, queries were created in Microsoft Access. In the query, the “Logging data and survey responses” table was selected with all its columns. However, only records with the applicable use-type names, for example the word “bath” in the “use type” column was included in the results. The query for “bath” was converted in a table named “Bath original”. The process was repeated for each end-use, until six tables had been created, containing relevant data for the six selected end-uses.

5.5.4. Unit Conversions

The units of measurement for the volume and flow rate fields in REUWS were given in gallons, and gallons per minute respectively. For the end-use model the units were converted to litres for the volume and litres per second for the flow rate. To perform the conversions, three additional columns were created in each of the end-use tables. Table 5.6 summarises the properties allocated to the three new columns in Microsoft Access.

Table 5.6: Unit conversion column properties

Column name	Data Type	Expression
Peak (ℓ/s)	Calculated	$[Peak\ (gal/min)]\ x\ 3.854118\ /\ 60$
Volume (ℓ)	Calculated	$[Volume\ (gal)]\ x\ 3.854118$
Mode (ℓ/s)	Calculated	$[Mode\ (gal/min)]\ x\ 3.854118\ /\ 60$

5.5.5. Household Size Calculation

Household size is one of the most notable parameters influencing water consumption. The number of times that a particular end-use was used during the day is a function of household size, described as the number of people in the household. For example, it makes sense that the toilet is flushed more frequently in a four-person household than in a two-person household.

The number of events occurring per day was thus related to household size in the model. The household size information was gathered from the survey responses data. The two relevant questions quoted from the questionnaires are as follows:

Question 30: How many people reside full-time at this address during the winter months of the year (generally December - February)? (*Enter the number of individuals in each age group.*)

Adults (18+) b) Teenagers (13 - 17) c) Children (under 13)

Question 31: How many people reside full-time at this address during the summer months of the year (generally June - August)? (*Enter the number of individuals in each age group.*)

Adults (18+) b) Teenagers (13 - 17) c) Children (under 13)

The responses to the two questions were provided in the columns entitled Q30adults, Q30teen, Q30child, Q31adults, Q31teen, and Q31child. The first step was to insert columns entitled “Q30total” and “Q31total” in the table which gave the total number of people living in the house in the winter and summer months, respectively. This was done by taking the sum of adults, teenagers and children reported in the survey responses, since it was not considered necessary in this study to distinguish between different consumer age groups.

The next step was to determine the month in which an event took place, so that each event was related to the corresponding winter or summer household size. A column entitled “Month” was added in each end-use table, which evaluated the “Date” column. Based on the event date, a number between 1 and 12 representing the month of the year in which the event took place was computed in the “Month” column. Months 3, 4, 5, 6, 7, and 8 (March to August) were considered summer months, while months 9, 10, 11, 12, 1 and 2 (September to February) were considered winter months, due to the fact that the REUWS data originated from the northern hemisphere.

A column entitled “Household size” was added to each end-use table which evaluated the “Month” column. If the month was a value between 3 and 8, then the household size value in the “Q31total” column was used. If the month value was not between 3 and 8, then the household size value in “Q30total” was used. This ensured that each event was assigned the appropriate household size value.

There were cases where the date of an event corresponded to a summer month, the Q31total column (summer month household size) contained no response, but the Q30total column (winter month household sizes) had a value. In such cases it was assumed that the same number of people residing in the winter months was present in the summer months, or visa versa. Therefore, in all instances where the household size column yielded values of 0, the appropriate value from either the Q30total or Q31total columns was manually copied. Table 5.7 summarizes the properties allocated to the new columns discussed above.

Table 5.7: Household size column properties

Column name	Data Type	Expression
Q30total	Calculated	$[Q30adults] + [Q30teen] + [Q30child]$
Q31total	Calculated	$[Q31adults] + [Q31teen] + [Q31child]$
Month	Calculated	$Month([Date])$
Household size	Calculated	$\mathrm{If}([\mathrm{Month}] \geq 3 \text{ and } [\mathrm{Month}] \leq 8, [\mathrm{Q31total}], \mathrm{Q30total})$

The range of household size from the REUWS data sample was 8 PPH, varying from 1 to 9. The proportion of the end-use events that occurred within each household size category was extracted and presented in Table 5.8.

Table 5.8: Household size category proportions

Household size (PPH)		1	2	3	4	5	6	7	8	9	Total
Proportion of events (%)	Bath	1.9	18.4	17.4	23.4	21.4	11.4	2.7	2.2	1.2	100.0
	Washing machine	1.9	20.0	19.3	26.5	17.5	9.0	3.1	2.0	0.8	100.0
	Dishwasher	1.7	23.9	18.4	28.2	18.0	6.1	2.3	0.8	0.5	100.0
	Tap	16.1	37.1	19.3	16.4	7.7	2.5	0.6	0.2	0.1	100.0
	Shower	6.2	31.8	21.1	22.7	12.0	4.1	1.0	0.9	0.2	100.0
	Toilet	7.5	34.9	19.9	21.4	10.8	3.6	1.2	0.5	0.2	100.0
	Total	12.9	35.2	19.5	18.4	9.2	3.2	0.9	0.4	0.2	100.0

From Table 5.8 it is clear that the lowest proportion of events occurred in the 6 PPH, 7 PPH, 8 PPH, and 9 PPH categories, which together were only responsible for 4.8% of the total number of events. It was therefore decided, for the purpose of the end-use model, to group the household size categories for 6 PPH, 7 PPH, 8 PPH, and 9 PPH into a single category, represented for the purpose of simplicity by 6 PPH. All the events occurring within households of the above mentioned sizes therefore formed part of the 6 PPH category.

Queries were created in Microsoft Access to disaggregate the data in each end-use table, based on household size. In the query, each end-use table was selected in turn, including all the columns in the table. However only records that, for example, had values for $0 < \mathrm{PPH} \leq 1$ in the “Household size” column were included in the results for the 1 PPH query. The process was repeated so that 2 PPH, 3 PPH, 4 PPH, 5 PPH and 6 PPH queries resulted for each of the six end-use categories. The individual household size tables were necessary so that the daily frequency of events for each end-use could be determined based on household size.

5.5.6. Export Microsoft Access Tables to Microsoft Excel

A separate Microsoft Excel workbook was created for each of the tap, toilet, washing machine, dishwasher, shower and bath end-uses. The data in the modified “Bath” Microsoft Access table was exported to a sheet in the “Bath”

Microsoft Excel workbook. The “Bath” queries for 1 PPH, 2 PPH, 3 PPH, 4 PPH, 5 PPH and 6 PPH were each exported to a separate worksheet in the Bath Microsoft Excel workbook. The export process was repeated for each end-use.

Microsoft Excel allowed only 65 534 entries to be imported at a time. This meant that in cases such as the toilet end-use that consisted of 348 345 event entries, the data was copied in sections, with each section exported as a new worksheet. The data in the separate worksheets were later combined into a single worksheet. Due to the number of event entries present in the tap end-use data, the tap data had to be exported as 20 separate sections, hence 20 worksheets were created. Furthermore, a single Microsoft Excel worksheet was limited to 1 048 576 rows. Since the number tap end-use event entries exceeded this number, the additional rows were positioned in an adjacent table on the same sheet.

5.5.7. Duration Between Cycle Calculation

As mentioned previously, the dishwasher and washing machine events consisted of a number of sequential cycles. Each cycle had a start time and duration, however the REUWS database did not explicitly provide the duration between cycles (in seconds). Instead only the start times were provided. Additional columns were therefore created in the relevant end-use workbooks to calculate the duration between cycles.

To improve clarity, each event was assigned a unique number, with each cycle in the event having the same number. A column was inserted where the start time of one cycle was subtracted from the end time of the previous cycle within one event. This served to calculate the time between cycles, given in hh:mm:ss format. Another column was inserted which served to convert the time into seconds, resulting in the “duration between cycles” variable.

5.5.8. Data Filtering in Microsoft Excel

The data provided in the REUWS database had already gone through a rigorous filtering process, which had been checked and verified by Mayer et al. (1999), as discussed earlier in this text. The REUWS verification was considered to be sufficient and no additional filtering of the given values was done. However, there were cases where obvious errors required correction. Most of the adjustments occurred in the start time, end time, or duration fields.

In the duration field, there were cases where the duration value was given as -86340, 0, or 82 seconds. Negative and zero durations are of no use, and since the logging measurements occurred at 10-second intervals, all event durations were expected to be factors of ten. Hence, the quoted duration of 82 seconds was suspicious. When the difference between the event start time and end time was calculated, the actual durations resulted in, for example, 60, 160 and 80 seconds, respectively. The calculated durations were given preference when there were differences between the given and calculated durations. In cases where the event start time and end time were equal, and the duration was zero seconds, the entire record was removed, since there was no method of establishing the actual duration, or whether any event had actually occurred.

Within the dishwasher and washing machine data sets there were events that appeared to have an improbably large number of cycles. When this was investigated it was usually because the first cycle of a new event was labelled “Dishwasher” instead of “Dishwasher1”. It was possible to identify such problems by means of inspection, by noting where new events should have started. This situated was typified by excessively long durations between cycles. There were also instances where the start time of a cycle in one event would occur before the previous cycle of the same event finished. Where there was an overlapping of cycles in a single event, one of the cycles was removed. Events with just one cycle, or duplicate cycles, were also removed.

5.6. Household Size Frequency Probability Distributions

Section 5.5.5 describes how the household size (number of persons per household) categories were calculated. The same information was used to obtain the total number of events within each household size category. The household size was considered a discrete variable, so an estimation of the probability distribution was obtained by calculating the relative frequency, and thereafter the cumulative relative frequency, of the household size categories. The cumulative relative frequency was necessary, because it was the distribution applied in the end-use model for the household size selection process. The resulting probabilities used for the household size are given in Table 5.9.

Table 5.9: Household size probability calculation

Household size (PPH)	Frequency	Relative Frequency (Probability)	Cumulative Relative Frequency
1	214932	0.129	0.129
2	586075	0.352	0.482
3	323637	0.195	0.676
4	305498	0.184	0.860
5	153744	0.092	0.952
6	79392	0.048	1.000
Sum	1663278	1.000	

5.7. Daily Event Frequency Probability Distributions

The daily frequency of events was related to the household size. Event frequencies were discrete variables, since it represented the precise number of events for a particular household. For example, it was not possible for 3.75 shower events to occur in a day, so the frequency values did not comprise an entire interval on the number line, as would be the case with continuous variables. Probability distributions for discrete variables were therefore determined by calculating the relative frequencies and cumulative frequencies of the data. The procedures followed to determine the probability distributions of daily event frequencies are described in section 5.7.1 and 5.7.2.

5.7.1. Bath, Shower, Toilet, and Tap Probability Distribution

Using the bath as an example, all the data in the previously disaggregated 1 PPH category was considered. The key-codes were sorted in ascending order, and within that arrangement the event dates were sorted in ascending order. The “subtotal” function in Microsoft Excel was used in such a way that at every change in date, a subtotal was inserted which counted the number of events occurring on the same date. The subtotal values were used to construct a table containing the number of bath events that occurred on individual days in each household. The table was then sorted so that the “number of events per day” was arranged in ascending order. The subtotal function was used on the resulting table, this time inserting a subtotal at every change in the “daily event frequency” column, and counting the number of events in each category. The values of the subtotals were used to construct a table containing the number of days on which the event frequencies took place. For the 1 PPH category, the bath end-use resulted in daily event frequencies ranging between one and four events per day. The daily event frequencies of zero were assumed to occur on the logged days that had no recorded bath event. The resulting probabilities used for the bath data in the 1 PPH category are given in Table 5.10.

Table 5.10: Bath probability calculation for household size of one person

Daily Event Frequency (Number of baths per day)	Frequency	Relative Frequency (Probability)	Cumulative Relative Frequency
0	284	0.612	0.612
1	123	0.265	0.877
2	39	0.084	0.961
3	14	0.030	0.991
4	4	0.009	1.000
Sum	464	1.000	

The result suggests that in households with one person, the probability that no bath event takes place on any given day is 0.612. The probability that three bath events occur on one day is 0.030, et cetera.

For the bath end-use the probability calculation procedure explained above was repeated for the 1 PPH, 2 PPH, 3 PPH, 4 PPH, 5 PPH, and 6 PPH categories. The entire process was also performed for the shower, toilet, and tap end-uses.

The cumulative relative frequencies can also be represented graphically, as shown in Figure 5.3 for the 1 PPH bath end-use. Figure 5.3 also illustrates the Monte Carlo methodology of how the model selected the number of events per day. A random number (with a uniform probability distribution) between zero and one was generated by Microsoft Excel (applied to the y-axis), and the appropriate number of events was selected (on the x-axis). For example, if the random number was 0.75 then one bath event would have occurred, as shown by the dotted lines on Figure 5.3. The cumulative relative frequencies used in the model for the bath, shower, toilet, and tap end-uses for the different household size categories are summarised in Appendix B.

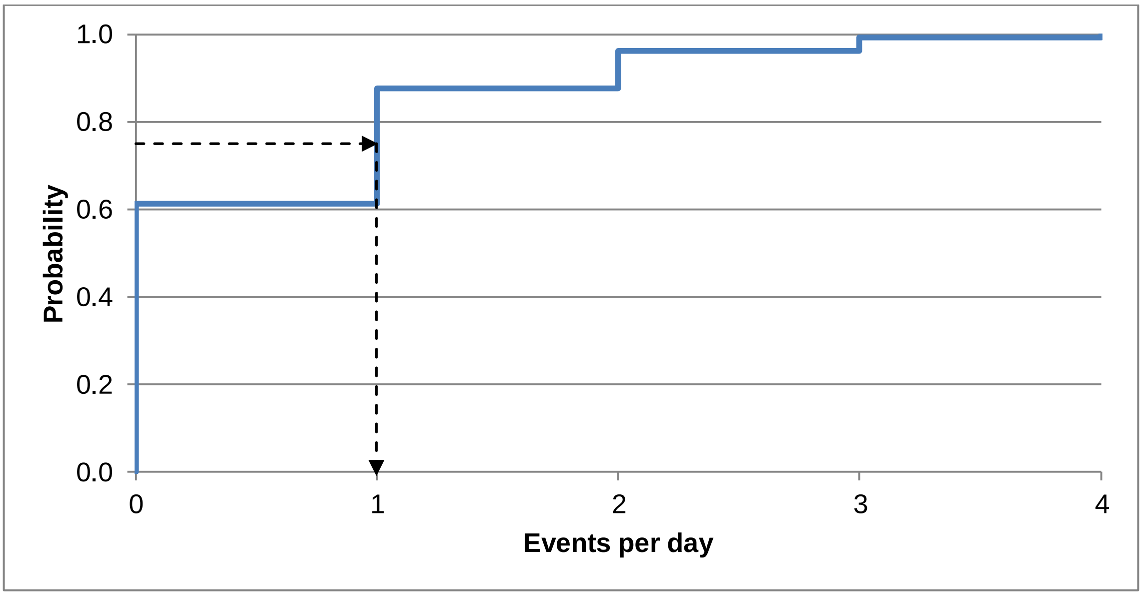


Figure 5.3: Cumulative probability distribution for 1 PPH bath end-use frequency

5.7.2. Dishwasher and Washing Machine Probability Distribution

The probability distributions for the daily frequency of dishwasher and washing machine events were determined, using the same method as discussed in section 5.7.1 for the other end-uses. The only difference was that within each dataset, only the first cycle in each event was considered. The first step therefore entailed sorting the data so that the use types entitled “Dishwasher1” or “Clotheswasher1” were grouped together, and the rest of the cycles were ignored. This ensured that the number of complete events per day was counted, regardless of the number of cycles per event. The number of cycles and the duration between cycles could be obtained from the signature patterns for these end-use events.

Blokker et al. (2010) assigned constant and specific patterns to dishwashers and washing machines in the SIMDEUM. In the SIMDEUM, dishwasher events consisted of 4 cycles, with a volume of 14 ℓ and duration of 84 s. Similarly, washing machine events consisted of 4 cycles, with a volume of 50 ℓ and duration of 5 minutes. During the inspection of the dishwasher and washing machine data used in this study, no clear pattern could be identified amongst the events. The number of cycles, duration of cycles, flow rates and volumes varied considerably, even within single households.

Since the logging data revealed relatively unique events, the dishwasher and washing machine events were not assigned fixed characteristics (in terms of the number of cycles, duration, and flow rate) in the end-use model, but also generated unique events each time those end-uses were used. This meant that for each dishwasher and washing machine event the number of cycles, the flow rate and duration for each cycle, as well as the duration between cycles, were specified, with probability distributions, to thus mimic the actual variation in these parameters as per the recorded data.

5.8. Number of Cycles Probability Distribution

For each household size category (1 PPH - 6 PPH) the number of cycles per dishwasher and washing machine event was counted. The events were sorted, based on their number of cycles, and then the subtotal command in Microsoft Excel was used. A subtotal was inserted each time a change in the “number of cycles” column changed, which then counted the number of occurrences within that category. The cumulative relative frequencies were then calculated to determine the corresponding PDFs (provided in Appendix B).

Figures 5.4 and 5.5 present the cumulative probability distributions for the washing machine and dishwasher number of cycle variable graphically. The REUWS data contained some dishwasher and washing machine events with up to 12 and 20 cycles respectively. Upon visual inspection it was not always possible to refute the large number of cycles by identifying where a single event should have been labelled as separate events. The recorded extreme events consisting of many cycles were therefore included in the model, because there was no scientific reason to exclude them (one possible explanation would be the repetitive use of an appliance, with different settings for each event).

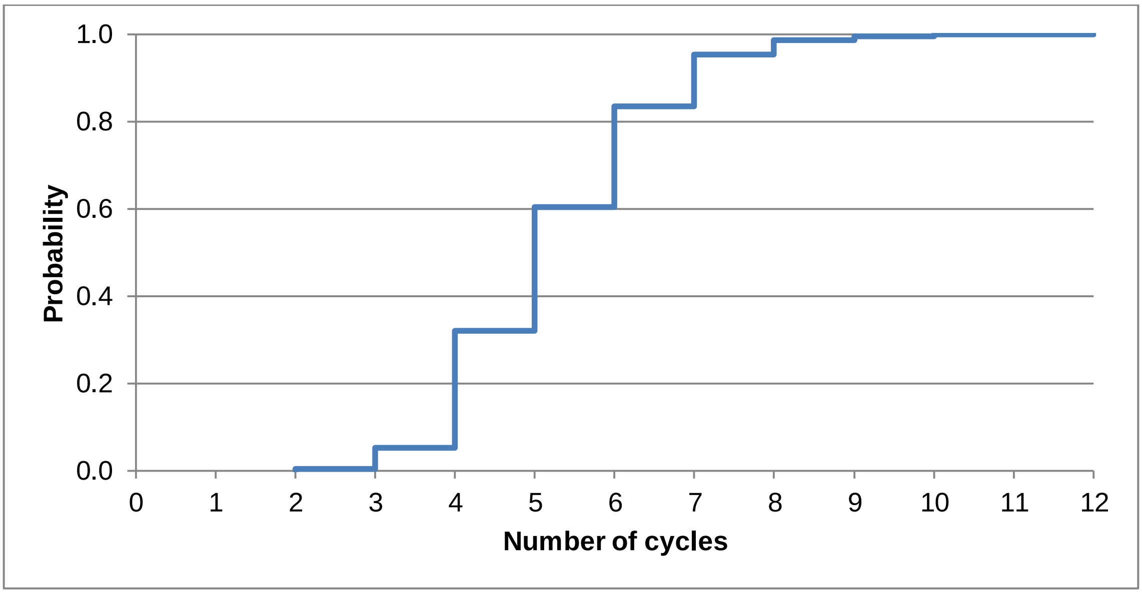


Figure 5.4: Cumulative probability distribution for dishwasher number of cycles

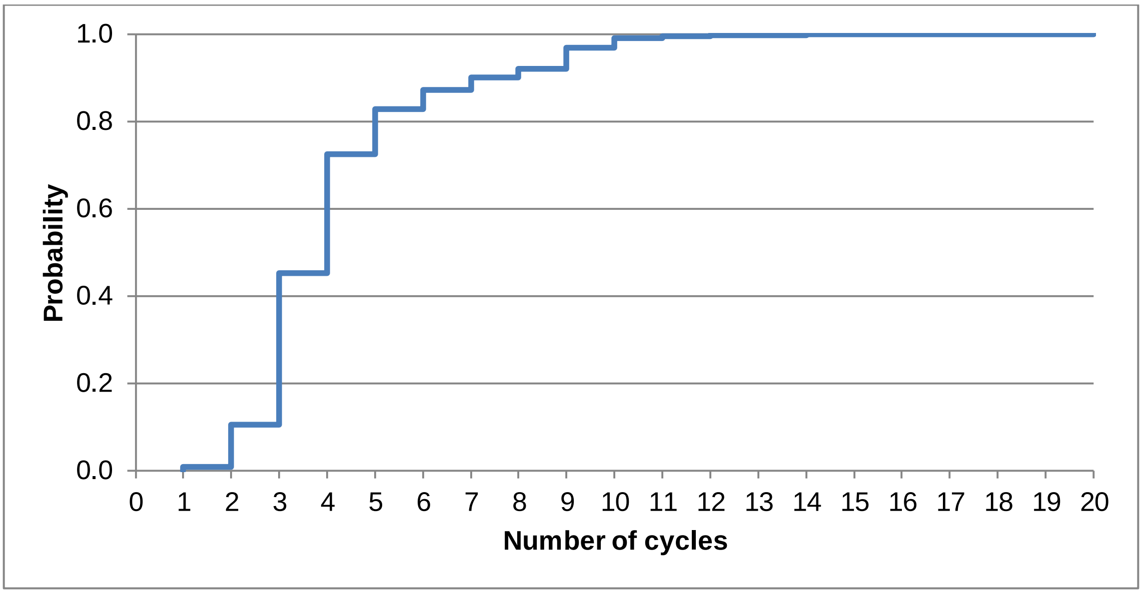


Figure 5.5: Cumulative probability distribution for washing machine number of cycles

5.9. Starting Hour Probability Distribution

The event starting times were obtained by selecting an hour of the day during which an event would take place according to the REUWS data, and then assigning a random minute and second within that hour. Any time interval could have been chosen instead of the 1 hour intervals, such as a 15 minute or 1 minute interval, to derive the probability distributions. Smaller time intervals would, however, have resulted in fewer data points being available in each category.

Since a limited number of data points were available from the REUWS, it was possible that smaller time intervals would skew the probability distribution and not be representative of the peak times that water is used in households. The hour categories allowed a clear probability trend to become visible. The exact minute and second within a certain hour that an event occurred was assumed to be random. Starting hour probability distributions were needed for the end-use model and were therefore derived.

For each end-use, the starting times of all recorded events were considered. An additional column was inserted in the data called "hour". In this column, the Microsoft Excel function with the expression "hour()" was used to evaluate the start time of each event and return the hour of the day as a number from 0 to 24. Where 0 represents 24h00 and 24 represents 12h00. The hour values were sorted in ascending order, and the subtotal function was used to determine the frequency of occurrence of each hour value. The cumulative relative frequencies (probabilities) representing the starting hour that were determined for each end-use is provided in Appendix B.

5.10. Goodness of Fit Tests

In the end-use model developed in this study, the volume, flow rate, and duration between cycles had to be determined for the relevant end-uses at the occurrence of each event. As discussed earlier, the REUWS database contained thousands of measured values for the above mentioned parameters. One possibility was to use the actual sample data directly as input variables in the end-use model. Such an approach would, however, necessitate the assumption that the volume, flow rate and duration variables are discrete, and the results would have been limited to the values within the sample. Alternatively, mathematical functions such as best fit trend lines could be fitted to the data, or the sample data could be fitted to theoretical probability distributions. In both cases the above mentioned parameters would be applied as continuous variables, and the model input variables could be mathematically described. For the purpose of this study, theoretical distribution functions were used. The disadvantage of the latter approach is that an additional error would have been introduced, due to slight mismatches between theoretical and actual distributions.

The input variables for an end-use model to estimate peak flows in a water distribution system have inherent uncertainties built into each parameter, with resulting error, which may even exceed the uncertainties, introduced by a slight mismatch of the theoretical distribution. The theoretical distributions did not yield statistically significant fits to the data in all cases. The “best fit” distributions were, however, still applied in the end-use model, since the sample dataset itself was not necessarily representative of all end-use volumes and flow rates, but merely provided a guide to a possible distribution of those parameters. The purpose of the theoretical distributions were, therefore, to yield possible volume and flow rate values for specific end-uses within a reasonable range, and not to replicate the REUWS dataset exactly. It would be advantageous if future work investigated the effect on the results of the end-use model if the sample data was used directly in the model, instead of fitted distributions.

For this research goodness of fit tests were used to determine which theoretical probability distributions provided the best fit to the sample data to the greatest degree. The @Risk software was used to apply the Chi-squared, A-D, and the K-S goodness of fit tests to seventeen different theoretical probability distributions and the data. For each of the three goodness of fit tests, @Risk ranked the seventeen distributions in ascending order, where a rank of 1 represented the best fit distribution, and 17 represented the worst.

Each goodness of fit test determined the best-fit distribution differently, and gave a larger weighting to different components of a distribution such as the tail or the centre range, as explained in section 3.7. Therefore, the resultant rank given by the three tests was not always the same, as could be expected.

For example, Table 5.11 shows the probability distribution rankings resulting from the three goodness of fit tests based on the shower flow rate data. The last three distributions were not ranked because the distributions were not applicable to the data.

Following the extensive literature review, no reference could be found to confirm that one test is preferred above the other. Milke et al. (2008) considered the results given by all three tests, and used a scoring system to determine the best-fit distribution of their data. A similar scoring system was used for this study.

Table 5.11: Goodness of test results for shower flow rate

Rank	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov
1	Log-Logistic	Log-Logistic	Log-Logistic
2	Pearson6	Pearson6	Erlang
3	Erlang	Erlang	Pearson6
4	Gamma	Gamma	Gamma
5	Log normal	Log normal	Log normal
6	Lognorm2	Lognorm2	Lognorm2
7	Weibull	Weibull	Weibull
8	Rayleigh	Rayleigh	Rayleigh
9	Inverse Gaussian	Exponential	Inverse Gaussian
10	Exponential	Triangle	Pearson5
11	Pearson5	Uniform	Exponential
12	Triangle	Inverse Gaussian	Chi-Squared
13	Uniform	Pearson5	Triangle
14	Chi-Squared	Chi-Squared	Uniform
-	Beta General	Beta General	Beta General
-	Pareto	Pareto	Pareto
-	Pareto2	Pareto2	Pareto2

The scoring system entailed using the ranking value as a proxy for score, and the sum of the three rankings then provided a total score per distribution. The total scores of all the distributions were then compared with each other, which allowed an overall placing to be determined, so as to select the "best" distribution in each case. By treating the results in this manner, it was ensured that all three tests were given equal emphasis.

Rearranging the results contained in Table 5.11 produces Table 5.12, which illustrates the scoring system with the distributions sorted alphabetically by name. The shaded row indicates the one with the best fit based on the "weight of all three tests".

Table 5.12: Goodness of test results for shower flow rate

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	-	-	-	-	-
Chi Squared	14	14	12	40	13
Erlang	3	3	2	8	3
Exponential	10	9	11	30	9
Gamma	4	4	4	12	4
Inverse Gaussian	9	12	9	30	9
Log Logistic	1	1	1	3	1
Log Normal	5	5	5	15	5
Log Normal2	6	6	6	18	6
Pareto	-	-	-	-	-
Pareto2	-	-	-	-	-
Pearson5	11	13	10	34	10
Pearson6	2	2	3	7	2
Rayleigh	8	8	8	24	8
Triangle	12	10	13	35	11
Uniform	13	11	14	38	12
Weibull	7	7	7	21	7

Since the Log-logistic distribution was ranked first in all three tests in this case, it had an overall score of 3 (1+1+1), which was the lowest score overall. The Log-Logistic distribution would therefore have been selected as the probability distribution for describing the shower volumes.

Data was available for the peak flow rate, as well as the mode flow rate. The peak flow represented the maximum flow rate (averaged over a 10 second interval) measured during the event, and the mode flow rate represented the flow rate that was recorded most often for any particular type of event. When the volume was divided by the duration, the calculated flow rates were closer to the mode flow rates than to the recorded peak flow rate. The mode flow rate logged data was therefore used to determine best fit distributions for the event flow rates. This was not considered to have a notable impact on the ultimate results and the decision was non-critical in terms of the research findings.

The comprehensive set of goodness of fit ranking results for all of the end-use model parameters are provided in Appendix A. The following sections describe the selection of the probability distribution functions for the end-use event volumes, flow rates and duration between cycles.

5.10.1. Shower Volume Probability Distribution Function

The Log-Logistic distribution was selected as the best fit distribution for shower volume since the Log-Logistic distribution ranked first in all three GOF tests, and ranked first overall. A total of 50 286 shower volume data points were used in the GOF test. The CDF of the fitted Log-Logistic distribution is graphically presented in Figure 5.6 together with the shower volume data. The graph shows that there is some variation between the lowest and highest values; however, in the centre region there is an excellent fit.

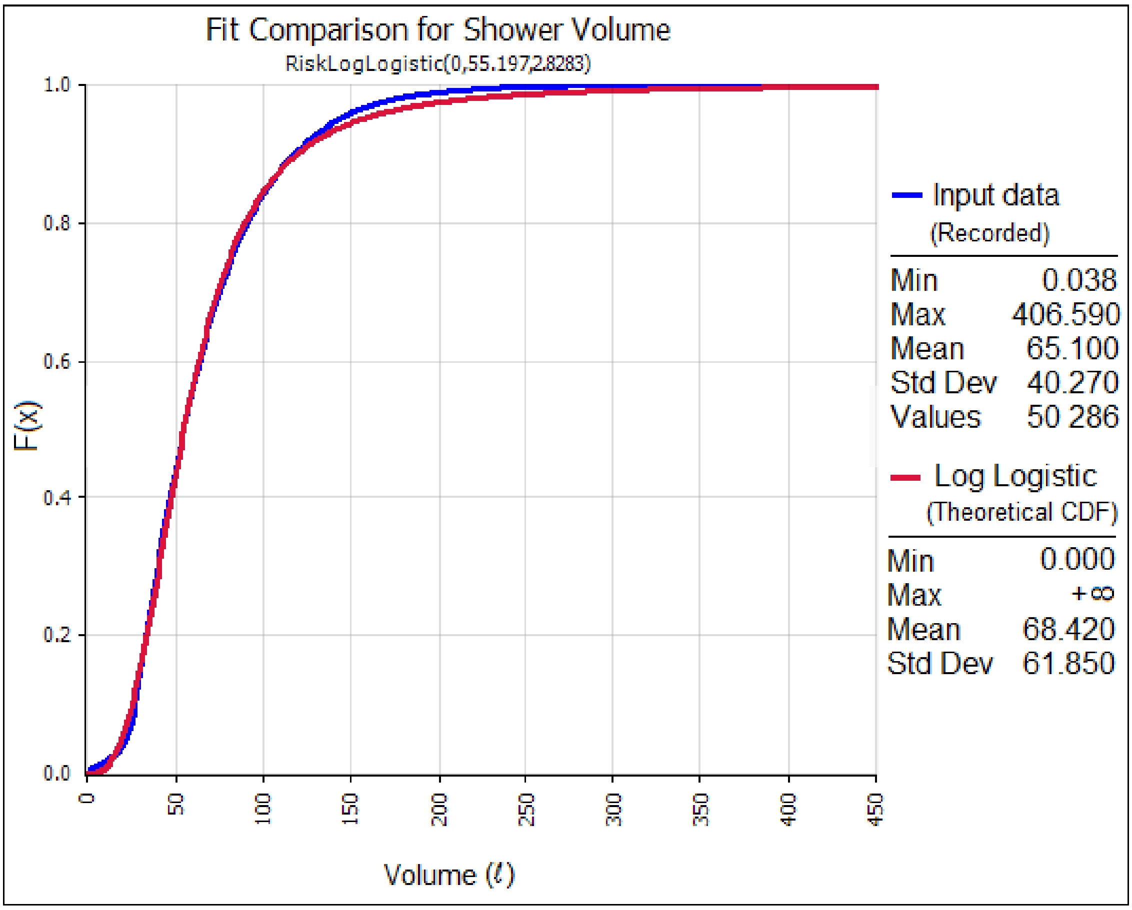


Figure 5.6: CDF plot of Log-Logistic distribution and shower volume data

5.10.2. Shower Flow Rate Probability Distribution Function

The Log-Logistic distribution was selected as the best fit distribution for shower flow rate since the Log-Logistic distribution ranked first overall based on the weight of all three GOF tests.

A total of 50 286 shower flow rate data points were used in the GOF test. The shower flow rate data and the CDF of the fitted Log-Logistic distribution is graphically presented in Figure 5.7. The graph shows that there is a good fit for most of the data range.

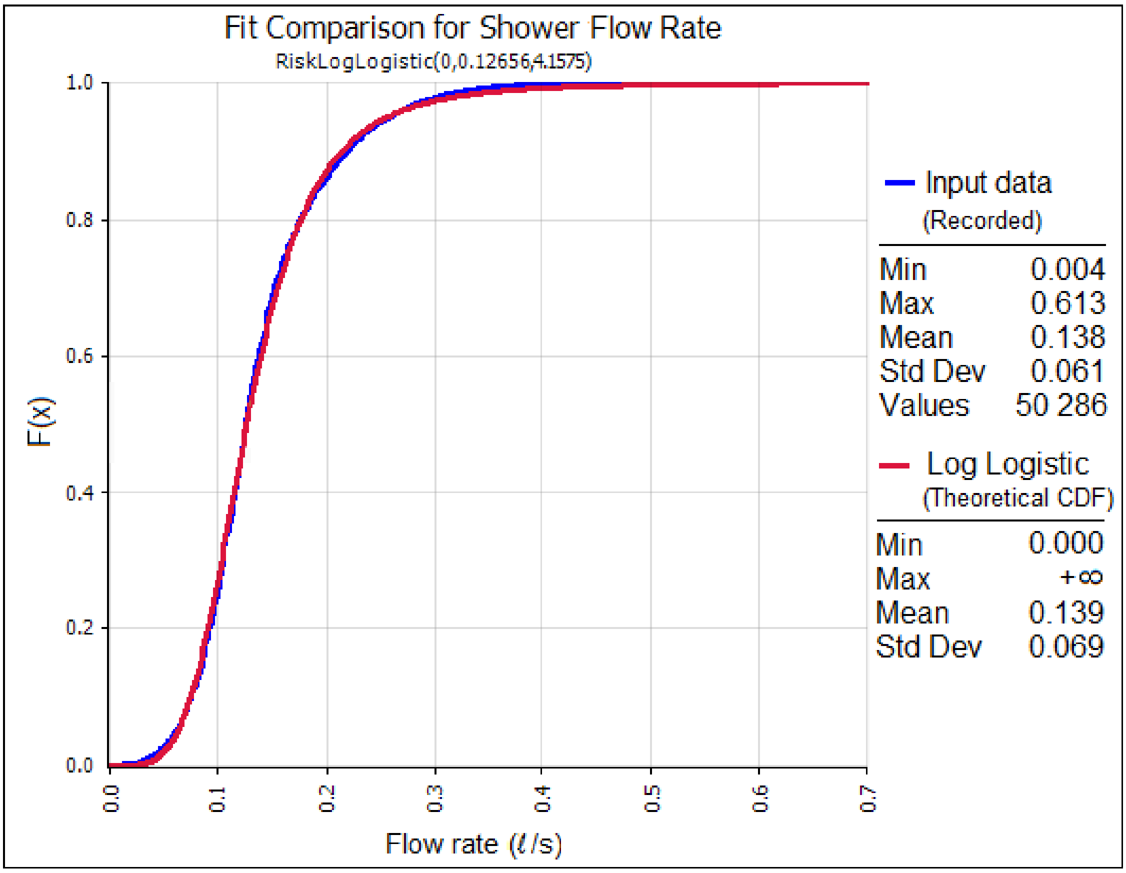


Figure 5.7: CDF plot of Log-Logistic distribution and shower flow rate data

5.10.3. Bath Volume Probability Distribution Function

The Rayleigh distribution was selected as the best fit distribution for bath volume since the Rayleigh distribution ranked first in all three GOF tests, and ranked first overall.

A total of 4 105 bath volume data points were used in the GOF test. The CDF of the fitted Rayleigh distribution is graphically presented in Figure 5.8 together with the bath volume data. The Rayleigh distribution does not fit the data for bath volume as well as the earlier graphs fit the corresponding data, but it is a better fit than all the other distributions.

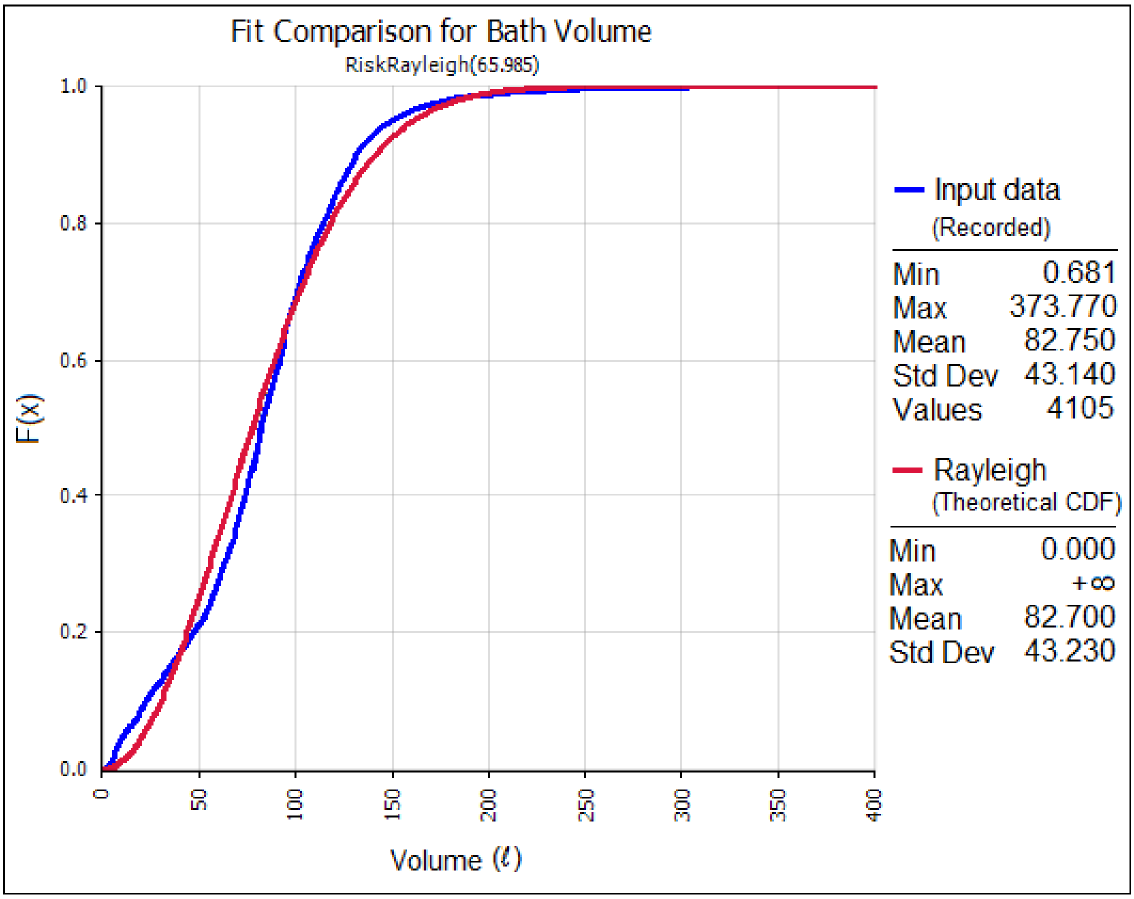


Figure 5.8: CDF plot of Rayleigh distribution and bath volume data

5.10.4. Bath Flow Rate Probability Distribution Function

The Weibull distribution was selected as the best fit distribution for bath flow rate since the Weibull distribution was ranked first overall based on the weight of all three GOF tests.

A total of 4 105 bath flow rate data points were used in the goodness of fit test. The bath flow rate data and the CDF of the fitted Weibull distribution is presented is graphically presented in Figure 5.9. This graph shows that the

Weibull distribution is not a perfect fit; however, it is a better fit than any of the other distributions.

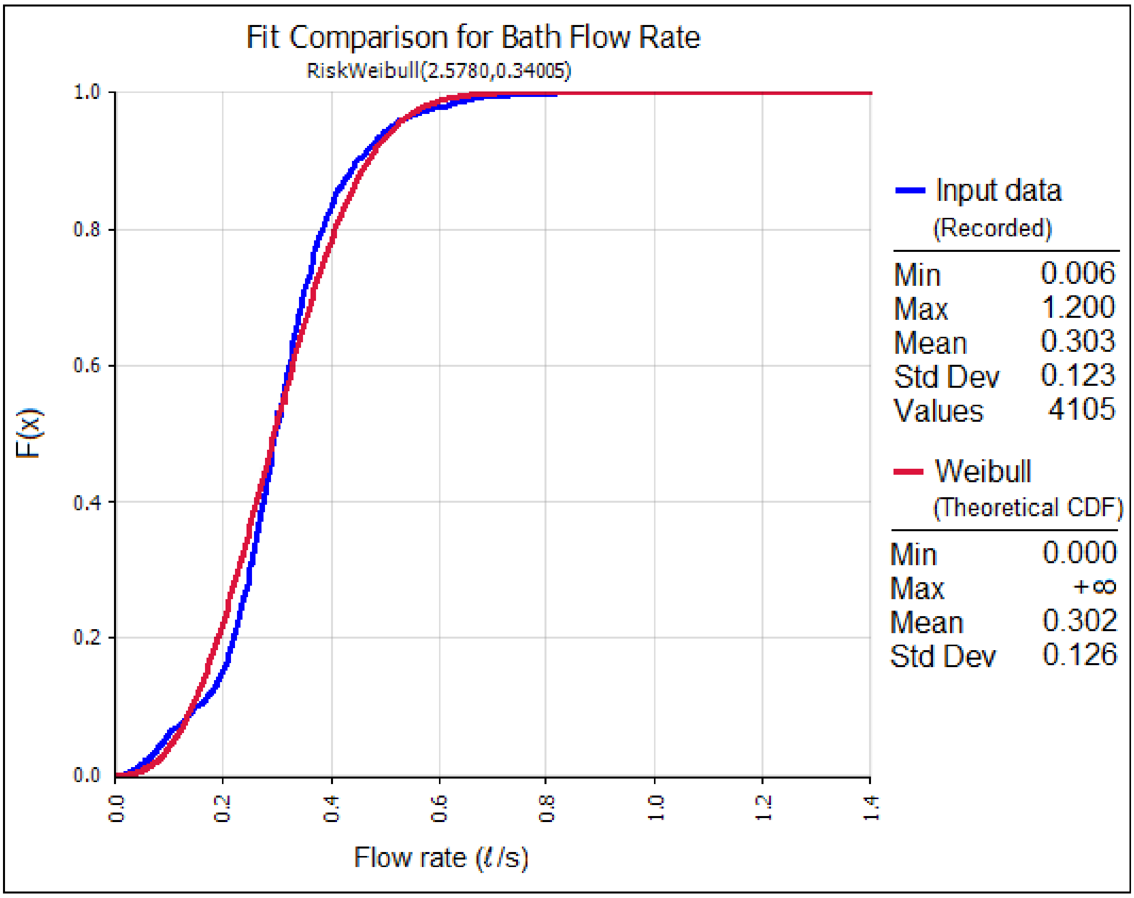


Figure 5.9: CDF plot of Weibull distribution and bath flow rate data

5.10.5. Toilet Volume Probability Distribution Function

The Weibull distribution was selected as the best fit distribution for toilet volume since the Weibull distribution ranked first in all three GOF tests, and ranked first overall.

A total of 289 477 toilet volume data points were used in the GOF test. The CDF of the fitted Weibull distribution is graphically presented in Figure 5.10 together with the toilet volume data. This graph shows that there is some variation in the lower range of values; however, for the rest of the range there is a reasonably good fit.

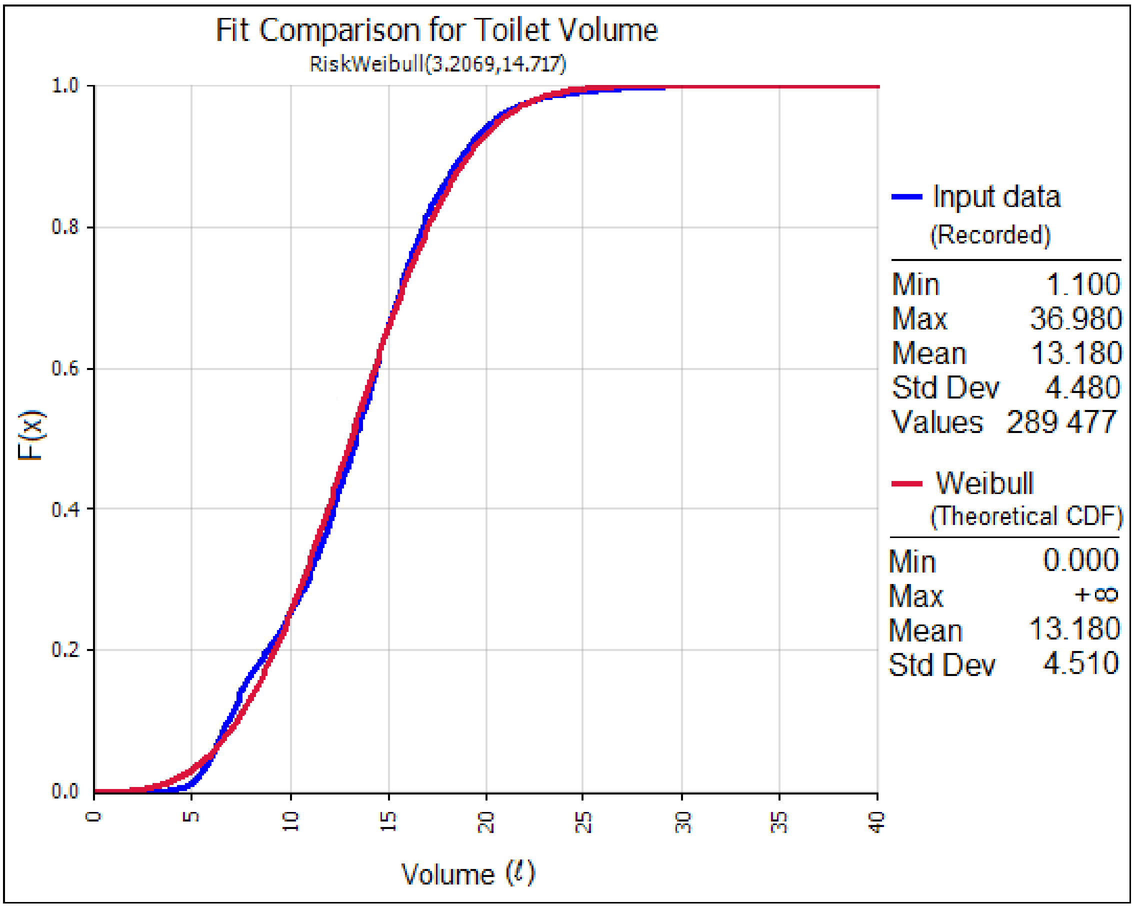


Figure 5.10: CDF plot of Weibull distribution and toilet volume data

5.10.6. Toilet Flow Rate Probability Distribution Function

A total of 289 477 toilet flow rate data points were used in the GOF tests. The Gamma distribution was ranked first by the A-D test and third by the Chi-squared and K-S tests, thus it achieved an overall top rank on scores. The Weibull distribution was ranked first by both the Chi-squared and K-S tests, but the A-D test ranked the Weibull distribution only twelfth, resulting in the Weibull distribution having a poor overall score of four.

When inspecting the fitted CDF plots for the Gamma distribution, shown in Figure 5.11, and the Weibull distribution shown in Figure 5.12, however, the Weibull distribution appears to provide a better fit. A subjective judgement was made in this case to use the Weibull distribution function as the best fit distribution for the toilet flow rate.

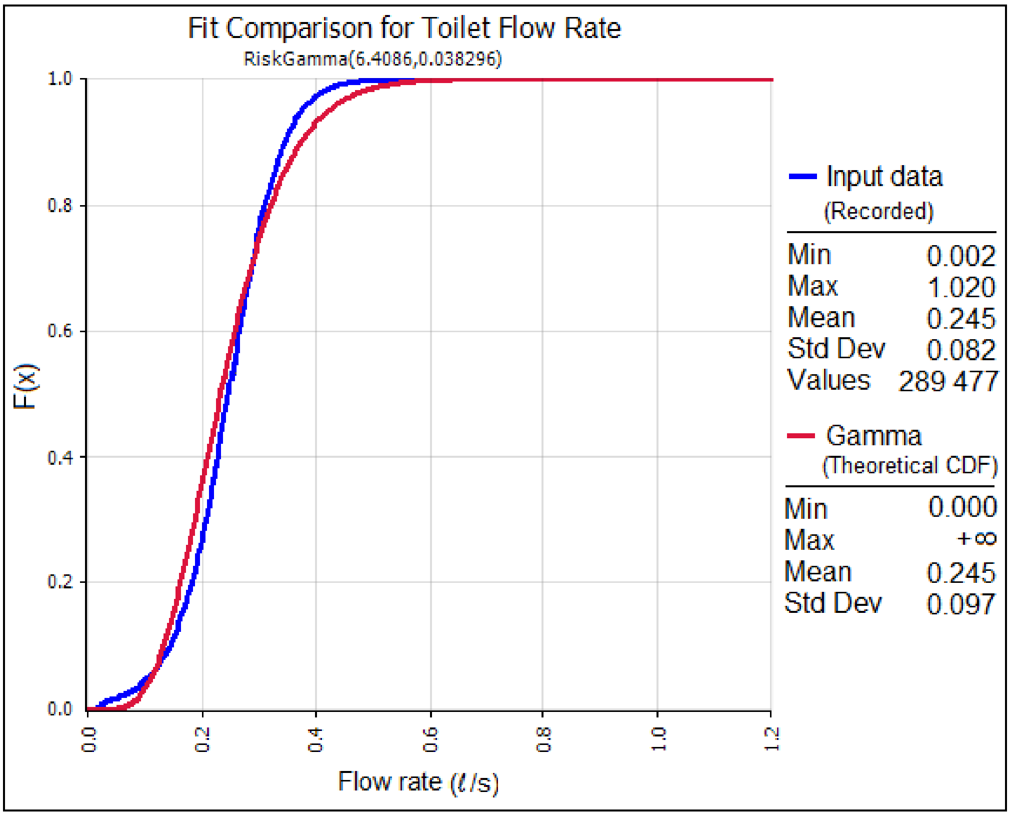


Figure 5.11: CDF plot of Gamma distribution and toilet flow rate data

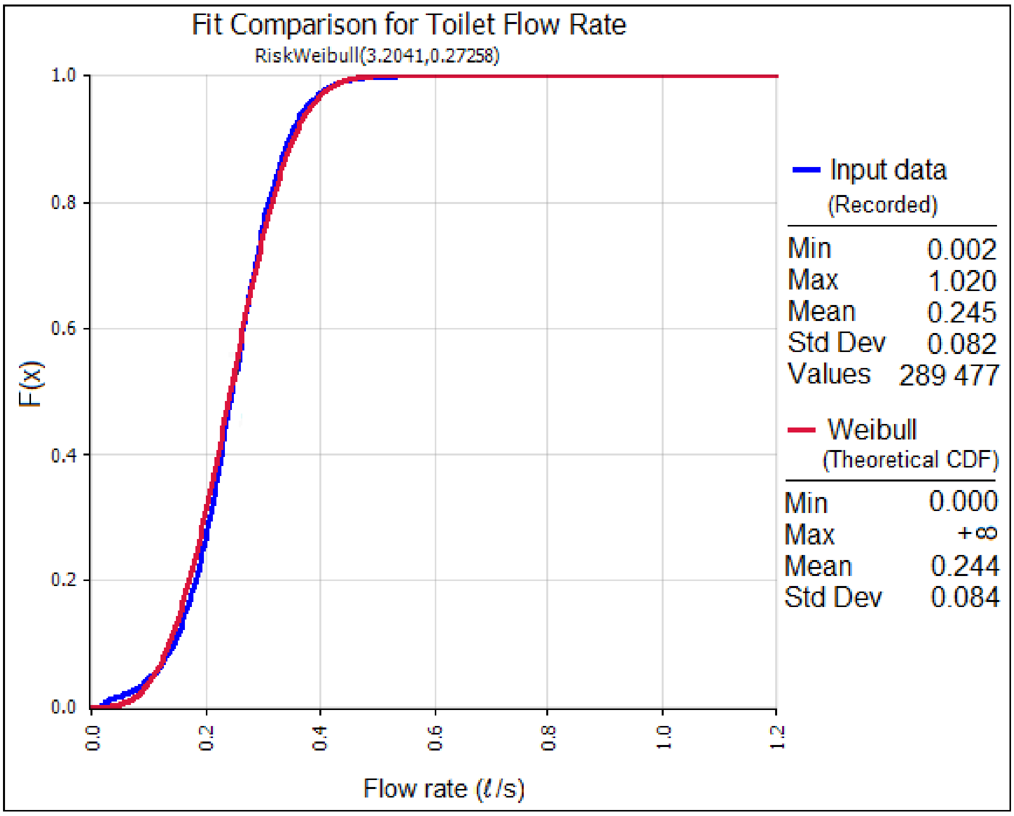


Figure 5.12: CDF plot of Weibull distribution and toilet flow rate data

5.10.7. Tap Volume Probability Distribution Function

A total of 1 150 583 tap volume data points were used in the GOF tests. The Inverse Gaussian, Pearson5 and Pearson6 distributions were the top three overall ranked distributions, closely followed by the much more common Log Normal distribution. The CDF of the first three distributions were mathematically complex and could not be represented by a Microsoft Excel equation or function. The Log Normal distribution showed a good fit as well, being ranked fourth. The Log Normal distribution was therefore selected to describe the tap volume. The CDF of the Log Normal distribution is graphically presented in Figure 5.13 together with the tap volume data.

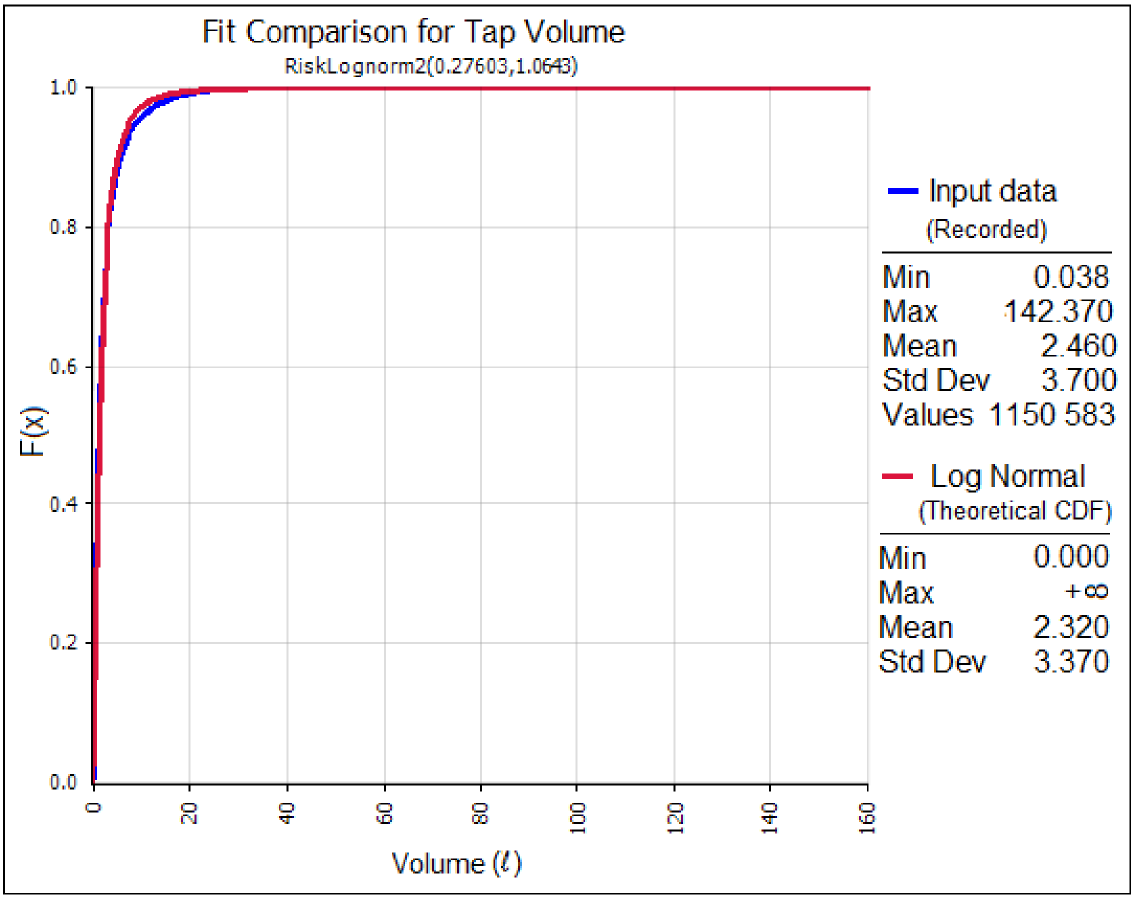


Figure 5.13: CDF plot of Log Normal distribution and tap volume data

5.10.8. Tap Flow Rate Probability Distribution Function

The Gamma distribution was selected as the best fit distribution for tap flow rate since the Gamma distribution was ranked first overall based on the weight of all three GOF tests.

A total of 1 150 583 tap flow rate data points were used in the GOF tests. The the tap flow rate data and the CDF of the fitted Gamma distribution is graphically presented in Figure 5.14. The graph shows that the Gamma distribution fits the data very well.

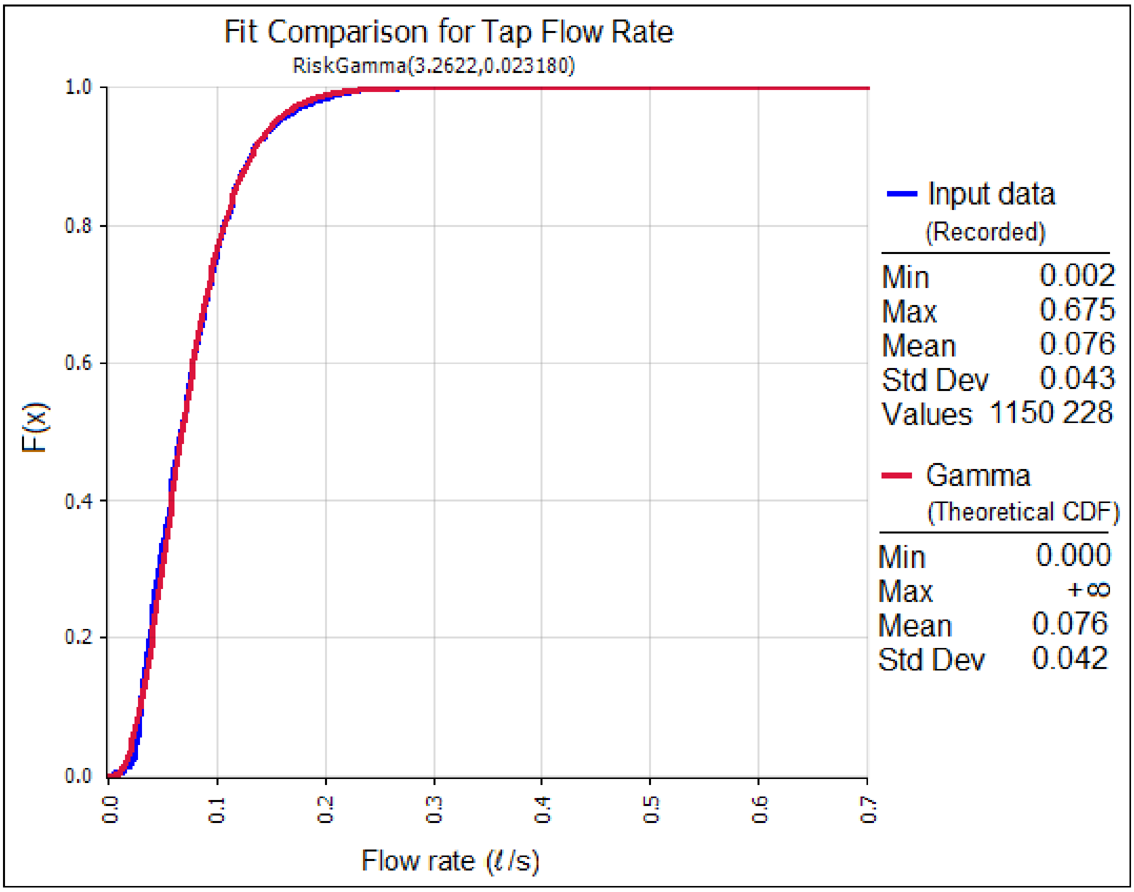


Figure 5.14: CDF plot of Gamma distribution and tap flow rate data

5.10.9. Dishwasher Cycle Volume Probability Distribution Function

The Log-Logistic distribution was selected as the best fit distribution for dishwasher cycle volume since the Log-Logistic distribution ranked first in all three goodness of fit tests, and ranked first overall.

A total of 33 652 dishwasher cycle volume data points were used in the GOF tests. The CDF of the fitted Log-Logistic distribution and the dishwasher cycle volume data is graphically presented in Figure 5.15.

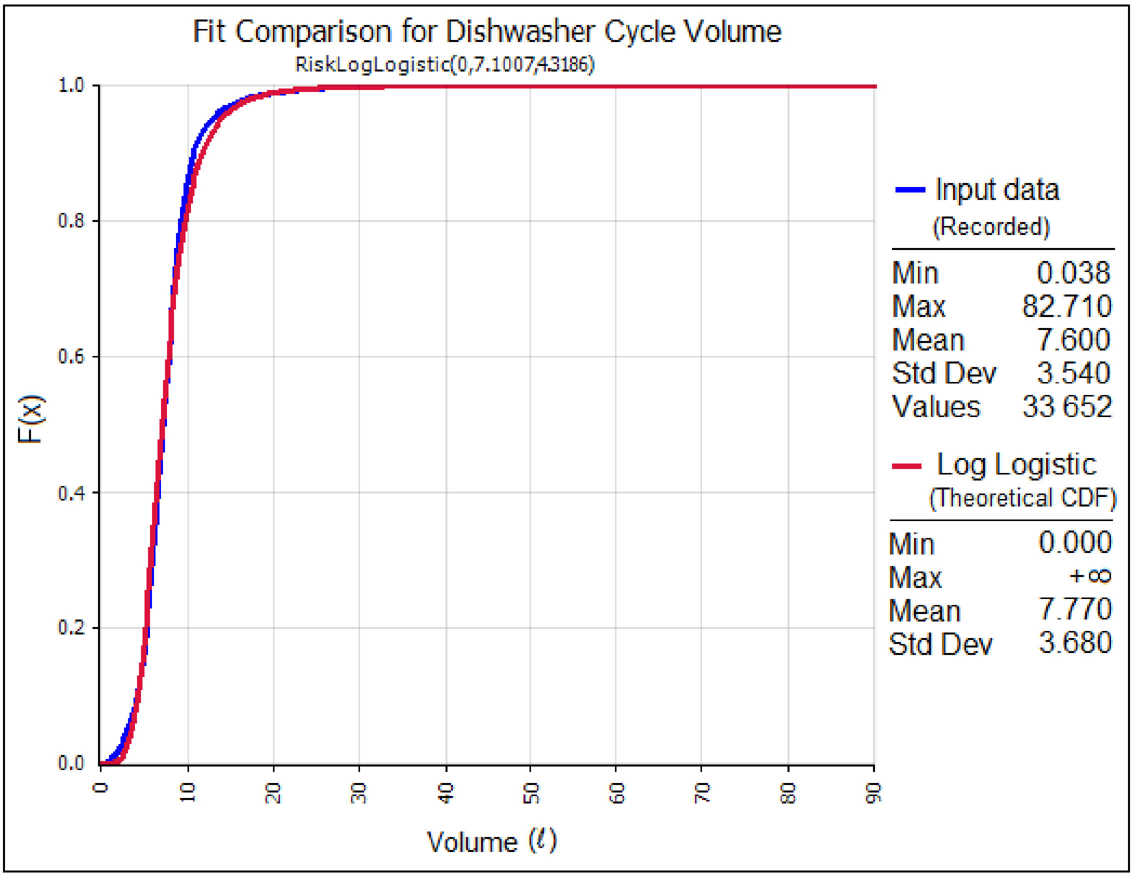


Figure 5.15: CDF plot of Log-Logistic distribution and dishwasher volume data

5.10.10. Dishwasher Cycle Flow Rate Probability Distribution Function

The Erlang distribution was selected as the best fit distribution for dishwasher cycle flow rate since the Erlang distribution was ranked first overall based on the weight of all three GOF tests.

A total of 33 652 dishwasher cycle flow rate data points were used in the GOF tests. The CDF of the fitted Erlang distribution and the dishwasher cycle flow rate data is graphically presented in Figure 5.15. The graph shows that there is some variation between the higher ranges of values; however, for the rest of the values there is a reasonably good fit.

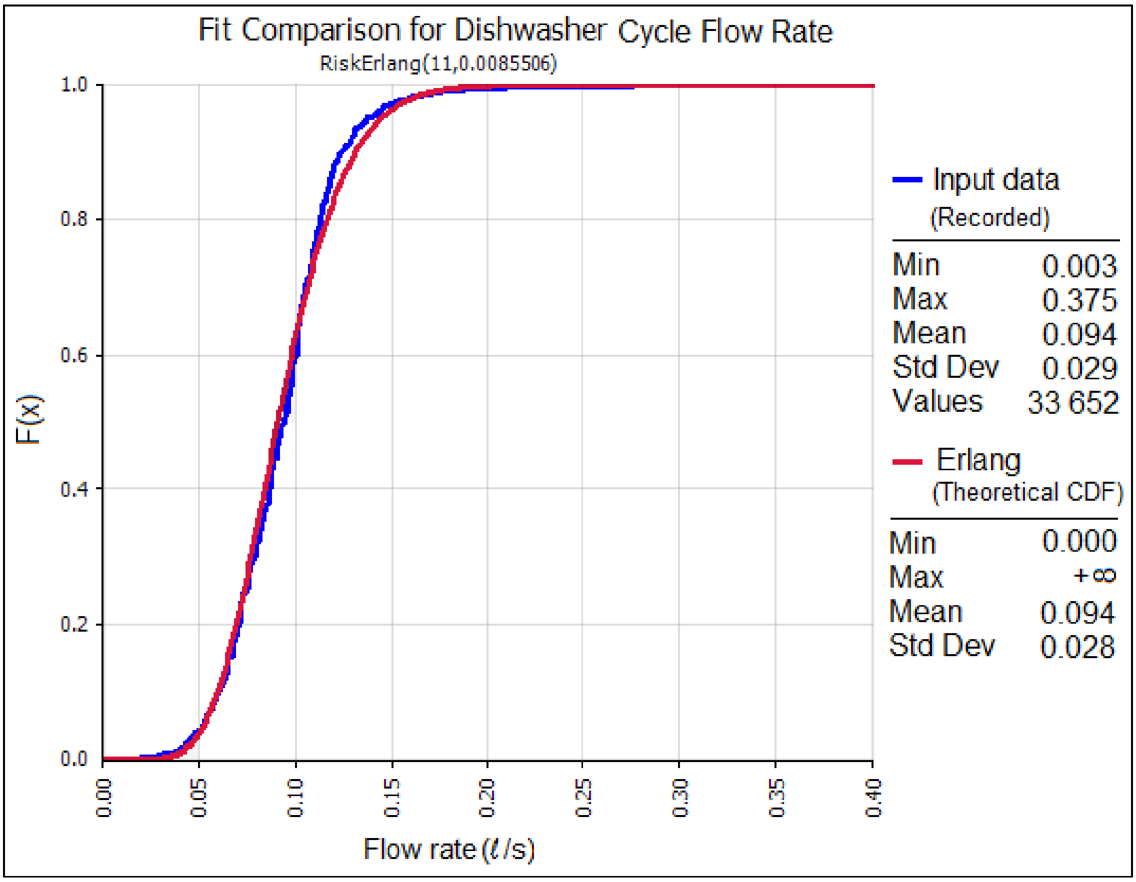


Figure 5.16: CDF plot of Erlang distribution and dishwasher flow rate data

5.10.11. Dishwasher Duration Between Cycles Probability Distribution Function

A total of 26 827 dishwasher duration between cycles data points were used in the GOF tests. The Pearson6 distribution ranked top overall, followed by the Log-Logistic distribution. The Log-Logistic distribution was selected in favour of the first, for the same reason given in section 5.10.7. The dishwasher duration between cycles data and the CDF of the fitted Log-Logistic distribution is graphically presented in Figure 5.17. The graph shows a relatively good fit.

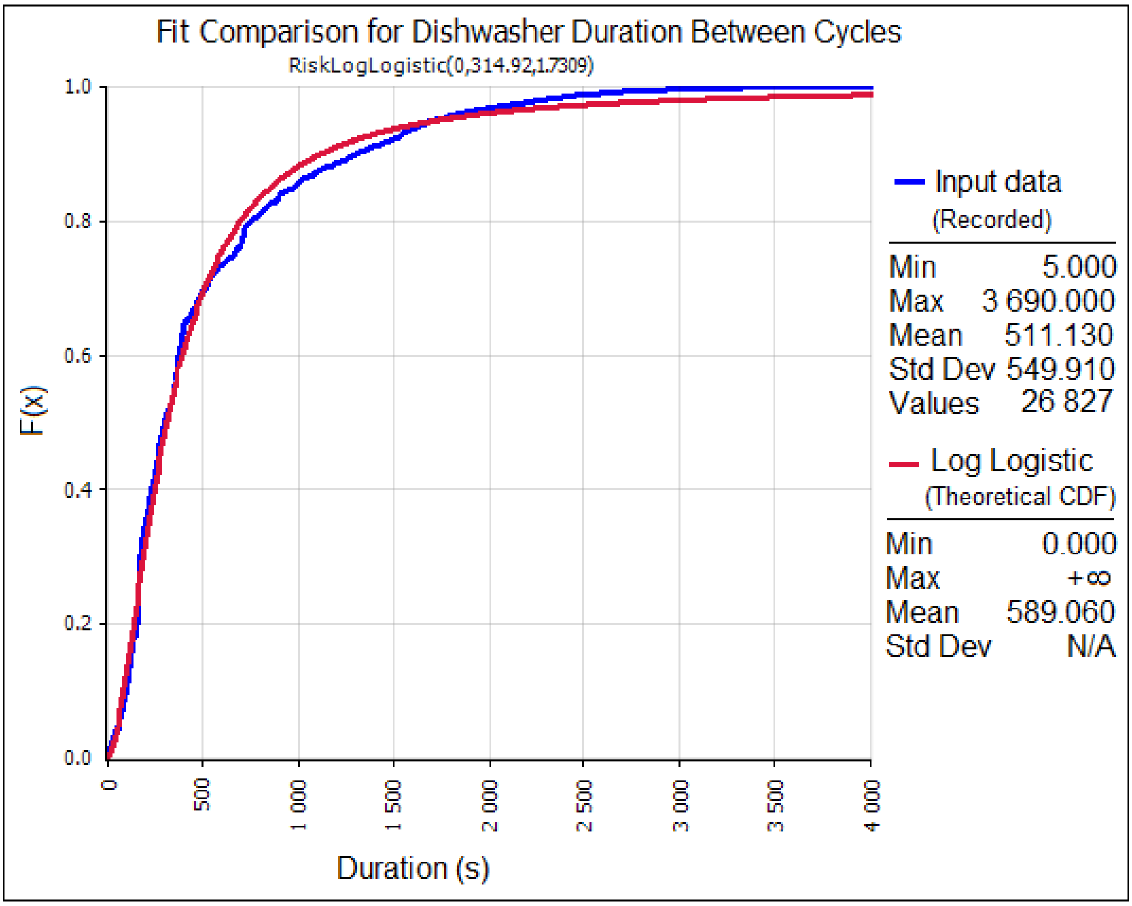


Figure 5.17: CDF plot of Log-Logistic distribution and dishwasher duration between cycles data

5.10.12. Washing Machine Cycle Volume Probability Distribution Function

The Weibull distribution was selected as the best fit distribution for washing machine cycle volume since the Weibull distribution was ranked first overall based on the weight of all three GOF tests.

A total of 114 887 washing machine cycle volume data points were used in the goodness of fit test. The CDF of the fitted Weibull distribution is graphically presented in Figure 5.18 together with the washing machine cycle volume data. The graph illustrates a peculiarity of the washing machine volume data, in the sense that it does not form a smooth curve. This may be ascribed in part to the fact that washing machines and the relevant manufacturer's settings are linked to various pre-defined volumes.

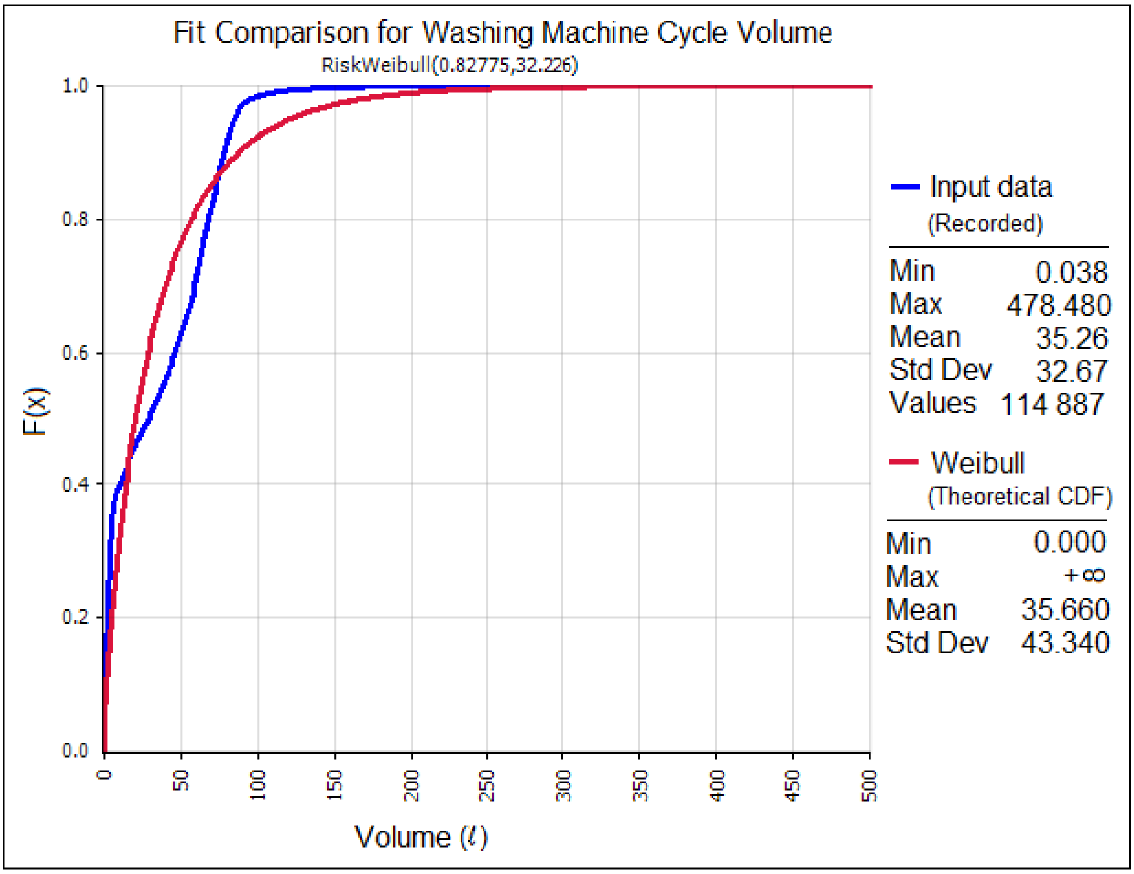


Figure 5.18: CDF plot of Weibull distribution and washing machine volume data

5.10.13. Washing Machine Cycle Flow Rate Probability Distribution Function

The Weibull distribution was selected as the best fit distribution for washing machine cycle flow rate since the Weibull distribution was ranked first overall based on the weight of all three GOF tests.

A total of 114 887 washing machine cycle flow rate data points were used in the GOF tests. The CDF of the fitted Weibull distribution and the washing machine cycle flow rate data is graphically presented in Figure 5.19. The graph shows that the Weibull distribution is a reasonably good fit.

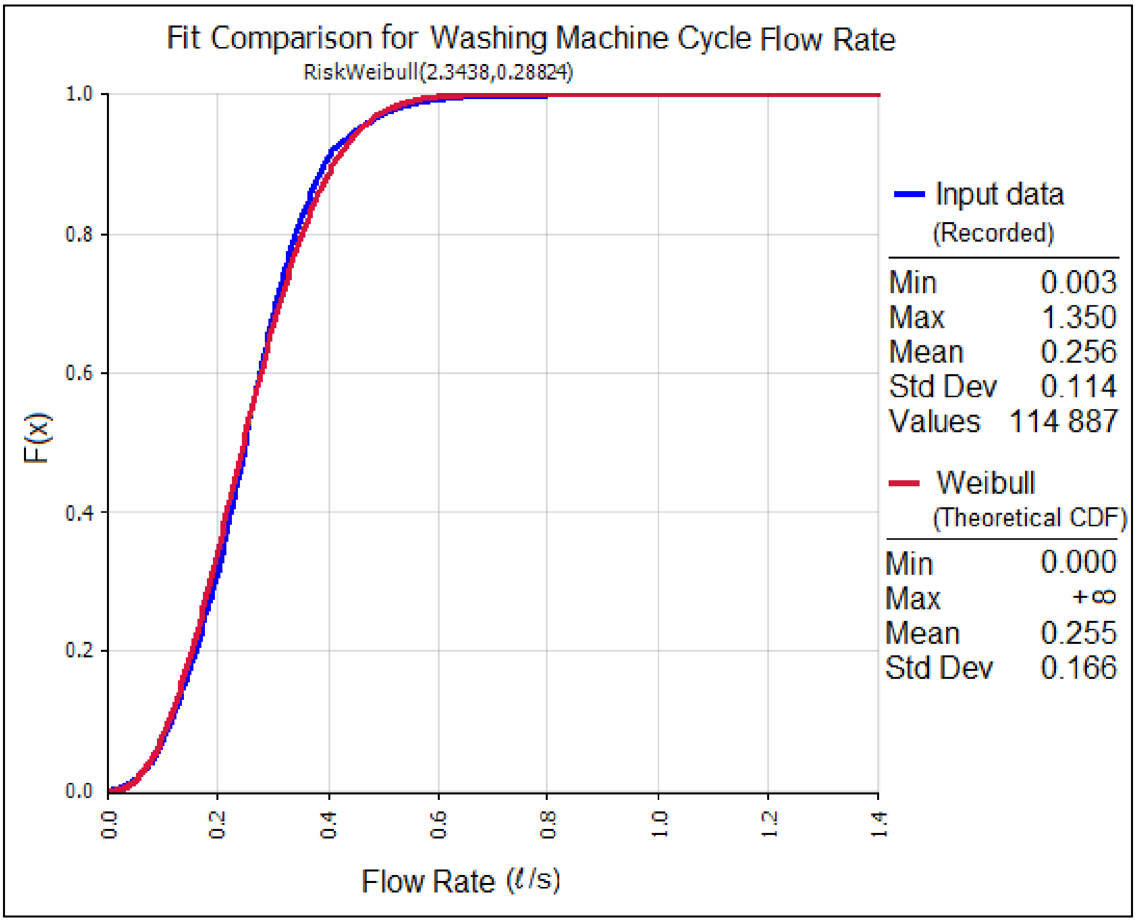


Figure 5.19: CDF plot of Weibull distribution and washing machine cycle flow rate data

5.10.14. Washing Machine Duration Between Cycles Probability Distribution Function

The Beta General distribution was selected as the best fit distribution for washing machine duration between cycles since the Beta General distribution was ranked first overall based on the weight of all three GOF tests.

A total of 86 785 washing machine duration between cycles data points were used in the GOF tests. The CDF of the fitted Beta General distribution and washing machine duration between cycles data is graphically presented in Figure 5.17. This graph shows that there is some variation between the values, but overall it is a reasonable fit.

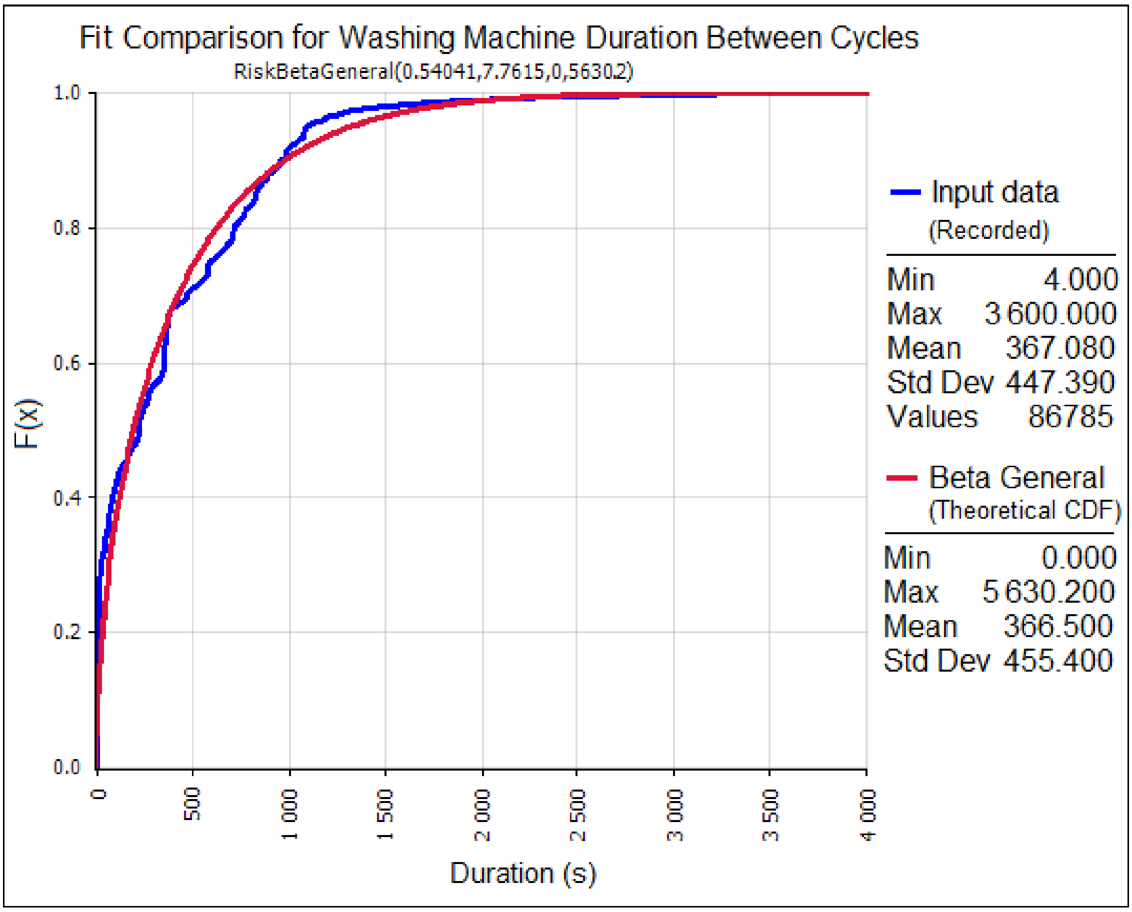


Figure 5.20: CDF plot of Beta General distribution and washing machine duration between cycles data

5.10.15. Distribution Parameters

A summary of the selected distributions and the values of the distribution parameters applied in the end-use model is presented in Tables 5.26 to 5.28.

Table 5.13: Event volume distributions

End-use	Distribution	Parameter	Parameter Value
Shower	Log-Logistic	γ	0.000
		β	55.197
		α	2.828
Bath	Rayleigh	β	65.985
Toilet	Weibull	α	3.207
		β	14.717

End-use	Distribution	Parameter	Parameter Value
Tap	Log Normal	β	0.276
		α	1.064
Dishwasher	Log-Logistic	γ	0.000
		β	7.101
		α	4.319
Washing machine	Weibull	α	0.823
		β	32.226

Table 5.14: Event flow rate distributions

End-use	Distribution	Parameter	Parameter Value
Shower	Log-Logistic	γ	0.000
		β	0.127
		α	4.158
Bath	Weibull	α	2.578
		β	0.340
Toilet	Weibull	α	3.204
		β	0.273
Tap	Gamma	α	3.262
		β	0.023
Dishwasher	Erlang	α	11.000
		β	0.009
Washing machine	Weibull	α	2.344
		β	0.288

Table 5.15: Event duration between cycles distributions

End-use	Distribution	Parameter	Parameter Value
Dishwasher	Log-Logistic	γ	0.000
		β	314.920
		α	1.731
Washing machine	Beta general	α_1	0.540
		α_2	7.762
		min	0.000
		max	5630.200

5.11. End-Use Model Construction

The construction of an electronic version of the model in Microsoft Excel was part of this research work and elaboration on its construction was considered essential in order to ensure that the work is repeatable. A detailed step-by-step explanation is thus presented in this section of how the model was compiled. The various tables presented in section 5.11 present summaries of the equations applied in Microsoft Excel in order to achieve the results of the end-use model. Reading section 5.11 is not crucial to those readers who would merely like to follow the research methodology and logic presented in order to derive the results.

The end-use model was fully constructed by means of tables and equations in a single Microsoft Excel workbook consisting of eight worksheets. The worksheets were entitled Household size, Shower, Bath, Toilet, Tap, Dishwasher, Washing machine, and Flow per house. The six end-use worksheets had an identical layout, except for the additional cycle requirements for the dishwasher and washing machine. The following sections provide an explanation of the information contained in each of the worksheets.

5.11.1. Household Size Worksheet

The purpose of the household size worksheet was to select the number of persons in the household so that the appropriate number of events distribution was applied in subsequent steps.

The cumulative relative frequency distribution table for household size was inserted in the household size worksheet range A1:C10. The relative frequency and cumulative relative frequency columns were shifted down one unit so that the equation in cell B11 referred to the appropriate values as shown in Table 5.16. The calculation procedure was as follows: After generating a random number in Microsoft Excel, the “lookup()” function was used to search for the random number (the lookup value) in the cumulative relative frequency column (lookup vector) and return the value in the PPH column (lookup range). If the lookup function could not find the lookup value, the function matched the largest value in the lookup vector that was less than or equal to the lookup value. For example, if the random number was 0.5, then a household size of 3 was selected and displayed in cell B11.

Table 5.16: Household size worksheet calculation

	A	B	C
1	Household size calculation		
2	Household size (PPH)	Relative Frequency (Probability)	Cumulative Relative Frequency
3			
4	1	0.000	0.00
5	2	0.129	0.129
6	3	0.352	0.482
7	4	0.195	0.676
8	5	0.184	0.86
9	6	0.092	0.952
10		0.048	1.00
11	Selected PPH	=LOOKUP(RAND(),C4:C10,A4:A10)	

5.11.2. End-Uses Worksheets

Each end-use was represented by its own worksheet, which contained the relevant data corresponding to the end-use. The shower and dishwasher end-uses are used for purposes of illustration, however the bath, toilet, and tap end-uses followed a similar process as the shower, and the washing machine was similar to the dishwasher.

The daily event frequency was firstly calculated using the cumulative relative frequencies as shown in Tables B1.1 to B1.6 in Appendix B. Each household size category had a cumulative frequency distribution in adjacent columns, with the leftmost column representing the possible number of events per day. An “if()” statement was used in combination with the “lookup()” function that was applied, similarly to the household size calculation procedure described in section 5.11.1. If, for example, the household size in the household size worksheet was 3, then the 3 PPH cumulative relative frequency column was selected as the lookup vector. The household size value therefore determined which one of the six possible cumulative relative frequency columns the lookup function used to match a random number, and return a daily event frequency value.

The daily starting hour cumulative relative frequency table unique to each end-use was also available in the worksheets for reference purposes (see Table B2.1 in Appendix B). The next step in the model was to determine the starting time of each event, for which a new table was created, as shown in Table 5.17. In the event number column, a number was displayed only if the daily event frequency value (obtained in cell C23) was greater than one or greater than the previous event number. If an event number was present, then the lookup function matched a random number to the daily starting hour cumulative relative frequency table. The cumulative relative frequency (lookup vector) was situated in range R4:R28, while the possible 24 hours (result vector) was situated in range P4:P28. The resulting hour value was used as the starting hour, while a random number was generated to establish the minutes and seconds within that hour that the event would start. The event numbers did not dictate the order in which events occurred, as the starting times of events were sorted in ascending order at a later stage. The equations used in Microsoft Excel to achieve the above results are shown in Table 5.18.

Table 5.17: Event starting time example

	T	U	V
1	Event starting time		
2	Event number	Hour	Starting time
3			
4	1	6	06:20:31 AM
5	2	11	11:14:17 AM
6	3	7	07:11:55 AM

Table 5.18: Event starting time equations

	T	U	V
1	Event starting time		
2	Event number	Hour	Starting time
3			
4	=IF($C$23>0,1,"")	=IF(T4="","",LOOKUP(RAND(), $R$4:$R$28,$P$4:$P$28))	=IF(T4="","",(U4/24)+ (RAND()/60))
5	=IF($C$23>T4,T4+1,"")	=IF(T5="","",LOOKUP(RAND(), $R$4:$R$28,$P$4:$P$28))	=IF(T5="","",(U5/24)+ (RAND()/60))
6	=IF($C$23>T4,T4+1,"")	=IF(T5="","",LOOKUP(RAND(), $R$4:$R$28,$P$4:$P$28))	=IF(T5="","",(U5/24)+ (RAND()/60))
7	=IF($C$23>T4,T4+1,"")	=IF(T5="","",LOOKUP(RAND(), $R$4:$R$28,$P$4:$P$28))	=IF(T5="","",(U5/24)+ (RAND()/60))

The dishwasher and washing machine made use of the same procedure as described above, except that an additional column entitled “number of cycles” was added to the event starting time table. The new column referred to the number of cycles cumulative relative frequency table as given in Table B2.2, in Appendix B. If an event number was present, then the lookup function was used to match a random number to the cumulative relative frequency, and return a number of cycles value. Due to new random numbers being generated for each event, it was possible for consecutive events to have different numbers of cycles.

The flow rate and volume of each event was determined as part of a new table entitled “Event characteristics”, an example of which is displayed in Table 5.19. The event number was simply repeated from the event starting time table, and the event types were dependant on the worksheet in which the table occurred. If an event number was present, then the volume and flow rate was calculated based on the continuous distribution function parameters selected in section 5.10.5 and 5.10.6.

Table 5.19: Event characteristics example

	AD	AE	AF	AG	AH	AI	AJ
1	Event Characteristics						
2 3	Event number	Event Type	Volume (ℓ)	Flow rate (ℓ/s)	Starting time	Duration (s)	Ending time
4	1	Shower	185	0.164	06:20:31 AM	1130	06:39:20 AM
5	2	Shower	79	0.140	07:11:55 AM	564	07:21:18 AM
6	3	Shower	63	0.131	11:14:17 AM	482	11:22:19 AM

In some cases the continuous distribution functions for the volume and flow rate were solved using the CDF typed into Microsoft Excel, and in other cases Microsoft Excel had built in functions which were used. If the shower volume Log-Logistic CDF is used as an example, the CDF is:

$$F(x) = \frac{1}{1 + \left(\frac{\beta}{x - \gamma}\right)^{\alpha}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5.1)$$

where $F(x)$ is a value between zero and one, and x is the corresponding volume.

The objective was, therefore, to solve the equation for x to obtain the volume, and a random number between zero and one was substituted in $F(x)$. The parameters $\alpha, \beta,$ and γ have known values unique to the end-use as calculated with @Risk software. The resulting equation for the Log-Logistic distribution solving for x was therefore:

$$x = \frac{\beta}{\left(\frac{1}{F(x)} - 1\right)^{1/\alpha}} + \gamma \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5.2)$$

Table 5.20 summarises the equations used to determine the event volume for the different end-uses, while Table 5.21 similarly describes the event flow rate equations.

Table 5.20: Event volume calculation equations

End-use	Distribution	Para-meter	Parameter cell reference	Microsoft Excel Equation
Shower	Log-Logistic	γ	Y4	=$Y$5/(((1/RAND())-1)^(1/$Y$6)))+ $Y$4
		β	Y5	
		α	Y6	
Bath	Rayleigh	β	Y4	=$Y$4*(-2*LN(1-RAND()))^(1/2)
Toilet	Weibull	α	Y4	=$Y$5*(-1*LN(1-RAND()))^(1/$Y$4)
		β	Y5	
Tap	Log Normal	β	Y4	=_xlfn.LOGNORM.INV(RAND(),$Y$4, $Y$5)
		α	Y5	
Dish-washer	Log-Logistic	γ	AD4	=($AD$5/(((1/RAND())-1)^(1/$AD$6))) +$AD$4
		β	AD5	
		α	AD6	
Washing machine	Weibull	α	AD4	=$AD$5*(-1*LN(1-RAND()))^(1/ $AD$4)
		β	AD5	

Table 5.21: Event flow rate calculation equations

End-use	Distribution	Para-meter	Parameter cell reference	Microsoft Excel Equation
Shower	Log-Logistic	γ	AB4	=($AB$5/(((1/RAND())-1)^(1/$AB$6))) +$AB$4
		β	AB5	
		α	AB6	
Bath	Weibull	α	AB4	=$AB$5*(-1*LN(1-RAND()))^(1/ $AB$4)
		β	AB5	
Toilet	Weibull	α	AB4	=$AB$5*(-1*LN(1-RAND()))^(1/ $AB$4)
		β	AB5	
Tap	Gamma	α	AB4	=_xlfn.GAMMA.INV(RAND(),$AB$4, $AB$5)
		β	AB5	
Dish-washer	Erlang	α	AJ4	=_xlfn.GAMMA.INV(RAND(),$AJ$4, $AJ$5)
		β	AJ5	
Washing machine	Weibull	α	AJ4	=$AJ$5*(-1*LN(1-RAND()))^(1/$AJ$4)
		β	AJ5	

The starting times were listed in ascending order by making use of Microsoft Excel's "small()" function. The "small()" function returns the k^{th} smallest number in a dataset, where the event number corresponding to the starting time is substituted as k. The duration of each event is calculated by dividing the volume by the flow rate. The end time is subsequently obtained by converting the duration to a time using the "time()" function, and adding it to the starting time.

The event characteristics table for the dishwasher and washing machine worksheets contained an additional column for the cycle number. For each event, the selected number of cycles received a cycle number and the volume and flow rates were only calculated if a cycle number was present. The event starting time table provided the starting times only for unique events, and not the starting times of individual cycles.

The duration between cycles was calculated by using the continuous distribution function parameters selected in sections 5.10.11 and 5.10.14. The starting time of a new cycle was therefore calculated as the starting time of the previous cycle plus the duration between cycles. Due to new random numbers being generated for each duration between cycles, it was possible for consecutive cycles to have different durations between cycles. Table 5.22 summarises the equations used to determine the duration between cycles for the different end-uses.

Table 5.22: Event duration between cycle calculation equations

End-use	Distribution	Para-meter	Parameter cell reference	Microsoft Excel Equation
Dish-washer	Log-Logistic	γ	AG4	=($AG$5/(((1/RAND())-1)^(1/$AG$6))) +$AG$4
		β	AG5	
		α	AG6	
Washing machine	Beta general	α_1	AG4	=_xlfn.BETA.INV(RAND(),$AG$4, $AG$5,$AG$6,$AG$7)
		α_2	AG5	
		min	AG6	
		max	AG7	

The final step in the end-use worksheets was to represent the flow rates when end-use events occurred, on a per second basis throughout the day. This was done by creating a table listing the time for each second of the day in one column and the flow rate in an adjacent column.

Table 5.23 shows extracts from the daily flow profile table. A very lengthy equation was used in the flow rate column to determine which value was displayed in each cell. The “if()” and “lookup()” function was used repeatedly such that if the time in column AL was greater than the starting time, but smaller than the ending time of any event in the event characteristics table, then the corresponding flow rate for that event was inserted in column AM. If the time did not overlap with any event, a value of zero was displayed. In this way each of the 86 400 time values in column AL was evaluated.

Since the flow rates were given in litres per second for each second that the flow rate occurred, the values in column AM essentially provided the volume of water flowing per second. The sum of values between 06:20:31 AM and 06:20:40 AM in Table 5.23, results in a volume of 1.64 litres (0.164 x 10). The sum of the entire AM column therefore provided the total volume of water attributed to the specific end-use.

Table 5.23: Daily end-use flow profile calculation

	AL	AM
1	Daily Flow	
2	Time	Flow rate (ℓ/s)
3		
4	12:00:00 AM	0
5	12:00:01 AM	0
6	12:00:02 AM	0
7	12:00:03 AM	0
8	12:00:04 AM	0
9	12:00:05 AM	0
395	06:20:31 AM	0.164
396	06:20:32 AM	0.164
397	06:20:33 AM	0.164
398	06:20:34 AM	0.164
399	06:20:35 AM	0.164

	AL	AM
1	Daily Flow	
2	Time	Flow rate (ℓ/s)
3		
400	06:20:36 AM	0.164
401	06:20:37 AM	0.164
402	06:20:38 AM	0.164
403	06:20:39 AM	0.164
404	06:20:40 AM	0.164
86394	11:59:50 PM	0
86395	11:59:51 PM	0
86396	11:59:52 PM	0
86397	11:59:53 PM	0
86398	11:59:54 PM	0
86399	11:59:55 PM	0
86400	11:59:56 PM	0
86401	11:59:57 PM	0
86402	11:59:58 PM	0
86403	11:59:59 PM	0

5.11.3. Household Summary Worksheet

The household summary worksheet contained the overall results from the end-use model. Table 5.24 presents extracts from the household summary worksheet.

Column A provided the row labels, which were constant. Column B represented the results for a single iteration of the model. Cell B2 simply repeated the household size value initially selected in the household size worksheet. Cell B4 repeated the total volume of the shower end-use, as calculated in the shower worksheet, by adding the per second flow rates in the daily flow table. Cells B5 to B9 similarly summarised the total volumes resulting from the respective end-uses, while cell B10 summed the end-use volumes to provide the total volume of water used by the household in one day. Cells B12 to B86411 added the flow rates occurring simultaneously from each of the six end-uses to obtain a total flow rate for the household at the particular time step. The values provided in the household summary worksheet were later utilised in the peak factor calculation procedure.

Column B was the only column that contained active equations in the cells. Each time the end-use model workbook was re-calculated it represented another iteration of the model. New random numbers were subsequently generated in all of the worksheets, which resulted in another possible scenario of household water demand and that changed the values in column B. It was therefore necessary to capture the constantly changing values in column B and save each unique scenario. This was done by copying column B and using the special paste function to paste the text values in an adjacent column.

Table 5.24: Household summary worksheet example

	A	B	C	D	ALM	ALN
1	Scenario		1	2	999	1000
2	No. of Persons	2	4	2	3	2
3	End-use	Total Volume (ℓ)				
4	Shower	94.23	338.02	26.90	379.06	94.23
5	Bath	95.15	36.63	74.79	118.81	95.15
6	Toilet	112.10	193.91	250.33	217.80	112.10
7	Tap	54.40	119.64	136.54	81.88	54.40
8	Dishwasher	15.04	49.36	83.28	0.00	15.04
9	Washing machine	31.13	336.19	384.28	291.10	31.13
10	Total	402.05	1073.74	956.12	1088.65	402.05
11	Time	Total Flow rate (ℓ/s)				
12	12:00:00 AM	0.00	0.00	0.00	0.00	0.00
13	12:00:01 AM	0.00	0.00	0.00	0.00	0.00
14	12:00:02 AM	0.00	0.00	0.23	0.00	0.00
15	12:00:03 AM	0.00	0.00	0.23	0.00	0.00
86410	11:59:57 PM	0.13	0.00	0.00	0.00	0.13
86410	11:59:58 PM	0.13	0.00	0.00	0.00	0.13
86411	11:59:59 PM	0.13	0.00	0.00	0.00	0.13

Thousands of iterations were required, the scenario saving procedure was therefore automated and repeated in Excel by means of a loop sequence. The loop was programmed as a macro in Microsoft Visual Basic. The calculation steps of the macro were as follows: On the flow per house sheet cell C2 was selected. Copied Range B2:B86411. Selected cell B2 so that it was the active cell. For w=1 (the first sequence in the loop) the active cell moved one column to the right and paste special values in that column. Once the values were pasted the workbook was re-calculated and new values were present in range B2:86411. For w=2, the second sequence in the loop, the active cell moved another column to the right and paste special new iteration values in that column. This was repeated for a set number of iterations, after which the workbook was saved.

By executing the macro for several different numbers of iteration loops, it was found that a maximum of one thousand iterations in a single workbook was successful. When more than a thousand iterations were done, the workbook stopped responding and all the data was lost. Figure 5.21 presents the code used in the macro to perform the saving procedure. When the macro was executed once, the result was the end-use model workbook containing 1 000 daily water demand scenarios in columns C to ALN.

```
Sub PF()
Sheets("Flow per house").Select
Range("C2").Select
Dim w As Integer
Range("$B$2:$B$86411").Select
Selection.Copy
Range("B2").Select
For w = 1 To 1000
    ActiveCell.Offset(0, 1).Select
    Selection.PasteSpecial  Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
        :=False, Transpose:=False
  Next w
ActiveWorkbook.Save
End Sub
```

Figure 5.21: Macro code to save individual model iterations

5.12. End-use Model Executions and Groupings

5.12.1. End-Use Model Size

The template of the Microsoft Excel end-use model workbook (before the macro was executed that saved the model iterations) had a size of about 265 MB. After the macro was executed, and data for 1 000 iterations had been saved in the workbook, it had an approximate size of between 393 MB and 397 MB. At the time of conducting this study, this was an exceptionally large file size.

The calculation time of a single iteration varied, depending on the computer on which the model was run. However, on average a single iteration had a calculation time of approximately 10 seconds, possibly due to the large number of calculations performed in the end-use model. Executing the macro once, and obtaining 1 000 iterations, therefore took almost three hours to complete. The relatively long duration of calculation proved to be a limitation of the model. Due to time constraints, the total number of executions performed in this study was limited. The end-use model macro was executed 100 times, resulting in 100 individual workbooks containing 1 000 iterations each. In total, therefore, 100 000 unique iterations of daily residential indoor water demand were available to be used in this study.

5.12.2. Household Groupings

It is important to note that a single iteration of the end-use model represented the indoor water demand of a single household on any given day. The water demand profiles resulting from different iterations did not explicitly characterise water demand on any particular day of the week or time of the year. Such temporal differences were taken into account in the distribution functions applied in the model. Water demand events corresponding to a weekday, weekend day, summer day or winter day were all included as possibilities that could be selected by any one of the iterations. Due to the probabilities associated with household sizes, event frequencies, starting hours, volumes, flow rates and so forth, it was expected that most of the iterations would yield average water demand results, while there would also be some extreme events. Not every iteration therefore represented the maximum water demand of that particular household. It was the intention to derive such typical “overall” water demand patterns, in order to assess the peak events in relation to the average demand over a given time period. The extreme cases were, however, the most relevant when dealing with PFs.

The objectives of this study included investigating how the water demand profiles (and ultimately the PFs) differed in differently sized residential areas. The differently sized residential areas were modelled by grouping the individual households together and taking the sum of their water demand throughout the day. A number of unique scenarios were produced for each residential area size (or household group size), in order to establish the variability of water demand within constant household group sizes, and provide a number of different possibilities.

It has been established that water demand is more variable within smaller household group sizes, and therefore it would be beneficial to have a greater number of different scenarios for the smaller groups. However, due to the limit of 100 000 unique iterations available, the number of scenarios per group was also limited. The original REUWS from which the raw data was obtained in this study consisted of 1 188 individual households. It was therefore decided that the total number of individual iterations used within each household group and scenario combination should not be less than 1 000, while the greatest number of scenarios is attributed the smallest household group sizes.

According to CSIR (2003), the PF is constant for more than 2 000 equivalent erven. A household group size of 2 000 was therefore used as the upper limit in this study, in order to compare the findings to the common PF curves of CSIR (2003). It would be easy to increase the group size in a future study by speeding up the computing time of the end-use model, and obtaining a greater number of iterations, but for the purpose of this study, 2 000 households as an upper limit was considered sufficient.

The household group sizes (the number of iterations summed to obtain a single water demand profile) and the number of scenarios selected for each group size is summarised in Table 5.25. In order to obtain the daily water demand profile for each group size, the required number of iterations was obtained from the end-use model workbooks and summed for each second of the day. This was repeated until the chosen number of scenarios was available. The format of the daily water demand profile remained the same as that presented in Table 5.24.

Table 5.25: Household group size summary

Household group size	Number of scenarios	Number of model iterations
1	1 000	1 000
2	1 000	2 000
3	1 000	3 000
4	250	1 000
5	200	1 000
6	200	1 200
7	200	1 400
8	200	1 600
9	200	1 800
10	100	1 000
20	100	2 000
30	100	3 000
40	50	2 000
50	50	2 500
60	50	3 000
70	50	3 500
80	50	4 000
90	50	4 500
100	10	1 000
200	10	2 000
300	10	3 000
400	10	4 000
500	10	5 000
600	10	6 000
700	10	7 000
800	10	8 000
900	10	9 000
1 000	5	5 000
2 000	5	10 000
Total	4 950	99 500

5.13. Peak Factor Calculation

Once a daily water demand profile was available for each scenario of household group size, the PFs were calculated for each profile. Peak factors were calculated as the ratio between the maximum flow rate (averaged over a selected short time period) and the average flow rate during a 24 hour period.

$$PF = \frac{(Q_{\delta t})_{max}}{(Q_{24h})_{avg}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5.3)$$

One of the objectives in this study involved investigating the effect on the magnitude of PFs of using different time intervals (δt) in the calculation procedure. Eight different time intervals were therefore selected to calculate the PFs. The time intervals used in this study are summarised in Table 5.26.

Table 5.26: Peak factor time intervals

$\boldsymbol{\delta t}$	$\boldsymbol{\delta t}$ **(seconds)**
1 second	1
10 seconds	10
1 minute	60
5 minutes	300
10 minutes	600
15 minutes	900
30 minutes	1800
60 minutes	3600

The volume of water in litres, sampled in one second intervals throughout the day, for a household or group of households was available from the end-use model results. A Microsoft Excel spreadsheet was created which calculated the PFs as follows:

1. Sum the volumes occurring between 12:00:00 AM and 11:59:59 PM (86400 seconds) to determine total volume during the day.
2. Divide the total daily volume (ℓ) by 86400 (s) to determine $(Q_{24h})_{avg}$, the average daily flow rate (ℓ/s).

3. For $\delta t = 3600s$, sum the volumes for each of the 24 consecutive 3600 second intervals. For example, between 12:00:00 AM - 12:59:59 AM, 01:00:00 AM - 01:59:59 AM, etc.
4. Identify the maximum volume of the twenty four 3600 second time intervals.
5. Divide the maximum volume (ℓ) by 3600 to determine $(Q_{3600s})_{max}$, the maximum flow rate (ℓ/s) averaged over 3600 seconds.
6. Calculate PF_{60min}, the 60 minute peak factor, by dividing $(Q_{3600s})_{max}$ by $(Q_{24h})_{avg}$.
7. Repeat steps 3 to 6 for $\delta t = 1800s$, $\delta t = 900s$, $\delta t = 600s$, $\delta t = 300s$, $\delta t = 60s$, $\delta t = 10s$, and $\delta t = 1s$.
8. Repeat steps 1 to 7 for each of the 4950 scenarios.

The 39 600 PFs calculated according to this δt-step procedure were summarised in a spreadsheet and used to construct various graphs, as discussed in Chapter 6.

6. RESULTS

6.1. End-Use Model Water Demand

The purpose of the end-use model developed in this research was not to simulate the average water demand of a particular geographical area, but rather to generate a myriad of different scenarios of water demand in order to determine the PFs from the diurnal water demand profiles.

A total of 99 500 of the end-use model iterations were included in the results of this study. For each iteration the household size, and the water demand from each of the six end-uses, were recorded. The average household size and end-use volumes for the 99 500 iterations could thus be obtained.

The model was not calibrated, nor was it the intention to duplicate existing datasets. Some sort of "agreement" would, however, be expected when comparing model results for daily total volumes and the relative contribution by end-uses to daily totals, with other data and formerly publicised results. A comparison of the average volume per capita per day results for the end-use model and the data from the REUWS is given in Table 6.1.

Table 6.1: Average volume per capita per day comparison

End-use	Volume per capita per day (ℓ)		Difference (ℓ)	Difference (%)
	End-use model	REUWS		
Toilet	60.2	71.3	-11.1	16.9
Shower	56.0	44.7	+11.3	22.4
Washing machine	112.4	57.8	+54.6	64.2
Tap	30.4	42.0	-11.6	32.0
Dishwasher	14.3	3.9	+10.4	114.3
Bath	31.3	4.6	+26.7	148.7
Total	304.6	224.3	+80.3	30.4

It is evident that that the end-use model over-estimated the per capita indoor water demand, when compared with the measured results in the REUWS for most end-uses. The average daily per capita water demand was over-estimated for the shower, washing machine, dishwasher, and bath, while for the toilet and tap it was underestimated. The most significant differences occurred with the washing machine, dishwasher and bath end-uses. A possible future improvement could be achieved by changing the methodology used in the end-use model to estimate the dishwasher and washing machine volumes.

The share that each end-use contributed to the overall indoor demand was also compared with other studies, as shown in Table 6.2. The percentages quoted for other studies in Table 6.2 are presented as a fraction of the water demand for the six relevant end-uses in this study. The shares of individual end-uses appear to be reasonably within the given ranges reported by others.

Table 6.2: End-use share comparison

End-use	Toilet	Shower	Washing Machine	Tap	Dish-washer	Bath	Total Indoor
(Edwards and Martin, 1995)	34.0	4.1	21.6	25.8	1.0	13.4	100.0
(DeOreo et al., 1996)	29.3	19.5	28.2	16.7	3.4	2.9	100.0
(Mayer et al., 1999)	31.9	19.9	25.7	18.7	1.8	2.0	100.0
(DeOreo et al., 2001) Pre-retrofit	33.0	15.9	26.1	16.1	2.5	6.5	100.0
(DeOreo et al., 2001) Post-retrofit	21.0	23.0	24.4	21.2	3.2	7.1	100.0
(Loh and Coghlan, 2003)	22.0	34.1	26.8	17.1	-	-	100.0
Mayer et al. (2003) Pre-retrofit	33.0	19.9	23.0	17.4	1.7	5.0	100.0
Mayer et al. (2003) Post-retrofit	22.5	24.6	20.2	24.1	2.1	6.4	100.0
(Roberts, 2005)	19.3	31.5	27.1	17.7	1.7	2.6	100.0
Heinrich, 2007	19.5	39.4	23.7	13.6	1.3	2.5	100.0
(Willis et al., 2009)	15.4	36.4	22.0	19.8	1.6	4.7	100.0
This study	19.8	18.4	36.9	10.0	4.7	10.3	100.0

The water demand profiles generated by the end-use model were deemed to be acceptable in terms of this basic verification.

6.2. Peak Factor Variance

Subsequently, the PFs were calculated for 29 different group size combinations, with groups ranging from 1 household to 2 000 households. These group sizes could be equated to “water use zones”. For each group size eight different time intervals were used in the PF calculation. Since the PFs were calculated as the maximum water demand in a short period divided by the average water demand in that day, the baseline demand applicable to these PFs is the average daily demand. Figures C1.1 to C1.24 in Appendix C show the resulting peak factors that were plotted in ascending order for each of these scenarios. The magnitude of the PF is presented on the y-axis and the percentiles on the x-axis. These plots provided a graphical presentation of the range of PFs that were obtained for the household group sizes and time intervals for all of the iterations. The PF values represented by various percentiles of the results are provided in Tables C2.1 to C2.8 in Appendix C. The cells in the table were colour coded based on their values, where green denoted the lowest values and red denoted the highest values. This allowed the variance within household group sizes to be clearly evident.

The actual variance values of the PFs for each scenario were also computed. For all of the different time intervals, the PFs within the group sizes between one and ten households showed the greatest variance, while the group sizes between 100 and 2 000 households varied very little. For example, the PF_{60min} for the single household group had a variance of 5.057, while the same time interval PF for the 2 000 household group had a variance of 0.001. This reduction in variance with increased sample size is typical and as expected. Within all of the household group sizes, the PF resulting from small time intervals showed greater variability than the PF resulting from longer time intervals. For example, for the 500 household group size, the PF_{60min} had a variance of 0.006, while the PF_{1s} had a variance of 0.073. An overview of all the variances is provided in Table 6.3.

Table 6.3: Peak factor variance

Variance		PF time intervals							
		60 min	30 min	15 min	10 min	5 min	1 min	10 sec	1 sec
Group size (number of combined households)	1	5.057	13.321	44.365	72.224	160.397	484.735	754.526	798.739
	2	2.086	5.198	11.665	18.297	39.057	95.824	134.737	152.978
	3	1.180	2.796	5.587	9.588	16.241	40.594	53.090	55.964
	4	0.760	1.789	4.566	5.600	8.975	19.314	26.011	27.671
	5	0.453	1.350	2.921	3.789	8.067	12.638	15.099	16.835
	6	0.498	1.134	1.805	3.166	4.362	10.939	18.220	20.628
	7	0.417	0.842	1.509	2.311	3.581	8.015	9.938	10.184
	8	0.243	0.581	1.218	2.117	3.155	7.038	9.984	10.990
	9	0.359	0.644	1.309	1.804	2.708	4.256	6.712	7.118
	10	0.248	0.517	0.994	1.463	2.492	4.049	4.117	4.658
	20	0.104	0.192	0.369	0.445	0.775	1.573	1.906	1.965
	30	0.044	0.132	0.240	0.375	0.551	0.826	1.074	1.156
	40	0.039	0.089	0.189	0.230	0.407	0.470	0.459	0.495
	50	0.046	0.107	0.119	0.180	0.258	0.284	0.313	0.441
	60	0.025	0.061	0.083	0.194	0.182	0.536	0.650	0.704
	70	0.042	0.091	0.127	0.169	0.248	0.508	0.649	0.635
	80	0.030	0.050	0.130	0.183	0.308	0.388	0.457	0.476
	90	0.021	0.056	0.101	0.119	0.205	0.277	0.312	0.316
	100	0.018	0.051	0.090	0.044	0.127	0.172	0.268	0.267
	200	0.016	0.035	0.067	0.036	0.082	0.229	0.310	0.312
	300	0.007	0.019	0.020	0.035	0.041	0.049	0.057	0.054
	400	0.009	0.013	0.013	0.038	0.032	0.035	0.037	0.039
	500	0.006	0.015	0.022	0.029	0.038	0.065	0.073	0.073
	600	0.007	0.014	0.015	0.052	0.059	0.068	0.110	0.116
	700	0.002	0.003	0.013	0.027	0.031	0.042	0.052	0.053
	800	0.005	0.010	0.024	0.011	0.018	0.022	0.027	0.037
	900	0.003	0.005	0.016	0.017	0.027	0.019	0.017	0.019
	1000	0.003	0.011	0.009	0.011	0.006	0.024	0.014	0.013
	2000	0.001	0.008	0.019	0.013	0.024	0.020	0.011	0.012

Legend:

low ← variance → high

6.3. Maximum Peak Factor Comparison

The purpose of PFs is to represent a safety factor which denotes a limiting demand condition. Although many possible PFs were obtained in this study, the maximum PFs are of most concern. Some authors such as Booyens (2000) and Johnson (1999) have highlighted the benefits of assigning a return period to the PFs, which gives an indication of the risk of exceedance. In this study the limited number of scenarios in each household category did not allow for a reliable estimate of PF return periods. Future research may involve increasing the number of scenarios so that the return period can also be investigated. Only the maximum PFs resulting from the model were considered for further analysis.

The maximum PFs obtained using each of the different time intervals in each household group size were extracted from the result sets and plotted, as presented in Figure 6.1. The number of households is given on the x-axis in logarithmic scale, and the logarithmic scale was also used for the PFs given on the y-axis. The time intervals were each plotted as separate series.

From the results shown in Figure 6.1, it is clear that for all values of δt, the PF is relatively large for a small number of houses, and decreases as the household group size increases. This is in agreement with other studies. This is because the variability of water demand decreases when the combined water demand of many houses is considered. The PFs are the highest when a small δt is used in the calculation. As δt increases, the PF decreases. When longer time intervals are used, the variation of flow tends to be averaged out, this decreases the ratio between the average and the peak flow.

For a small number of houses there is a large difference between the PF calculated with different time intervals. In almost all cases the difference between PFs decreases as the number of houses increases. From the results in this study the difference between the PF_{1s} and PF_{10s} is almost negligible. This suggests that a 1 second logging frequency would not increase flow measurement results significantly, compared to a 10 second logging frequency.

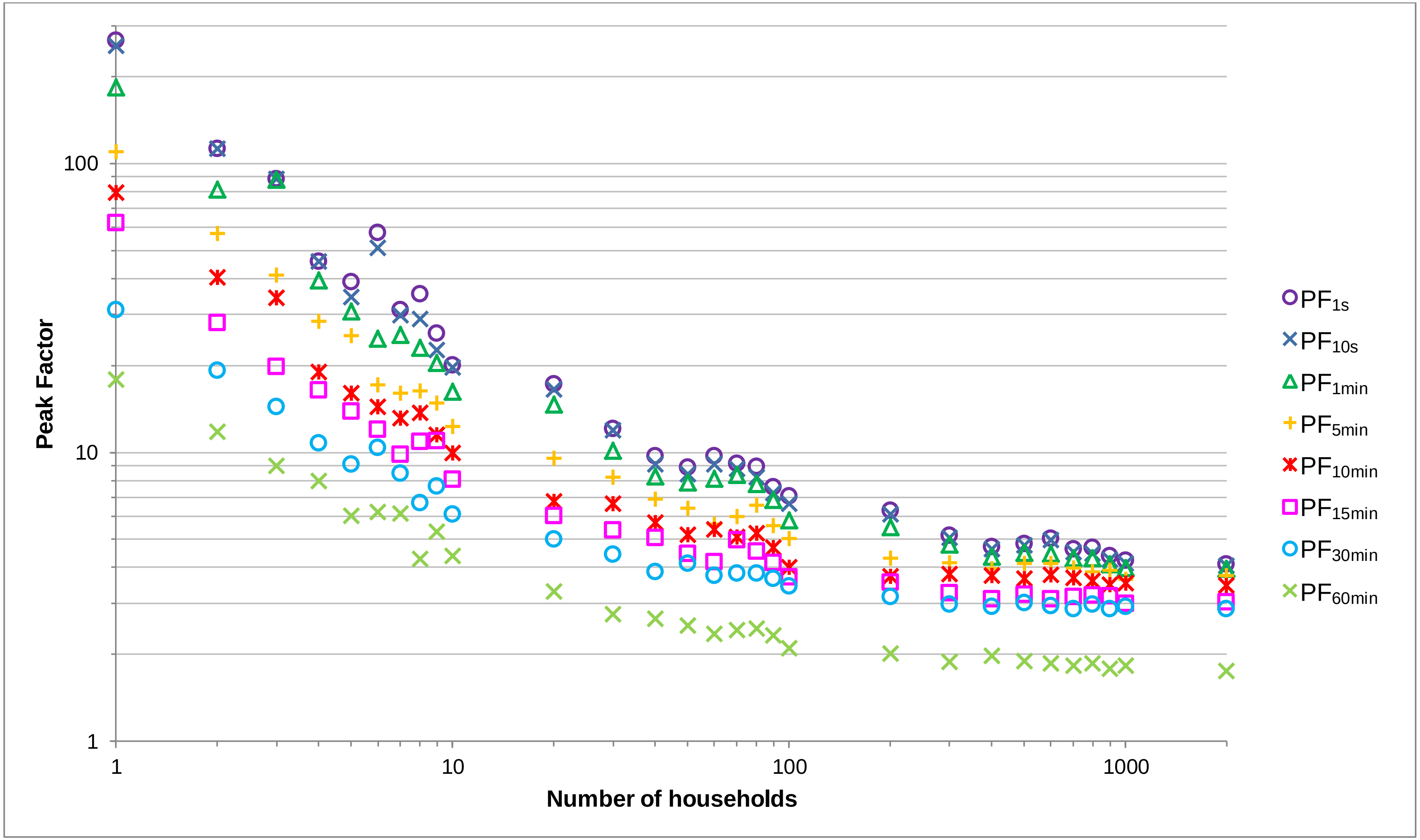


Figure 6.1: Comprehensive result set of all extracted maximum peak factors

The maximum PFs were then compared to the PFs reported by CSIR (2003), as shown in Figure 6.2. It is important to note that the PF results presented account for indoor water demand only, while the CSIR (2003) PFs included outdoor water demand. The CSIR (2003) provided the PF curve plotted against ee, where 1 ee = 1000 ℓ AADD. For the purpose of the comparison, it was necessary to convert ee to the number of households. The average water demand per iteration per household in this study was calculated to be 794 ℓ. It was therefore approximated that the AADD needed for conversion was 794 ℓ, hence a conversion of 0.8 ee = 1 household was used to plot the CSIR (2003) curve in Figure 6.2.

In the 2 000 household group size category, the PFs between the PF_{1s} and PF_{30min} categories range from 2.88 to 4.10. This means that the difference is only 1.23, which is relatively small. The PF_{60min} is consistently lower, to a greater extent than the other time intervals, and this is especially evident for the 2 000 household group size.

If it were assumed that the instantaneous PFs given in CSIR (2003), are equivalent to δt of one second, then the PF_{1s} curve derived in this study yields lower PFs for groups of 20 households and more. For 2 000 households the CSIR (2003) PF and the PF_{1s} from this study are approximately equal. The most significant difference between the curves occurs for household group sizes between one and 20. The magnitudes of the PFs from this study focusing on indoor use are notably larger than those reported for combined indoor and outdoor demand by the CSIR (2003).

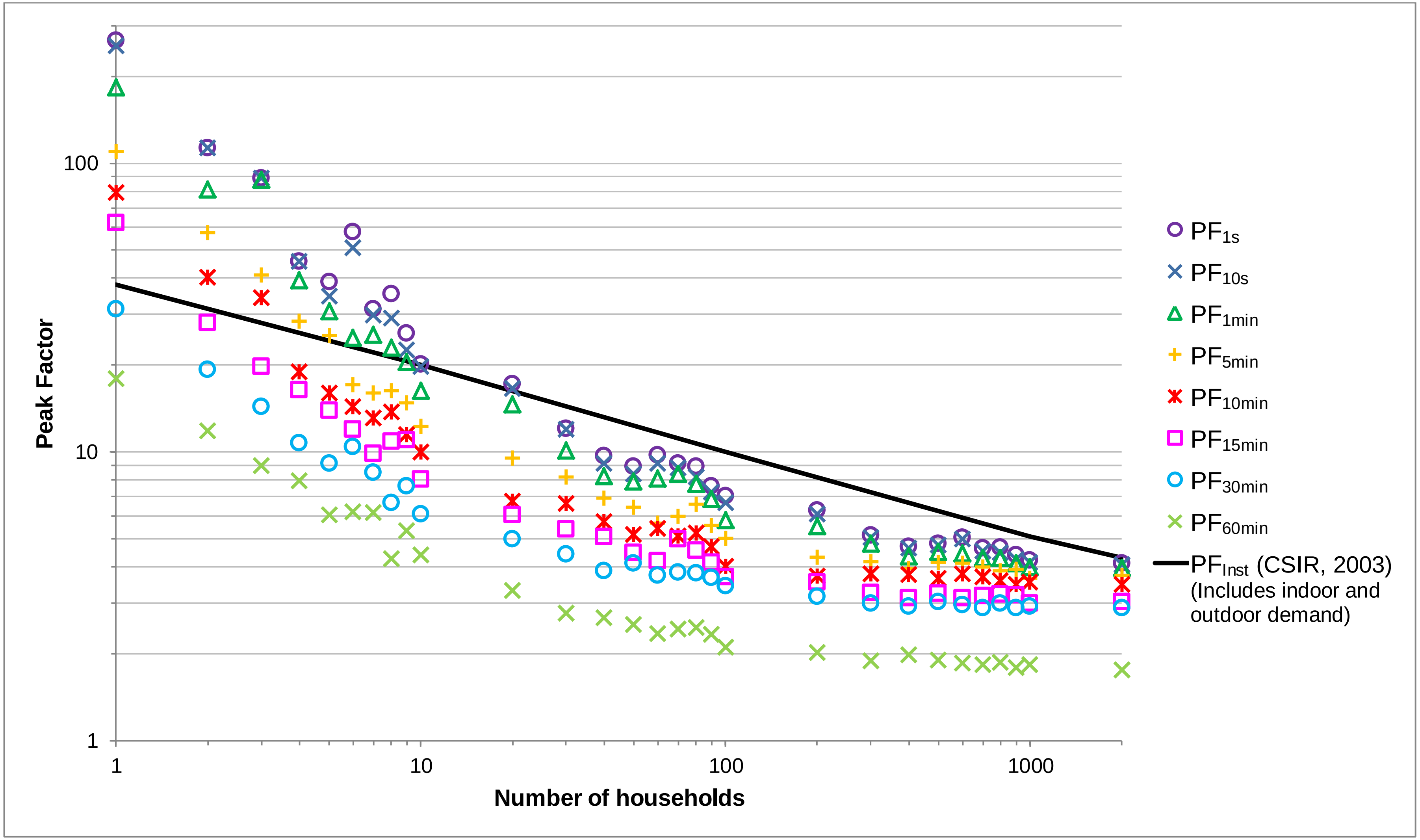


Figure 6.2: Comparison of peak factor results with CSIR (2003)

The maximum PFs were also compared to the PFs reported by Booyens (2000), as shown in Figure 6.3.

Booyens (2000) calculated the PF for different time intervals for residential areas. The residential areas consisted of 69, 444, and 794 erven. For the purpose of the comparison it was assumed that 1 erven = 1 household, to plot the points shown in Figure 6.3. The water consumption measured by Booyens included the total water consumption, including outdoor consumption.

It is interesting to note that although the magnitudes of the Booyens (2000) PFs are not the same as the results obtained in this study, the general trend is similar. The range of PFs achieved by applying different time intervals is larger for the smaller household group sizes than the larger groups. The most noticeable difference between the two studies is the variance evident between PFs for different values of δt for a given household group size. The PF results from Booyens (2000) imply that there is a difference of 0.3 between PF_{1min} and PF_{60min} for 794 households, while a difference of about 2.4 is observed in this study. The larger PF ranges achieved in this study may be attributed to the fact that water consumption characteristics of households throughout North America were applied in this study, while Booyens (2000) used local data. The water measured by Booyens (2000) represented homogeneous areas, which may have had similar water demand characteristics, resulting in less variation.

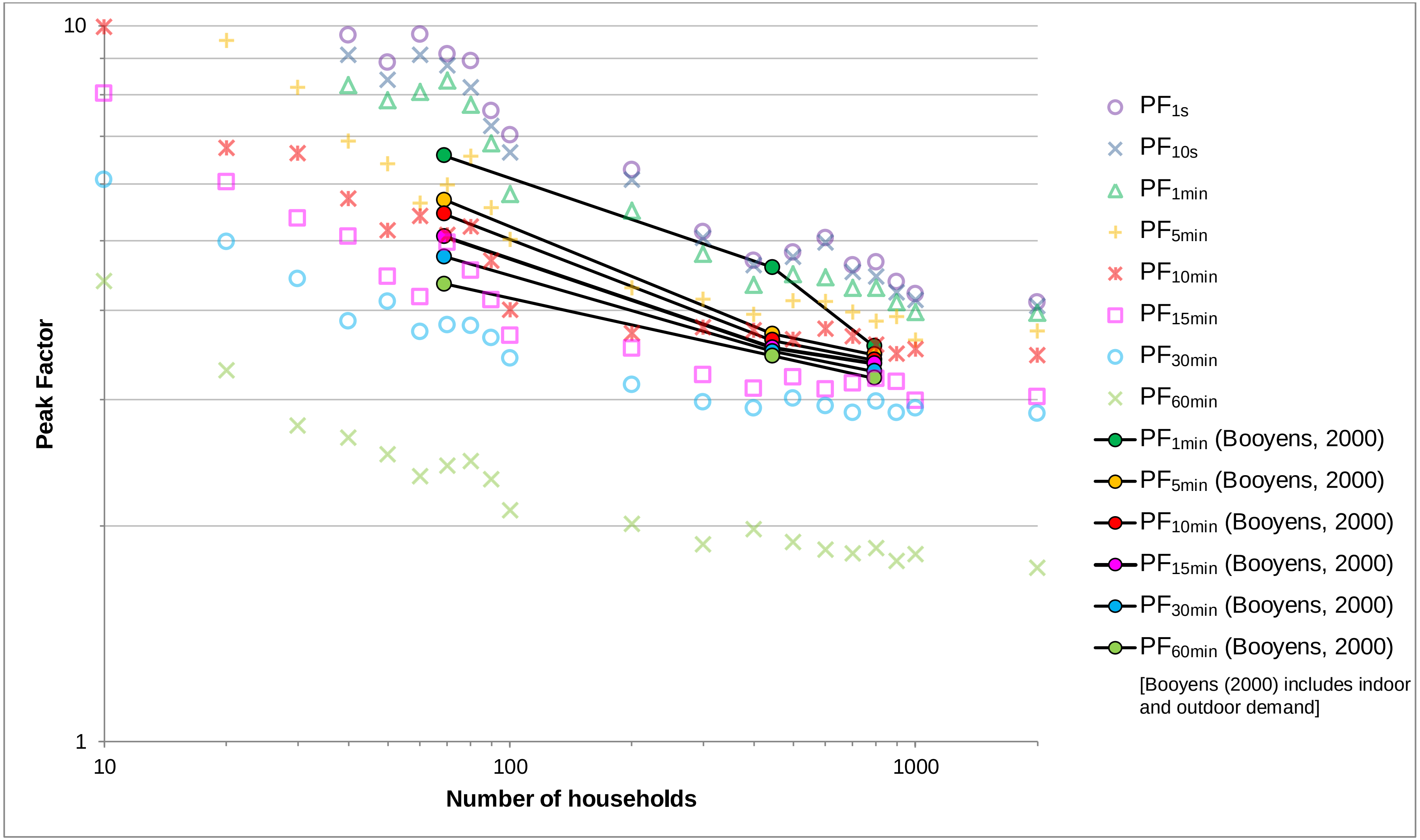


Figure 6.3: Comparison of peak factor results with Booyens (2000)

7. CONCLUSIONS AND RECOMMENDATIONS

7.1. Summary of Findings

The design of a water distribution system is often based on the most limiting demand conditions of the system. The estimated peak demand is one of the limiting demand conditions taken into consideration when determining the capacity of pipelines.

The literature reviewed suggests that the most widespread method of estimating peak demand is by multiplying the average demand by a peak factor. Some studies have derived peak factors from locally measured water consumption data, or by providing empirical equations or figures where the PF is specified as a function of population. In addition, studies such as van Zyl (1996), Zhang et al. (2005), and Tricarico et al. (2007) have investigated the use of probability theory to derive peak factors, but no reports could be found where end-use modelling was used as a basis to derive demand profiles and subsequent peak factors.

Various end-use models and tools were reviewed for the purpose of deriving probability based peak factors, including the flow trace method and models REUM and SIMDEUM. A similar approach to that used in SIMDEUM was eventually employed.

As part of this research a computer based stochastic end-use model was developed to estimate the daily residential water demand for a single house in one second time steps. Water demand was assumed to occur in rectangular pulses, where the water pulses described an end-use specific volume and flow rate. The REUWS database, containing measured end-use consumption data, was utilised to derive probability distributions for each of the end-model parameters. A single iteration of the end-use model represented a possible water demand scenario for a single household. The end-use model was executed 100 times and a total of 99 500 iterations of the end-use model was eventually used in the study. The water demand from individual iterations was summed to obtain the combined water demand from a group of households. The daily water demand was calculated for 29 different group sizes that ranged from one to 2 000 households. For each group size, a number of daily water demand scenarios were available for comparison.

It was found that the average total indoor water demand per household estimated by the end-use model was within a reasonable range. The share of water demand from the different end-uses was, however, different from that of other studies. It was found that the end-use model overestimated the share of water used by baths, dishwashers and washing machines, but this was not considered problematic.

The daily water demand results from the end-use model were then applied to determine the peak factors for each scenario and household group size. Eight different time intervals were used for the purpose of determining peak factors for each water demand scenario. The time intervals consisted of 1 second, 10 seconds, 1 minute, 5 minutes, 10 minutes, 15 minutes, 30 minutes and 1 hour. The maximum peak factor for all the time intervals was plotted against each household group size. Comparisons were made of the PF results to the PF presented by CSIR (2003) and Booyens (2000).

In the literature, peak factors that are often recommended to be used in design are PF_d, PF_h, or PF_{inst}, while empirical peak factors are often derived from 10 minute or 15 minute logging frequencies (Booyens, 2000; Tricarico, 2007). Booyens (2000) concluded that for developments with an AADD of less than $100\ \mathrm{k}\ell$ a PF_{15min} could be applied, while a PF_{1h} is acceptable for developments with an AADD greater than $100\ \mathrm{k}\ell$. No recent additional evidence was found in the literature that presented the circumstances for which the application of different peak factor time intervals would be appropriate.

7.2. Conclusion

The end-use model presented in this study yielded indoor water demand estimations that compare well with results from other studies. Several improvements can be made to the model, however it can be concluded that the probability based end-use model presented here is a useful method for deriving residential daily water demand profiles on a one second temporal scale.

Within any one water demand scenario, varying peak factors can be obtained by changing the time interval over which the peak flow is calculated. It is therefore very important that the peak factor term should be not quoted in isolation. Any statement of a peak factor must be accompanied by information on the associated time interval.

The peak factors across all time intervals were found to be inversely related to the number of households studied. As the number of households increased, the peak factors decreased. By visually inspecting the magnitudes of the peak factors, three distinct gradient changes were apparent, due to the rate of change of peak factor values. For the category of one to ten households the magnitude of peak factors decreased relatively rapidly as the number of households increased. The gradient decreased for the category of ten to 100 households. The flattest gradient resulted for the 100 to 2 000 household category, indicating that relative to the other categories the peak factor did not decrease as much with increased number of households.

A strong inverse relationship was evident between the magnitude of the peak factors and the peak factor time interval. As the time interval increased, the magnitude of the peak factor decreased. The degree to which the peak factor decreased with longer durations was affected by the number of households. For a small number of households the time interval had a significant effect on the peak factors, while the effect decreased as the number of households increased. This suggests that the peak factor time interval is not a pertinent consideration for more than 1 000 households. However, when smaller household group sizes are involved, the peak factor time interval will impact peak factor results to a greater extent.

The design of water distribution systems should incorporate different peak factor intervals when considering separate components of the system. For large pipelines distributing water to an area exceeding 1 000 households, a peak factor with a long time interval could be applied, since lower flow rate variability is likely to occur. A peak factor with a short time interval should be applied to pipelines directly servicing a street of ten households, for example.

In the absence of site-specific knowledge the derived peak factors from this research could be applied to estimate the indoor residential component of peak flow rate in a WDS.

7.3. Suggestions for Further Research

Only indoor water demand was considered in this study. It would be beneficial to include outdoor water demand in peak factor calculations in a future study.

This study focused on residential water demand, a typical urban water demand profile, however, includes water losses, industrial, commercial, and institutional water demand. Future work may consider including other components of the total water demand that a WDS may need to cater for.

The share of average water demand resulting from the end-use model's simulation of the bath, dishwasher and washing machine was overestimated. Those results may be improved if more stringent end-use model parameter limits are introduced. For example, the model could be assigned smaller ranges of volumes and flow rates for each end-use and the daily frequency or number of possible cycles per event could be decreased. The use of alternative theoretical distributions, or the direct incorporation of measured sample data could be investigated as a means to possibly improving the results.

Water flow rates in a WDS are dependant not only on the water demand, but also on the pressure in the system. If the pressure is extremely low, then a limited flow rate is available which may be less than the water demand. The large peak factors that were observed for the small number of households may have occurred because pressure was not taken into account. If the peak demand is limited by the pipe infrastructure, and especially the plumbing system, on the household property, then perhaps such large peaks will be less evident. Considerable scope remains to improve the end-use model by describing peak flow rate as a function of residual WDS pressure.

In the absence of local data, the probability distributions used to describe the parameters of the residential end-use model in this study were obtained from North American water measurements. The probability distributions may yield more representative results for water demand of South African households if South African data is used, but such data is not yet available.

The maximum peak factors were used in this study, however, the frequency with which peak flow rates occurred, was not taken into account. It may, for example, not be necessary to apply a design peak factor that only occurs 1% of the time. It may be beneficial to associate acceptable reliabilities with peak factors.

Microsoft Excel was used to construct the residential end-use model. The total computing time necessary to produce the results of this study was approximately 275 hours to complete all 99 500 model runs. The computation speed of the end-use model could be improved significantly by restructuring the model in a more efficient manner.

7.4. Summary of Contributions

An end-use model describing indoor residential water demand in a probabilistic manner has not been presented before in South Africa. Although the water demand characteristics described by the end-use model in this study may not be representative of all South African household types, it does provide a basis from which to conduct further research and improve the model.

In this research the effect that time intervals have on the magnitude of peak factors across a wide range of household group sizes was investigated. This study presented a detailed investigation that explicitly demonstrated the notable effect that time intervals have on peak factors. This study therefore emphasised the need to further investigate the incorporation of different peak factor time intervals in design guidelines in order to achieve optimum water distribution system designs.

REFERENCE LIST

Ang, A. H-S. & Tang, W.H. 1984. *Probability concepts in engineering planning and design*. Volume II - Decision, risk, and reliability. New York, USA: John Wiley & Sons.

American Water Works Association. 1989. *Distribution network analysis for water utilities*. Denver (CO), USA: American Water Works Association.

American Water Works Association. 1999. *Design and construction of small water systems: an AWWA small systems resource book*. Denver (CO), USA: American Water Works Association.

American Water Works Association. 2004. *Sizing water service lines and meters*. Denver (CO), USA: American Water Works Association.

Billings, R.B. & Jones, C.V. 2008. *Forecasting Urban Water Demand*. Denver (CO), USA: American Water Works Association.

Blokker, E.M.J., Buchberger, S.G., Vreeburg, J.H.G. & van Dijk, J.C. 2008. Comparison of water demand models: PRP and SIMDEUM applied to Milford, Ohio, data. *Proceedings of the 10th Annual Water Distribution Systems Analysis Conference, Water Distribution System Analysis 2008, van Zyl, J.E., Ilemobade, A.A., Jacobs, H.E. (eds.), 17-20 August 2008, Kruger National Park, South Africa.*

Blokker, E.J.M., Vreeburg, J.H.G. & van Dijk, J.C. 2010. Simulating residential water demand with a stochastic end-use model. *Journal of Water Resources Planning and Management*, 136(1), 19-26.

Booyens, J. 2000. Spitsvloei in munisipale waterverspeidingsnetwerke. Unpublished master’s thesis, Randse Afrikaanse Universiteit, Johannesburg, South Africa.

Bowen, P. T., Harp, J.F., Baxter, J.W. & Shull, R.D. 1993. *Residential water use patterns*. Denver (CO), USA: American Water Works Association Research Foundation.

Brière, F.G. 2007. *Drinking-water distribution, sewage, and rainfall collection*. Montréal, Canada: Presses Internationales Polytechnique.

Britton, T., Cole, G., Stewart, R. & Wiskar, D. 2008. Remote diagnosis of leakage in residential households. *Water: Journal of the Australian Water Association*, 35(6), 89-93.

Buchberger, S.G. & Schade, T. 1997. Poisson pulse queuing model for residential water demands. *Proceedings of the 27th Congress of the International Association of Hydraulic Research*, 10-15 August 1997, *San Francisco (CA), USA,* 488-493.

Buchberger, S.G. & Wells, G. 1996. Intensity, duration, and frequency of residential water demands*. Journal of Water Resources Planning and Management*, 122(1), 11-19.

Buchberger, S.G. & Wu, L. 1995. Model for instantaneous residential water demands. *Journal of Hydraulic Engineering*, 121(3), 232-246.

Buchberger, S.G., Carter, J.T., Lee, Y. & Schade, T.G. 2003. *Random demands, travel times, and water quality in dead ends*., Denver (CO), USA: American Water Works Association Research Foundation, AWWARF Report 90963F.

Buchberger, S.G., Blokker, M. & Vreeburg, J. 2008. Sizes for self-cleaning pipes in municipal water supply systems. *Proceedings of the 10th Annual Water Distribution Systems Analysis Conference WDSA2008, van Zyl, J.E., Ilemobade, A.A., Jacobs, H.E. (eds.), 17-20 August 2008, Kruger National Park, South Africa.*

Burn, L.S., De Silva, D. & Shipton, R.J. 2002. Effect of demand management and system operation on potable water infrastructure costs. *Urban Water,* 4(3), 229-236.

Butler, D. & Memon, F.A. 2006. *Water demand management.* London, UK: IWA Publishing.

Chandapillai, J. 1991. Realistic simulation of water distribution system. *Journal of Transportation Engineering*, 117(2), 258-263.

Council for Scientific and Industrial Research. 1983. *Guidelines for the provision of engineering services and amenities in residential townships*. Compiled for the Department of Community Development by the CSIR.

Council for Scientific and Industrial Research. 2003. *Guidelines for Human Settlement and Design*. A report compiled under the patronage of the Department of Housing by the CSIR. Pretoria, South Africa: CSIR Building and Construction Technology. Report No. BOU/E2001.

Day, D. & Howe, C. 2003. Forecasting peak demand- what do we need to know? *Water Science and Technology: Water Supply*, 3(3), 177-184.

DeOreo, W. B., Dietemann, A., Skeel, T., Mayer, P. W., Lewis, D. M. & Smith, J. 2001. Retrofit realities. *Journal of the American Water Works Association*, 93(3), 58-72.

DeOreo, W.B., Heaney, J.P. & Mayer, P.W. 1996. Flow trace analysis to assess water use. *Journal of the American Water Works Association*, 88(1), 79-90.

Devore, J. L. 2004. Probability and statistics for engineering and the sciences. 6th edition. Belmont (CA), USA: Brooks/Cole Thomson Learning.

Diao, K., Barjenbruch, M. & Bracklow, U. 2010. Study on the impacts of peaking factors on a water distribution system in Germany. *Water Science & Technology: Water Supply*, 10(2), 165-172.

Edwards, K. and Martin, L. 1995. A methodology for surveying domestic water consumption. *Water and Environment Journal*, 9(5), 477-488.

Forbes, C., Evans, M., Hastings, N. & Peacock, B. 2011. *Statistical Distributions*. 4th edition. Hoboken (NJ), USA: John Wiley & Sons.

GLS Software. 2012. *SWIFT*. Available from: <http://www.gls.co.za/software/products/swift/html> [Accessed: 5 May 2012].

Gravetter, F. J. & Wallnau, L. B. 2000. *Statistics for the behavioural sciences.* 5th edition. Belmont (CA), USA: Wadsworth/Thomson Learning.

Gupta, R. & Bhave, P.R. 1996. Comparison of methods for predicting deficient-network performance. *Journal of Water Resources Planning and Management,* 122(3), 214-217.

Hare, A.H. 1989. Aspects on peak flows in water mains. *Institute of Municipal Engineering of Southern Africa (IMIESA)*, 14(4), 24-28.

Heinrich, M. 2007. *Water end use and efficiency project (WEEP) - Final report*. BRANZ Study report 159. Judgeford, New Zealand: BRANZ.

Howell, D. C. 2002. *Statistical methods for psychology.* 5th edition. Pacific Grove (CA), USA: Duxbury/Thomson Learning.

Hyun, I.-H., Dockko, S. & Lee, C.-I. 2006. Determination of design capacity for water transmission systems based on the characteristics of regional peak factors. *Water Science and Technology: Water Supply*, 6(5), 85-92.

Jacobs, H.E. 2004*. A conceptual end-use model for residential water demand and return flow*. Unpublished doctoral dissertation, Rand Afrikaans University, Johannesburg, South Africa.

Jacobs, H.E., Fair, K.A., Geustyn, L., Daniels, J. & Du Plessis, J.A. 2007. Analysis of water savings: A case study during the 2004-2005 water restrictions in Cape Town. *Journal of the South African Institution of Civil Engineers*, 49(3), 16-26.

Jacobs, H.E. & Fair, K.A. 2012. Evaluating a tool to increase information-processing capacity for consumer water meter data. *SA Journal of Information Management*, 14(1), Article No. 500, 7 pages. Available from: <http://dx.doi.org/10.4102/ sajim.v14i1.500> [Accessed: 7 July 2012].

Jacobs, H.E., Geustyn, L.C. & Loubser, B.F. 2006. Water - how is it used at home? *Proceedings of the Biennial Water Institute of South Africa (WISA) Conference, 21-25 May 2006, Durban, South Africa.*

Jacobs, H.E. & Strijdom, J.L. 2009. Evaluation of minimum residual pressure as design criterion for South African water distribution systems. *Water SA*, 35(2) WISA 2008 Special Edition, 183-191.

Johnson, R. A. 1994. *Miller and Freund's probability and statistics for engineers*. 5th edition. London, United Kingdom: Prentice Hall International.

Johnson, E.H. 1999. Degree of utilisation - The reciprocal of the peak factor. Its application in the operation of a water supply and distribution system. *Water SA,* 25(1), 111-114.

Lingireddy, S., Wood, D.J. & Nelson, A. 1998. Modified pipe network model for incorporating peak demand requirements. *Journal of Water Resources Planning and Management*, 124(5), 296-299.

Loh, M. & Coghlan, P. 2003. *Domestic water use study*. Perth, Australia: Water Corporation.

Loucks, D. P. & van Beek, E. 2005. *Water resources systems planning and management. An Introduction to methods, models and applications*. Paris, France: United Nations Educational, Scientific and Cultural Organisation (UNESCO).

Loureiro, D., Coelho, S.T., Machado, A., Santos, A., Alegre, H. & Covas, D. 2006. Profiling residential water consumption. *Proceedings of the 8th Annual Water Distribution Systems Analysis Symposium, 27-30 August 2006, Cincinnati, Ohio, USA*.

Lugoma, M.F.T., van Zyl, J.E. & Ilemobade, A.A. 2012. The extent of on-site leakage in selected suburbs of Johannesburg. *Water SA,* 38(1), 127-132.

Martínez-Solano, J., Iglesias-Rey, P.L., Pérez-García, R., and López-Jiménez, P.A. 2008. Hydraulic analysis of peak demand in looped water distribution systems. *Journal of Water Resources Planning and Management*, 134(6), 504-510.

Mayer, P.W., DeOreo, W.B., Opitz, E.M, Kiefer, J.C, Davis, W.Y., Dziegielewski, B. & Nelson, J.O. 1999. *Residential end uses of water*. Denver (CO), USA: American Water Works Association Research Foundation and American Water Works Association.

Mayer, P.W., DeOreo, W.B., Towler, E. & Lewis, D.M. 2003. *Residential indoor water conservation study: Evaluation of high efficiency indoor plumping fixture retrofits in single-family homes in the East Bay municipal utility district service area.* Boulder (CO), USA: Aquacraft, Inc. Available from: <http://www.aquacraft.com/sites/default/files/pub/Aquacraft-%282003%29-EBMUD-Residential-Indoor-Water-Conservation-Study.pdf> [Accessed: 7 July 2012].

McKenzie R. & Seago, C. 2005. *Benchmarking of leakage from water reticulation systems in South Africa*. Gezina, Gauteng, South Africa: Water Research Commission. WRC Report No. TT 244/05.

Mckenzie, R. & Lambert, A. 2004. Best practice performance indicators for non-revenue water and water loss components: A practical approach. *Water 21*, 6(4), 43-45*.* Available from: <http://www.iwapublishing.com/pdf/August_2004.pdf> [Accessed: 7 July 2012].

Milke, M.W., Wong, V. & Mc Bean, E.A. 2008. Describing variability of MSW composition data with the Log-Logistic distribution. *Waste Management and Research*, 26(4), 355-361.

Myburgh, E. 2012. *Evaluating methods for fire protection and related fire risk categories in rural towns of the Western Cape, South Africa*. Unpublished master's thesis, University of Stellenbosch, South Africa.

Palisade Corporation. 2010. *Guide to using @Risk*. Risk analysis and simulation add-in for Microsoft Excel, Version 5.7. Available from: <http://palisade.com/downloads/ manuals/EN/RISK5_EN.pdf> [Accessed: 24 August 2011].

Perelman, L. & Osfeld, A. 2006. Demand loading conditions: To what extent are they representative? *Proceedings of the 8th Annual water distribution system analysis symposium, 27-30 August 2006, Cincinnati, Ohio, USA.*

Queensland Water Commission. 2012. *Permanent water conservation measures*. Available from: <http://www.qwc.qld.gov.au/efficiency/water-conservation.html.> [Accessed: 18 August 2012].

Race, J. & Burnell, D. 2004. Water distribution systems analysis: Analysis of domestic demand patterns. *Proceedings of the joint conference on water resource engineering and water resources planning and management, 30 July - 2 August 2000, Minneapolis, Minnesota, USA.*

Roberts, P. 2005. *Yarra Valley Water 2004 residential end use measurement study*. Melbourne, Australia: Yarra Valley Water.

Tanyimboh, T.T., Burd, R., Burrows, R. & Tabesh, M. 1999. Modelling and reliability analysis of water distribution systems. *Journal of Water, Science and Technology*, 39(4), 249-255.

Tanyimboh, T.T., Tahar, B. & Templeman, A. 2003. Pressure driven modelling of water distribution systems. *Journal of Water, Science and Technology,* 3(1), 255-261.

Tricarico, C., De Marinis, G., Gargano, R. & Leopardi, A. 2007. Peak residential water demand. *Proceedings of the Institution of Civil Engineers: Water Management*, 160(2), 115-122.

Trifunović, N. 2006. *Introduction to urban water distribution.* London, UK: Taylor & Francis/Balkema.

Turner, R., Fowler, T., Manson, N. & Stephenson, D. 1997. *Optimisation of Rand Water's distribution system*. Pretoria, South Africa: Water Research Commission. WRC Report No. 488/1/97.

Vairavamoorthy, K., Gorantiwar, S.D. & Mohan, S. 2007. Intermittent water supply under water scarcity situations. *Water International*, 32(1), 121-132.

van Vuuren, S.J. & van Beek, J.C. 1997. *Her-evalueering van die bestaande riglyne vir stedelike en industriële watervoorsiening gebaseer op gemete waterverbruike fase 1: Pretoria Voorsieningsgebied*. Pretoria, South Africa: Water Research Commission. WRC Report No. 705/1/97.

van Zyl, H.J., van Zyl, J.E., Geustyn, L., Ilemobade, A. & Buckle, J.S. 2007. *Water consumption levels in selected South African cities*. Pretoria, South Africa: Water Research Commission. WRC Report No. 1536/1/06.

van Zyl, J.E. 1996. Peak factors in municipal water reticulation networks. *Proceedings of the Water Institution of South Africa Biennial Conference and Exhibition, 20-23 May 1996, Port Elizabeth, South Africa*.

Vorster, J., Geustyn, L., Loubser, E., Tanner, A. & Wall, K. 1995. Strategy and master plan for water supply, storage and distribution in the East Rand region. *Journal of the South African Institution of Civil Engineering*, 37(2),1-5.

White, S., Milne, G. & Riedy, C. 2004. End use analysis: issues and lessons. *Journal of Water Science and Technology: Water Supply,* 4(3), 57-65.

Wilkes, C.R., Mason, A.D. & Hern, S.C. 2005. Probability distributions for showering and bathing water-use behavior for various U.S. subpopulations. *Risk Analysis*, 25(2), 317-337.

Willis, R., Stewart, R.A., Panuwatwanich, K., Capati, B. & Giurco, D. 2009. Gold Coast domestic water end use study. *Water: Journal of the Australian Water Association*, 36(6), 79-85.

Zhang, X. 2005. *Estimating water and wastewater peaking factors with Extreme Value Theory*. Unpublished master's thesis, University of Cincinnati, USA.

Zhang, X., Buchberger, S.G. & van Zyl, J.E. 2005. A theoretical explanation for peaking factors. *Proceedings of the American Society of Civil Engineers Environmental and Water Resources Institute Conference*, *15-19 May 2005, Anchorage (AK), USA.*

APPENDIX A

A1. Goodness of Fit Tests Ranking Results

Table A1.7.1: Shower volume GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	-	-	-	-	-
Chi Squared	14	14	12	40	13
Erlang	3	3	2	8	3
Exponential	10	9	11	30	9
Gamma	4	4	4	12	4
Inverse Gaussian	9	12	9	30	9
Log Logistic	1	1	1	3	1
Log Normal	5	5	5	15	5
Log Normal2	6	6	6	18	6
Pareto	-	-	-	-	-
Pareto2	-	-	-	-	-
Pearson5	11	13	10	34	10
Pearson6	2	2	3	7	2
Rayleigh	8	8	8	24	8
Triangle	12	10	13	35	11
Uniform	13	11	14	38	12
Weibull	7	7	7	21	7

Table A1.7.2: Shower flow rate GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	-	-	-	-	-
Chi Squared	15	14	14	43	12
Erlang	4	5	6	15	5
Exponential	12	11	11	34	9
Gamma	2	6	5	13	4
Inverse Gaussian	1	7	7	15	5
Log Logistic	3	1	1	5	1
Log Normal	6	3	3	12	3
Log Normal2	7	4	4	15	5
Pareto	-	-	-	-	-
Pareto2	13	12	12	37	10
Pearson5	8	8	8	24	6
Pearson6	5	2	2	9	2
Rayleigh	10	10	10	30	8
Triangle	11	13	13	37	11
Uniform	14	15	15	44	13
Weibull	9	9	9	27	7

Table A1.7.3: Bath volume GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	3	3	4	10	3
Chi Squared	15	15	13	43	14
Erlang	5	4	5	14	4
Exponential	9	9	9	27	9
Gamma	4	5	6	15	5
Inverse Gaussian	11	11	12	34	11
Log Logistic	6	6	3	15	6
Log Normal	7	7	7	21	7
Log Normal2	8	8	8	24	8
Pareto	-	-	-	-	-
Pareto2	10	10	10	30	10
Pearson5	12	12	11	35	12
Pearson6	-	-	-	-	-
Rayleigh	1	1	1	3	1
Triangle	14	14	15	43	15
Uniform	13	13	14	40	13
Weibull	2	2	2	6	2

Table A1.7.4: Bath flow rate GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	2	2	3	7	2
Chi Squared	14	13	13	40	13
Erlang	3	4	4	11	4
Exponential	11	12	12	35	12
Gamma	4	5	5	14	5
Inverse Gaussian	9	9	9	27	9
Log Logistic	5	3	1	9	3
Log Normal	7	6	6	19	6
Log Normal2	8	7	7	22	7
Pareto	-	-	-	-	-
Pareto2	12	11	11	34	11
Pearson5	10	10	10	30	10
Pearson6	-	-	-	-	-
Rayleigh	6	8	8	22	8
Triangle	15	15	15	45	15
Uniform	13	14	14	41	14
Weibull	1	1	2	4	1

Table A1.7.5: Toilet volume GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	2	2	2	6	2
Chi Squared	5	6	8	19	6
Erlang	3	4	4	11	3
Exponential	12	12	12	36	12
Gamma	4	3	5	12	4
Inverse Gaussian	8	9	9	26	9
Log Logistic	9	5	3	17	5
Log Normal	6	7	6	19	7
Log Normal2	7	8	7	22	8
Pareto	-	-	-	-	-
Pareto2	13	13	13	39	13
Pearson5	10	10	10	30	10
Pearson6	-	-	-	-	-
Rayleigh	11	11	11	33	11
Triangle	15	15	15	45	15
Uniform	14	14	14	42	14
Weibull	1	1	1	3	1

Table A1.7.6: Toilet flow rate GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	2	13	2	17	6
Chi Squared	13	10	12	35	12
Erlang	4	2	4	10	2
Exponential	10	9	11	30	11
Gamma	3	1	3	7	1
Inverse Gaussian	8	6	8	22	8
Log Logistic	-	-	-	-	-
Log Normal	5	3	5	13	3
Log Normal2	6	4	6	16	5
Pareto	-	-	-	-	-
Pareto2	11	8	10	29	10
Pearson5	9	7	9	25	9
Pearson6	-	-	-	-	-
Rayleigh	7	5	7	19	7
Triangle	14	14	14	42	14
Uniform	12	11	13	36	13
Weibull	1	12	1	14	4

Table A1.7.7: Tap volume GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	9	10	8	27	8
Chi Squared	11	12	7	30	9
Erlang	-	-	-	-	-
Exponential	10	11	9	30	9
Gamma	-	-	-	-	-
Inverse Gaussian	1	1	3	5	1
Log Logistic	8	6	6	20	6
Log Normal	3	4	4	11	4
Log Normal2	4	5	5	14	5
Pareto	-	-	-	-	-
Pareto2	2	7	11	20	6
Pearson5	6	3	1	10	3
Pearson6	5	2	2	9	2
Rayleigh	12	13	12	37	11
Triangle	14	14	14	42	12
Uniform	13	8	13	34	10
Weibull	7	9	10	26	7

Table A1.7.8: Tap flow rate GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	-	-	-	-	-
Chi Squared	12	10	12	34	10
Erlang	2	2	6	10	2
Exponential	7	9	11	27	7
Gamma	3	3	2	8	1
Inverse Gaussian	6	7	7	20	5
Log Logistic	11	6	5	22	6
Log Normal	4	4	3	11	3
Log Normal2	5	5	4	14	4
Pareto	-	-	-	-	-
Pareto2	-	-	-	-	-
Pearson5	10	8	10	28	8
Pearson6	8	1	1	10	2
Rayleigh	9	13	8	30	9
Triangle	14	14	14	42	12
Uniform	13	11	13	37	11
Weibull	1	12	9	22	6

Table A1.7.9: Dishwasher cycle volume GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	3	12	4	19	6
Chi Squared	6	5	10	21	7
Erlang	4	13	5	22	8
Exponential	13	7	13	33	12
Gamma	2	11	2	15	3
Inverse Gaussian	10	4	9	23	9
Log Logistic	1	1	1	3	1
Log Normal	7	3	6	16	4
Log Normal2	8	3	7	18	5
Pareto	-	-	-	-	-
Pareto2	14	8	14	36	13
Pearson5	12	6	12	30	10
Pearson6	5	2	3	10	2
Rayleigh	11	15	11	37	14
Triangle	15	9	15	39	15
Uniform	16	10	16	42	16
Weibull	9	14	8	31	11

Table A1.7.10: Dishwasher cycle flow rate GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	-	-	-	-	-
Chi Squared	14	13	14	41	13
Erlang	1	1	2	4	1
Exponential	11	9	10	30	10
Gamma	3	2	3	8	2
Inverse Gaussian	2	6	7	15	5
Log Logistic	7	3	1	11	3
Log Normal	5	4	4	13	4
Log Normal2	6	5	5	16	6
Pareto	-	-	-	-	-
Pareto2	12	10	11	33	11
Pearson5	4	7	8	19	7
Pearson6	-	-	-	-	-
Rayleigh	9	8	9	26	8
Triangle	10	11	12	33	11
Uniform	13	12	13	38	12
Weibull	8	14	6	28	9

Table A1.7.11: Dishwasher duration between cycles GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	-	-	-	-	
Chi Squared	14	14	12	40	11
Erlang	2	6	5	13	3
Exponential	3	7	6	16	4
Gamma	1	9	9	19	6
Inverse Gaussian	10	10	10	30	7
Log Logistic	9	1	1	11	2
Log Normal	7	3	3	13	3
Log Normal2	8	4	4	16	4
Pareto	-	-	-	-	
Pareto2	5	5	8	18	5
Pearson5	-	-	-	-	
Pearson6	6	2	2	10	1
Rayleigh	11	11	11	33	8
Triangle	13	12	13	38	9
Uniform	12	13	14	39	10
Weibull	4	8	7	19	6

Table A1.7.12: Washing machine cycle volume GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	1	9	1	11	2
Chi Squared	12	12	10	34	9
Erlang	-	-	-	-	-
Exponential	3	6	7	16	5
Gamma	-	-	-	-	-
Inverse Gaussian	8	7	8	23	6
Log Logistic	7	2	3	12	3
Log Normal	5	3	4	12	3
Log Normal2	6	4	5	15	4
Pareto	-	-	-	-	-
Pareto2	4	5	6	15	4
Pearson5	-	-	-	-	-
Pearson6	-	-	-	-	-
Rayleigh	10	10	9	29	8
Triangle	11	11	12	34	9
Uniform	9	8	11	28	7
Weibull	2	1	2	5	1

Table A1.7.13: Washing machine cycle flow rate GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	1	15	2	18	5
Chi Squared	15	12	13	40	12
Erlang	5	3	5	13	3
Exponential	11	10	11	32	10
Gamma	4	2	4	10	2
Inverse Gaussian	9	8	9	26	8
Log Logistic	6	4	3	13	3
Log Normal	7	6	7	20	6
Log Normal2	8	7	8	23	7
Pareto	-	-	-	-	-
Pareto2	12	11	12	35	11
Pearson5	10	9	10	29	9
Pearson6	-	-	-	-	-
Rayleigh	3	5	6	14	4
Triangle	13	13	14	40	12
Uniform	14	14	15	43	13
Weibull	2	1	1	4	1

Table A1.7.14: Washing machine duration between cycle GOF ranking results

Distribution	Chi-squared	Anderson-Darling	Kolmogorov-Smirnov	Sum of score	Overall Ranking
Beta General	3	1	1	5	1
Chi Squared	13	13	11	37	13
Erlang	-	-	-	-	
Exponential	1	9	8	18	7
Gamma	-	-	-	-	
Inverse Gaussian	9	7	9	25	9
Log Logistic	7	3	3	13	3
Log Normal	5	4	4	13	4
Log Normal2	6	5	5	16	5
Pareto	-	-	-	-	
Pareto2	2	8	7	17	6
Pearson5	-	-	-	-	
Pearson6	8	6	6	20	8
Rayleigh	11	12	10	33	11
Triangle	12	11	13	36	12
Uniform	10	10	12	32	10
Weibull	4	2	2	8	2

APPENDIX B

B1. Daily Event Frequency Probability Distributions

Table B1.7.1: Shower daily frequency cumulative relative frequency

Event Frequency	1 PPH	2 PPH	3 PPH	4 PPH	5 PPH	6 PPH
0	0.0433	0.0085	0.0164	0.0176	0.0424	0.1363
1	0.6532	0.4249	0.3013	0.2550	0.2183	0.2911
2	0.8826	0.7494	0.6082	0.5373	0.4264	0.4839
3	0.9630	0.8974	0.8185	0.7447	0.6457	0.6349
4	0.9836	0.9581	0.9184	0.8783	0.8067	0.7527
5	0.9937	0.9847	0.9617	0.9432	0.8972	0.8462
6	0.9976	0.9929	0.9815	0.9770	0.9504	0.9065
7	0.9981	0.9971	0.9925	0.9904	0.9780	0.9562
8	0.9986	0.9984	0.9961	0.9951	0.9888	0.9805
9	0.9995	0.9987	0.9973	0.9977	0.9944	0.9893
10	1.0000	0.9990	0.9982	0.9981	0.9954	0.9951
11		0.9991	0.9989	0.9991	0.9974	0.9990
12		0.9992	0.9991	0.9993	0.9990	1.0000
13		0.9994	0.9995	1.0000	0.9995	
14		0.9996	0.9998		1.0000	
15		0.9997	1.0000			
16		0.9999				
17		1.0000				

Table B1.7.2: Bath daily frequency cumulative relative frequency

Event Frequency	1 PPH	2 PPH	3 PPH	4 PPH	5 PPH	6 PPH
0	0.6121	0.1282	0.2192	0.1995	0.3301	0.5304
1	0.8772	0.7654	0.7861	0.7898	0.8010	0.8198
2	0.9612	0.9297	0.9580	0.9558	0.9450	0.9393
3	0.9914	0.9829	0.9895	0.9893	0.9871	0.9757
4	1.0000	0.9924	0.9974	0.9933	0.9903	0.9899
5		1.0000	0.9987	0.9973	0.9951	0.9939
6			1.0000	0.9987	0.9984	1.0000
7				1.0000	1.0000	

Table B1.7.3: Toilet daily frequency cumulative relative frequency

Event Frequency	1 PPH	2 PPH	3 PPH	4 PPH	5 PPH	6 PPH
0	0.0200	0.0063	0.0138	0.0151	0.0390	0.1206
1	0.0644	0.0233	0.0241	0.0246	0.0452	0.1305
2	0.1416	0.0450	0.0374	0.0355	0.0495	0.1332
3	0.2377	0.0811	0.0580	0.0545	0.0590	0.1395
4	0.3346	0.1256	0.0789	0.0751	0.0705	0.1467
5	0.4338	0.1767	0.1116	0.1005	0.0814	0.1548
6	0.5132	0.2319	0.1478	0.1329	0.0976	0.1728
7	0.5932	0.2898	0.1873	0.1696	0.1195	0.1953
8	0.6573	0.3590	0.2382	0.2156	0.1457	0.2178
9	0.7209	0.4248	0.3013	0.2610	0.1771	0.2484
10	0.7692	0.4888	0.3565	0.3122	0.2138	0.2736
11	0.8164	0.5514	0.4196	0.3657	0.2619	0.3159
12	0.8537	0.6079	0.4807	0.4144	0.3105	0.3573
13	0.8825	0.6587	0.5323	0.4728	0.3581	0.3960
14	0.9089	0.7022	0.5856	0.5219	0.4110	0.4338
15	0.9234	0.7445	0.6413	0.5776	0.4605	0.4761
16	0.9397	0.7873	0.6946	0.6234	0.5100	0.5185
17	0.9522	0.8230	0.7318	0.6709	0.5748	0.5581
18	0.9631	0.8523	0.7680	0.7115	0.6352	0.5959
19	0.9697	0.8764	0.8042	0.7536	0.6829	0.6238
20	0.9747	0.8983	0.8348	0.7839	0.7219	0.6526
21	0.9792	0.9155	0.8617	0.8157	0.7571	0.6886
22	0.9817	0.9288	0.8841	0.8399	0.7895	0.7237
23	0.9836	0.9424	0.9011	0.8624	0.8157	0.7588
24	0.9856	0.9517	0.9165	0.8840	0.8452	0.7876
25	0.9867	0.9591	0.9302	0.9036	0.8695	0.8164
26	0.9883	0.9669	0.9428	0.9187	0.8890	0.8308
27	0.9900	0.9729	0.9531	0.9346	0.9033	0.8560
28	0.9908	0.9777	0.9606	0.9435	0.9138	0.8704
29	0.9931	0.9810	0.9685	0.9521	0.9290	0.8902
30	0.9953	0.9844	0.9737	0.9573	0.9405	0.9082
31	0.9958	0.9869	0.9786	0.9666	0.9471	0.9253
32	0.9961	0.9891	0.9816	0.9730	0.9600	0.9316
33	0.9964	0.9909	0.9846	0.9794	0.9686	0.9433

Event Frequency	1 PPH	2 PPH	3 PPH	4 PPH	5 PPH	6 PPH
34	0.9969	0.9924	0.9867	0.9827	0.9743	0.9523
35	0.9972	0.9932	0.9889	0.9856	0.9786	0.9568
36	0.9975	0.9944	0.9907	0.9872	0.9829	0.9667
37	0.9978	0.9955	0.9917	0.9882	0.9876	0.9721
38	0.9983	0.9961	0.9933	0.9920	0.9900	0.9748
39	0.9989	0.9964	0.9943	0.9928	0.9914	0.9829
40	0.9992	0.9970	0.9955	0.9936	0.9929	0.9847
41	0.9994	0.9973	0.9956	0.9944	0.9943	0.9865
42	0.9997	0.9975	0.9964	0.9953	0.9957	0.9883
43	1.0000	0.9977	0.9970	0.9961	0.9967	0.9892
44		0.9978	0.9974	0.9963	0.9971	0.9910
45		0.9984	0.9976	0.9971	0.9981	0.9928
46		0.9988	0.9978	0.9975	0.9986	0.9937
47		0.9990	0.9984	0.9977	0.9990	0.9946
48		0.9992	0.9986	0.9986	0.9995	0.9955
49		0.9992	0.9990	0.9988	1.0000	0.9964
50		0.9993	0.9994	0.9990		0.9973
51		0.9994	0.9998	0.9992		0.9982
52		0.9995	1.0000	0.9994		0.9991
53		0.9996		0.9996		1.0000
54		0.9997		0.9998		
55		0.9998		1.0000		
56		0.9999				
57		1.0000				

Table B1.7.4: Tap daily frequency cumulative relative frequency

Event Frequency	1 PPH	2 PPH	3 PPH	4 PPH	5 PPH	6 PPH
0	0.0134	0.0063	0.0153	0.0189	0.0571	0.1976
1	0.0239	0.0133	0.0220	0.0231	0.0641	0.2000
2	0.0331	0.0205	0.0287	0.0280	0.0667	0.2012
3	0.0445	0.0277	0.0325	0.0317	0.0712	0.2024
4	0.0650	0.0349	0.0378	0.0338	0.0756	0.2036
5	0.0819	0.0438	0.0451	0.0375	0.0763	0.2060
6	0.1027	0.0549	0.0518	0.0425	0.0840	0.2096
7	0.1225	0.0672	0.0598	0.0503	0.0846	0.2132
8	0.1447	0.0803	0.0675	0.0556	0.0878	0.2180
9	0.1710	0.0948	0.0791	0.0616	0.0910	0.2251
10	0.1955	0.1091	0.0906	0.0687	0.0936	0.2311
11	0.2223	0.1229	0.1013	0.0779	0.0955	0.2395
12	0.2449	0.1368	0.1147	0.0875	0.1019	0.2455
13	0.2704	0.1519	0.1277	0.0980	0.1083	0.2563
14	0.2914	0.1700	0.1407	0.1109	0.1154	0.2575
15	0.3122	0.1883	0.1563	0.1240	0.1237	0.2671
16	0.3377	0.2057	0.1707	0.1363	0.1301	0.2790
17	0.3620	0.2235	0.1837	0.1476	0.1436	0.2850
18	0.3816	0.2418	0.2024	0.1651	0.1609	0.2922
19	0.4034	0.2583	0.2229	0.1785	0.1660	0.2994
20	0.4227	0.2788	0.2399	0.1948	0.1763	0.3090
21	0.4454	0.3004	0.2609	0.2087	0.1872	0.3269
22	0.4686	0.3184	0.2831	0.2252	0.1974	0.3353
23	0.4900	0.3391	0.3005	0.2414	0.2096	0.3461
24	0.5098	0.3594	0.3209	0.2585	0.2212	0.3557
25	0.5314	0.3789	0.3393	0.2734	0.2346	0.3749
26	0.5483	0.3979	0.3565	0.2891	0.2462	0.3844
27	0.5666	0.4181	0.3750	0.3046	0.2583	0.4024
28	0.5843	0.4368	0.3956	0.3269	0.2737	0.4180
29	0.6003	0.4564	0.4119	0.3457	0.2929	0.4407
30	0.6203	0.4752	0.4308	0.3633	0.3122	0.4575
31	0.6372	0.4929	0.4516	0.3840	0.3256	0.4719
32	0.6540	0.5075	0.4690	0.4034	0.3487	0.4850
33	0.6639	0.5283	0.4878	0.4181	0.3744	0.4982

Event Frequency	1 PPH	2 PPH	3 PPH	4 PPH	5 PPH	6 PPH
34	0.6757	0.5465	0.5057	0.4364	0.3910	0.5102
35	0.6892	0.5633	0.5185	0.4509	0.4083	0.5210
36	0.7034	0.5792	0.5325	0.4684	0.4212	0.5305
37	0.7166	0.5962	0.5474	0.4841	0.4340	0.5473
38	0.7250	0.6136	0.5638	0.5030	0.4506	0.5593
39	0.7370	0.6302	0.5816	0.5208	0.4712	0.5665
40	0.7446	0.6435	0.5927	0.5429	0.4853	0.5808
41	0.7547	0.6576	0.6074	0.5578	0.5006	0.5880
42	0.7644	0.6703	0.6231	0.5751	0.5186	0.5988
43	0.7767	0.6837	0.6370	0.5887	0.5340	0.6084
44	0.7858	0.6970	0.6479	0.6058	0.5532	0.6204
45	0.7956	0.7076	0.6586	0.6202	0.5692	0.6323
46	0.8065	0.7208	0.6722	0.6354	0.5853	0.6419
47	0.8160	0.7316	0.6848	0.6511	0.5987	0.6551
48	0.8236	0.7442	0.6959	0.6682	0.6179	0.6683
49	0.8308	0.7540	0.7066	0.6815	0.6359	0.6778
50	0.8380	0.7661	0.7171	0.6938	0.6545	0.6922
51	0.8434	0.7774	0.7269	0.7106	0.6699	0.7042
52	0.8514	0.7887	0.7362	0.7211	0.6891	0.7186
53	0.8601	0.7974	0.7500	0.7358	0.7006	0.7257
54	0.8662	0.8070	0.7588	0.7465	0.7103	0.7341
55	0.8718	0.8151	0.7693	0.7586	0.7263	0.7461
56	0.8767	0.8229	0.7792	0.7706	0.7449	0.7581
57	0.8821	0.8321	0.7861	0.7822	0.7538	0.7665
58	0.8885	0.8402	0.7932	0.7924	0.7635	0.7689
59	0.8934	0.8461	0.8039	0.8039	0.7737	0.7808
60	0.8987	0.8516	0.8125	0.8142	0.7872	0.7844
61	0.9043	0.8599	0.8201	0.8244	0.7981	0.7916
62	0.9092	0.8665	0.8284	0.8341	0.8103	0.8060
63	0.9131	0.8721	0.8366	0.8417	0.8179	0.8144
64	0.9187	0.8783	0.8429	0.8493	0.8327	0.8228
65	0.9232	0.8840	0.8507	0.8582	0.8429	0.8335
66	0.9280	0.8888	0.8572	0.8650	0.8500	0.8455
67	0.9331	0.8936	0.8660	0.8718	0.8609	0.8515
68	0.9376	0.8984	0.8719	0.8792	0.8667	0.8611

Event Frequency	1 PPH	2 PPH	3 PPH	4 PPH	5 PPH	6 PPH
69	0.9405	0.9037	0.8792	0.8862	0.8718	0.8671
70	0.9446	0.9093	0.8867	0.8925	0.8776	0.8743
71	0.9467	0.9143	0.8935	0.8986	0.8859	0.8778
72	0.9490	0.9196	0.8995	0.9041	0.8942	0.8874
73	0.9516	0.9249	0.9056	0.9088	0.9026	0.8994
74	0.9557	0.9294	0.9111	0.9117	0.9096	0.9030
75	0.9578	0.9352	0.9153	0.9166	0.9179	0.9102
76	0.9617	0.9389	0.9203	0.9219	0.9231	0.9126
77	0.9636	0.9424	0.9262	0.9271	0.9269	0.9162
78	0.9673	0.9457	0.9304	0.9308	0.9321	0.9222
79	0.9695	0.9490	0.9327	0.9358	0.9372	0.9269
80	0.9714	0.9531	0.9371	0.9389	0.9417	0.9341
81	0.9735	0.9567	0.9409	0.9421	0.9455	0.9365
82	0.9753	0.9605	0.9451	0.9463	0.9481	0.9389
83	0.9763	0.9646	0.9480	0.9505	0.9526	0.9485
84	0.9778	0.9668	0.9507	0.9549	0.9558	0.9545
85	0.9788	0.9708	0.9547	0.9573	0.9590	0.9557
86	0.9794	0.9730	0.9606	0.9620	0.9622	0.9581
87	0.9805	0.9748	0.9637	0.9649	0.9654	0.9617
88	0.9825	0.9772	0.9685	0.9670	0.9679	0.9629
89	0.9858	0.9794	0.9711	0.9688	0.9712	0.9665
90	0.9874	0.9820	0.9750	0.9730	0.9724	0.9725
91	0.9891	0.9843	0.9773	0.9748	0.9744	0.9760
92	0.9907	0.9862	0.9801	0.9782	0.9769	0.9820
93	0.9932	0.9883	0.9826	0.9817	0.9795	0.9856
94	0.9938	0.9902	0.9847	0.9861	0.9840	0.9868
95	0.9951	0.9923	0.9872	0.9893	0.9853	0.9892
96	0.9965	0.9938	0.9910	0.9927	0.9891	0.9904
97	0.9977	0.9952	0.9933	0.9940	0.9923	0.9940
98	0.9984	0.9970	0.9956	0.9963	0.9949	0.9964
99	0.9992	0.9985	0.9977	0.9979	0.9974	0.9976
100	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table B1.7.5: Dishwasher daily frequency cumulative relative frequency

Event Frequency	1 PPH	2 PPH	3 PPH	4 PPH	5 PPH	6 PPH
0	0.3779	0.0324	0.0781	0.0628	0.1722	0.4654
1	0.9666	0.9425	0.9437	0.9298	0.9004	0.9402
2	0.9982	0.9947	0.9947	0.9939	0.9865	0.9944
3	1.0000	1.0000	0.9985	1.0000	0.9975	1.0000
4			1.0000		0.9988	
5					1.0000	

Table B1.7.6: Washing machine daily frequency cumulative relative frequency

Event Frequency	1 PPH	2 PPH	3 PPH	4 PPH	5 PPH	6 PPH
0	0.1174	0.0166	0.0310	0.0298	0.0728	0.1868
1	0.6385	0.5074	0.4466	0.4214	0.3949	0.4188
2	0.8486	0.7505	0.6886	0.6884	0.6456	0.6265
3	0.9422	0.8757	0.8319	0.8340	0.7943	0.7552
4	0.9725	0.9453	0.9055	0.9106	0.8919	0.8654
5	0.9890	0.9781	0.9516	0.9555	0.9467	0.9188
6	0.9963	0.9900	0.9717	0.9759	0.9685	0.9524
7	0.9991	0.9940	0.9899	0.9883	0.9857	0.9733
8	1.0000	0.9978	0.9961	0.9959	0.9917	0.9884
9		0.9993	0.9977	0.9989	0.9977	0.9919
10		0.9998	0.9981	1.0000	0.9992	0.9942
11		1.0000	0.9988		1.0000	0.9965
12			0.9996			0.9988
13			1.0000			1.0000

B2. Starting Hour and Number of Cycles Probability Distributions

Table B2.1: Starting hour cumulative relative frequency

Hour	Bath	Shower	Toilet	Tap	Dishwasher	Washing machine
0	0.011	0.010	0.024	0.014	0.025	0.006
1	0.016	0.014	0.039	0.022	0.035	0.010
2	0.019	0.017	0.052	0.028	0.042	0.012
3	0.023	0.022	0.064	0.034	0.045	0.013
4	0.028	0.034	0.079	0.042	0.049	0.015
5	0.046	0.084	0.106	0.060	0.057	0.021
6	0.083	0.193	0.157	0.099	0.081	0.040
7	0.133	0.305	0.222	0.158	0.122	0.088
8	0.195	0.400	0.282	0.221	0.179	0.163
9	0.251	0.475	0.336	0.280	0.233	0.254
10	0.299	0.534	0.385	0.333	0.284	0.344
11	0.330	0.580	0.428	0.384	0.328	0.426
12	0.358	0.615	0.471	0.435	0.373	0.495
13	0.379	0.644	0.513	0.482	0.418	0.559
14	0.403	0.669	0.554	0.525	0.455	0.615
15	0.426	0.695	0.597	0.569	0.485	0.671
16	0.463	0.727	0.645	0.620	0.526	0.725
17	0.511	0.767	0.696	0.687	0.579	0.778
18	0.584	0.812	0.750	0.763	0.661	0.831
19	0.699	0.857	0.801	0.829	0.752	0.884
20	0.818	0.902	0.852	0.884	0.828	0.930
21	0.902	0.945	0.907	0.934	0.901	0.966
22	0.964	0.977	0.960	0.973	0.960	0.989
23	1.000	1.000	1.000	1.000	1.000	1.000

Table B2.2: Number of cycles cumulative relative frequency

Number of Cycles	Washing machine	Dishwasher
1	0.0069	0.0000
2	0.1045	0.0040
3	0.4506	0.0527
4	0.7253	0.3196
5	0.8282	0.6033
6	0.8721	0.8338
7	0.9006	0.9528
8	0.9209	0.9865
9	0.9674	0.9951
10	0.9896	0.9987
11	0.9947	0.9996
12	0.9960	1.0000
13	0.9974	
14	0.9982	
15	0.9989	
16	0.9992	
17	0.9995	
18	0.9997	
19	0.9999	
20	1.0000	

APPENDIX C

C1. Peak Factor Plots for All Scenarios

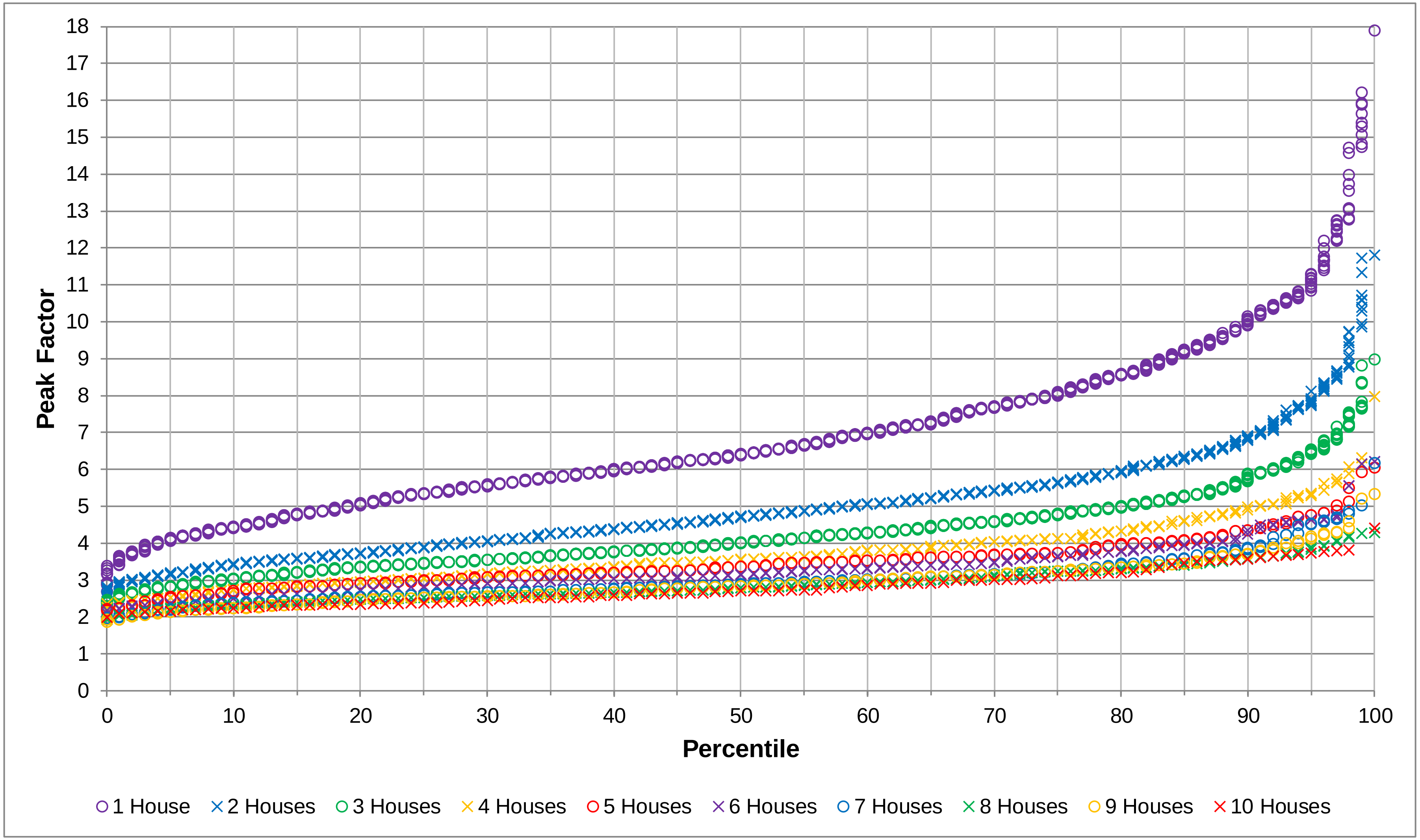


Figure C1.1: Result set for 60 minute peak factor and household group sizes ranging from 1 to 10

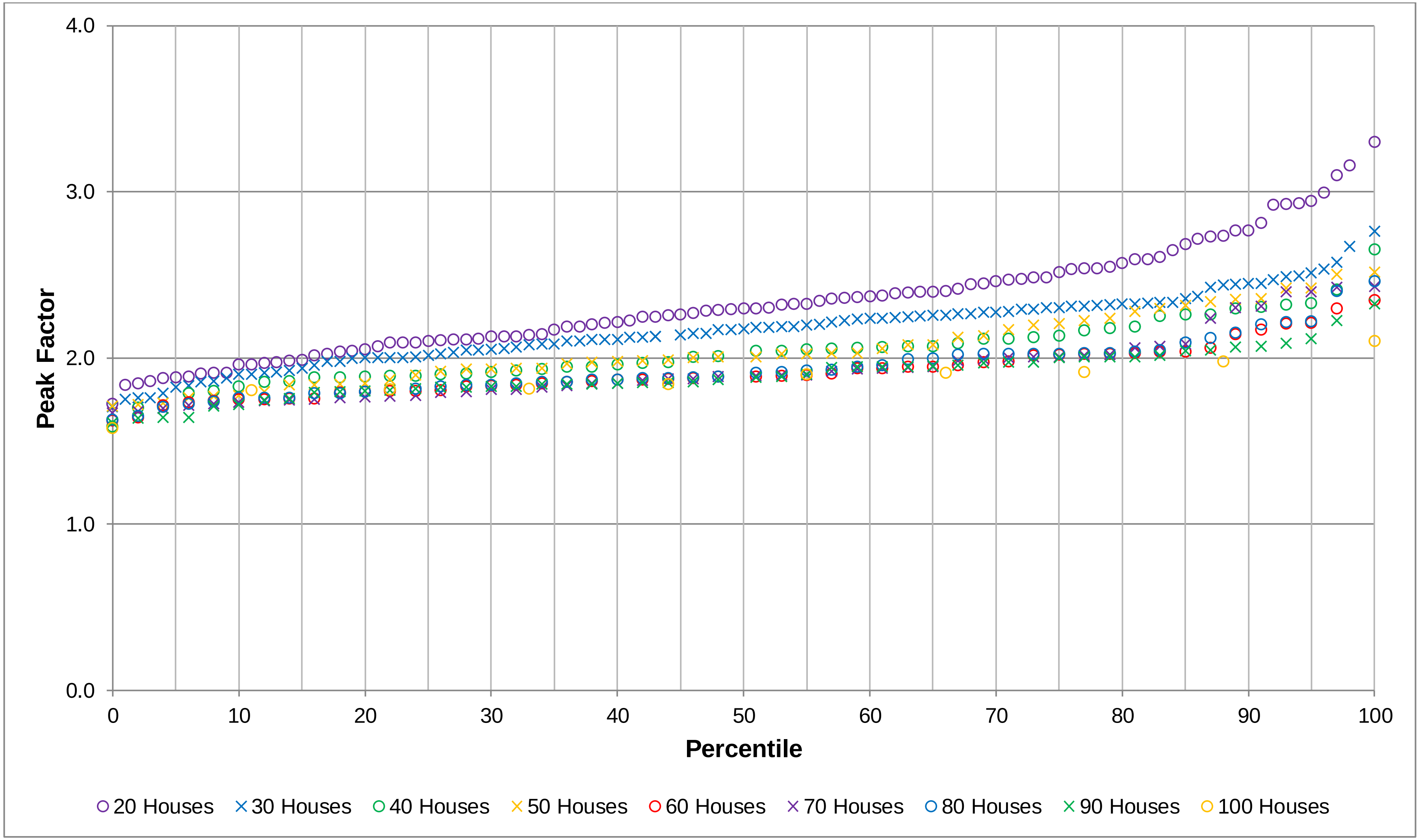


Figure C1.2: Result set for 60 minute peak factor and household group sizes ranging from 20 to 100

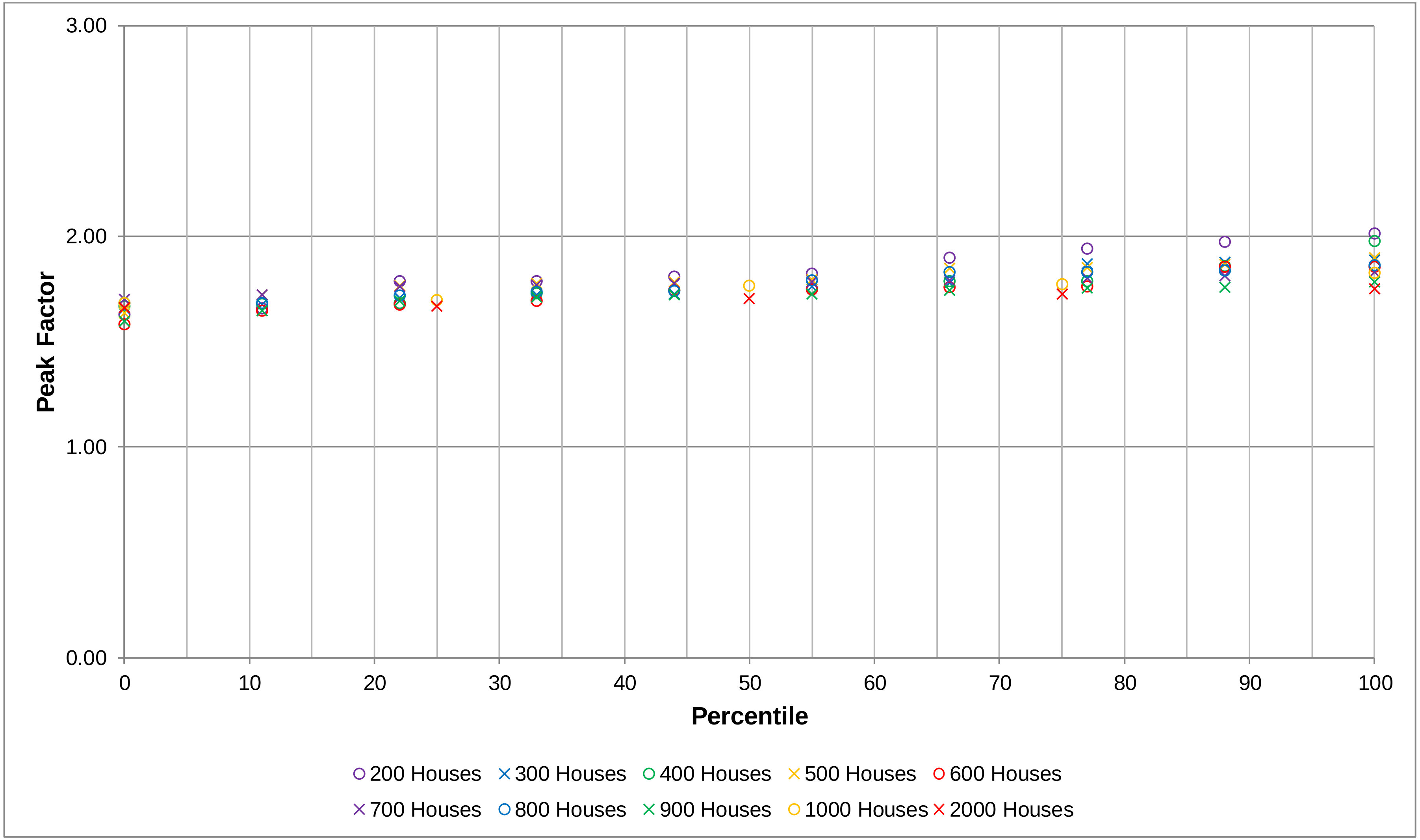


Figure C1.3: Result set for 60 minute peak factor and household group sizes ranging from 200 to 2000

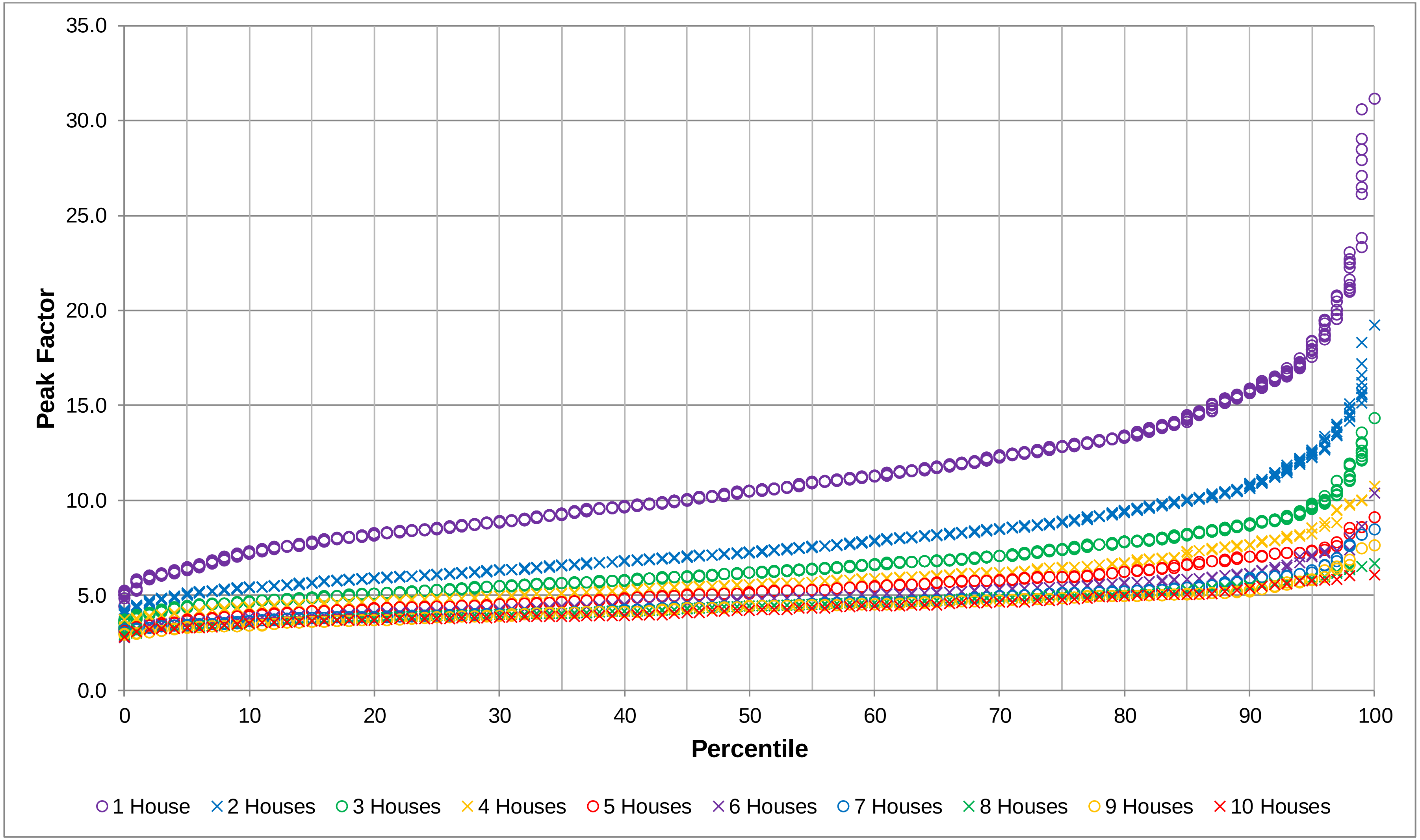


Figure C1.4: Result set for 30 minute peak factor and household group sizes ranging from 1 to 10

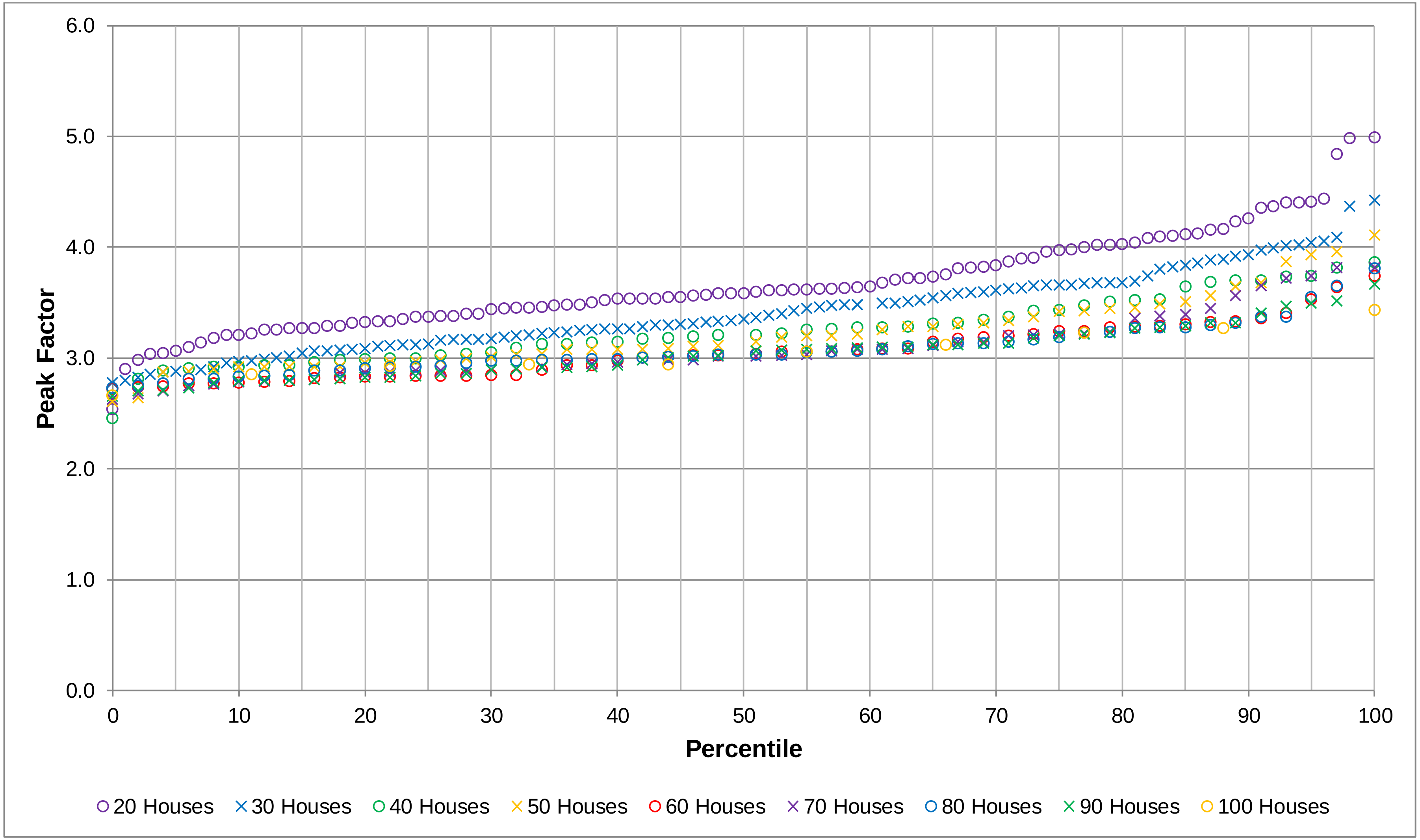


Figure C1.5: Result set for 30 minute peak factor and household group sizes ranging from 20 to 100

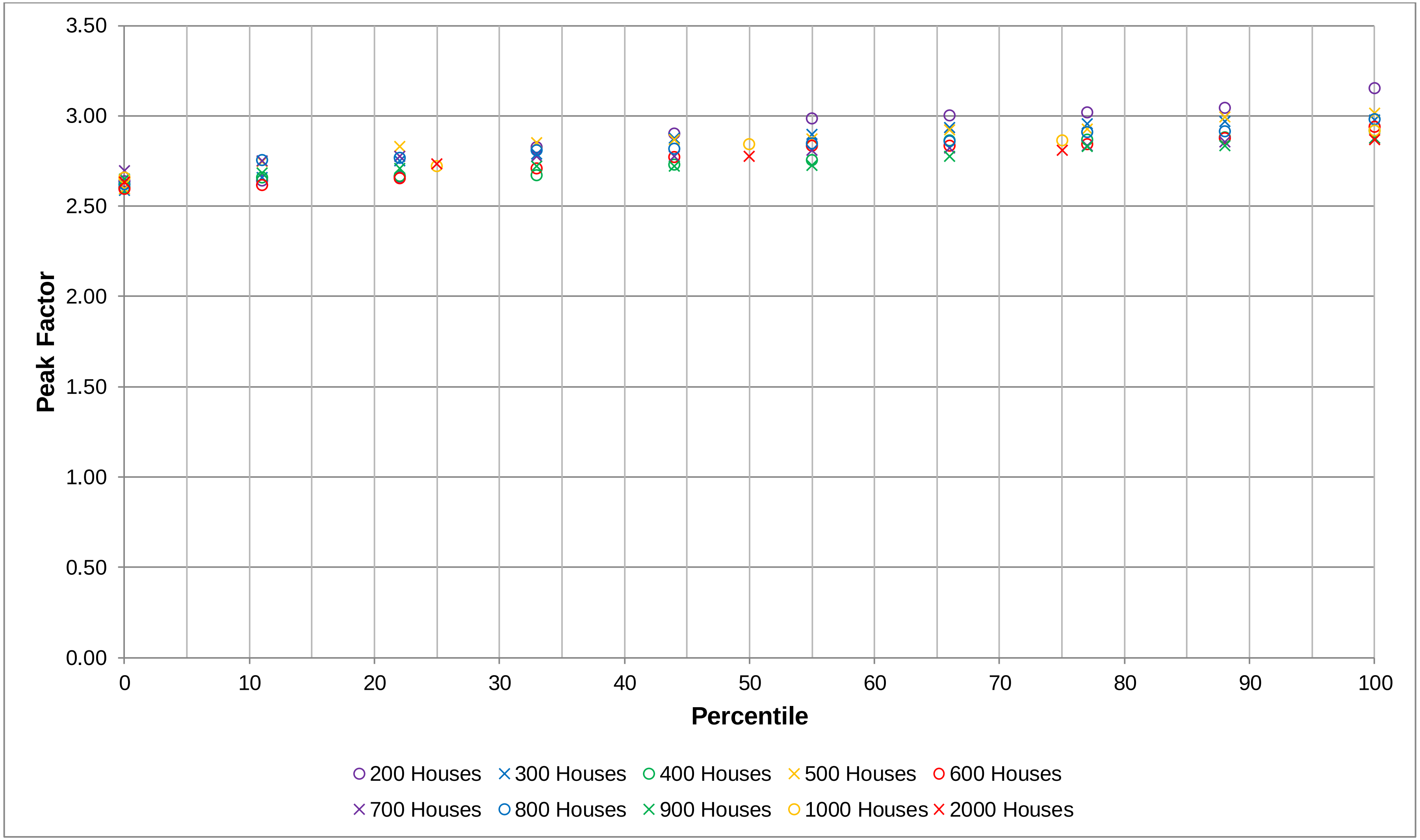


Figure C1.6: Result set for 30 minute peak factor and household group sizes ranging from 200 to 2000

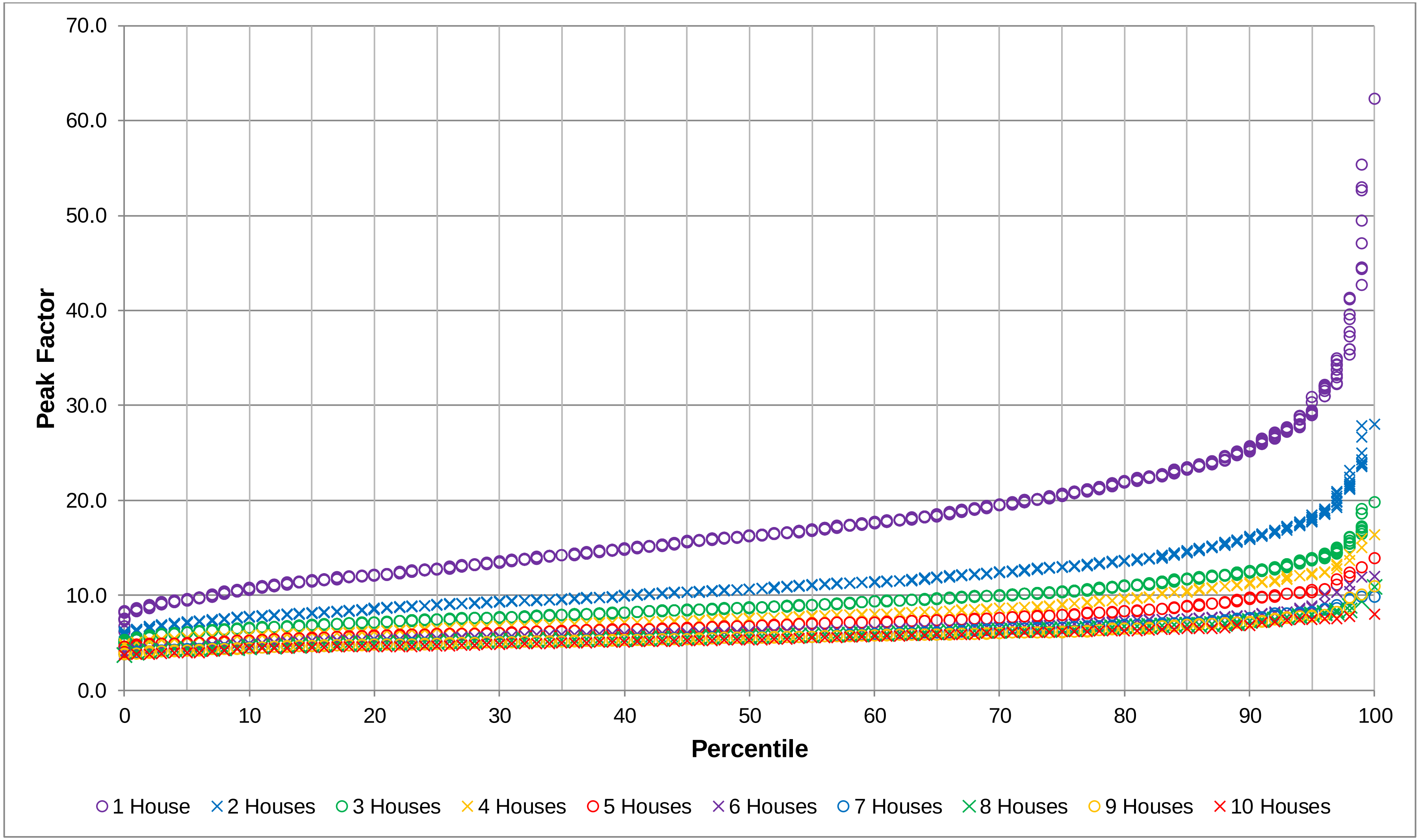


Figure C1.7: Result set for 15 minute peak factor and household group sizes ranging from 1 to 10

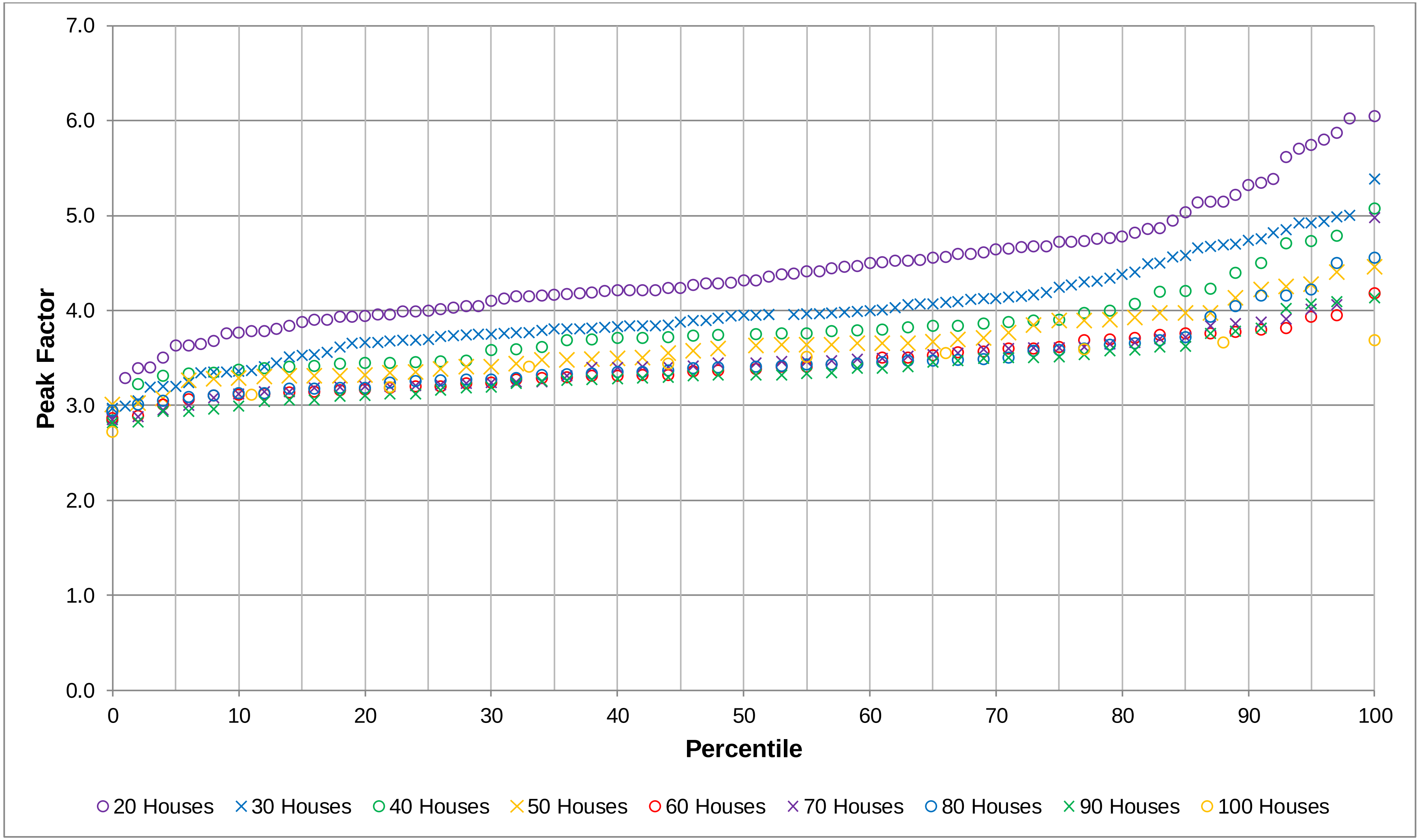


Figure C1.8: Result set for 15 minute peak factor and household group sizes ranging from 20 to 100

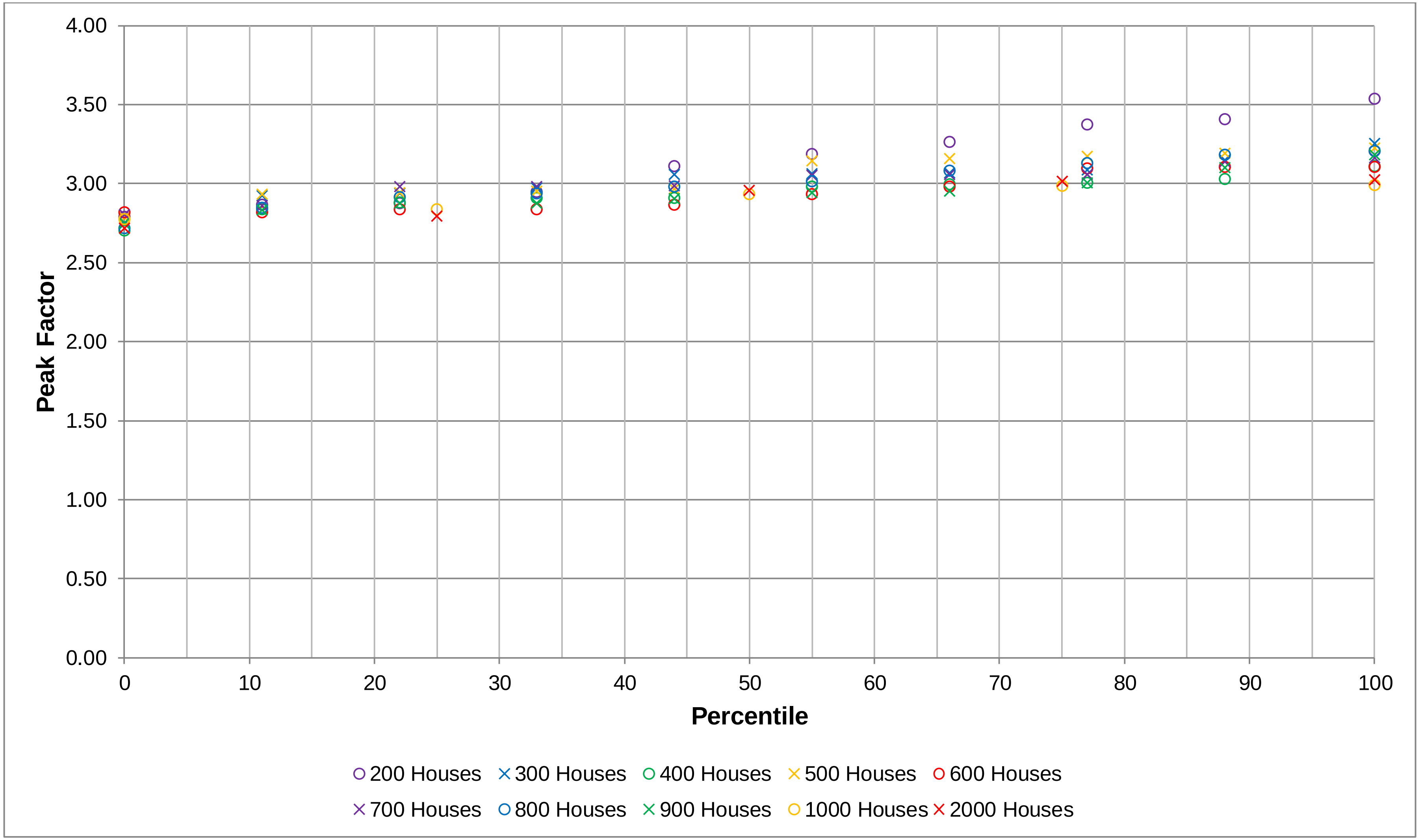


Figure C1.9: Result set for 15 minute peak factor and household group sizes ranging from 200 to 2000

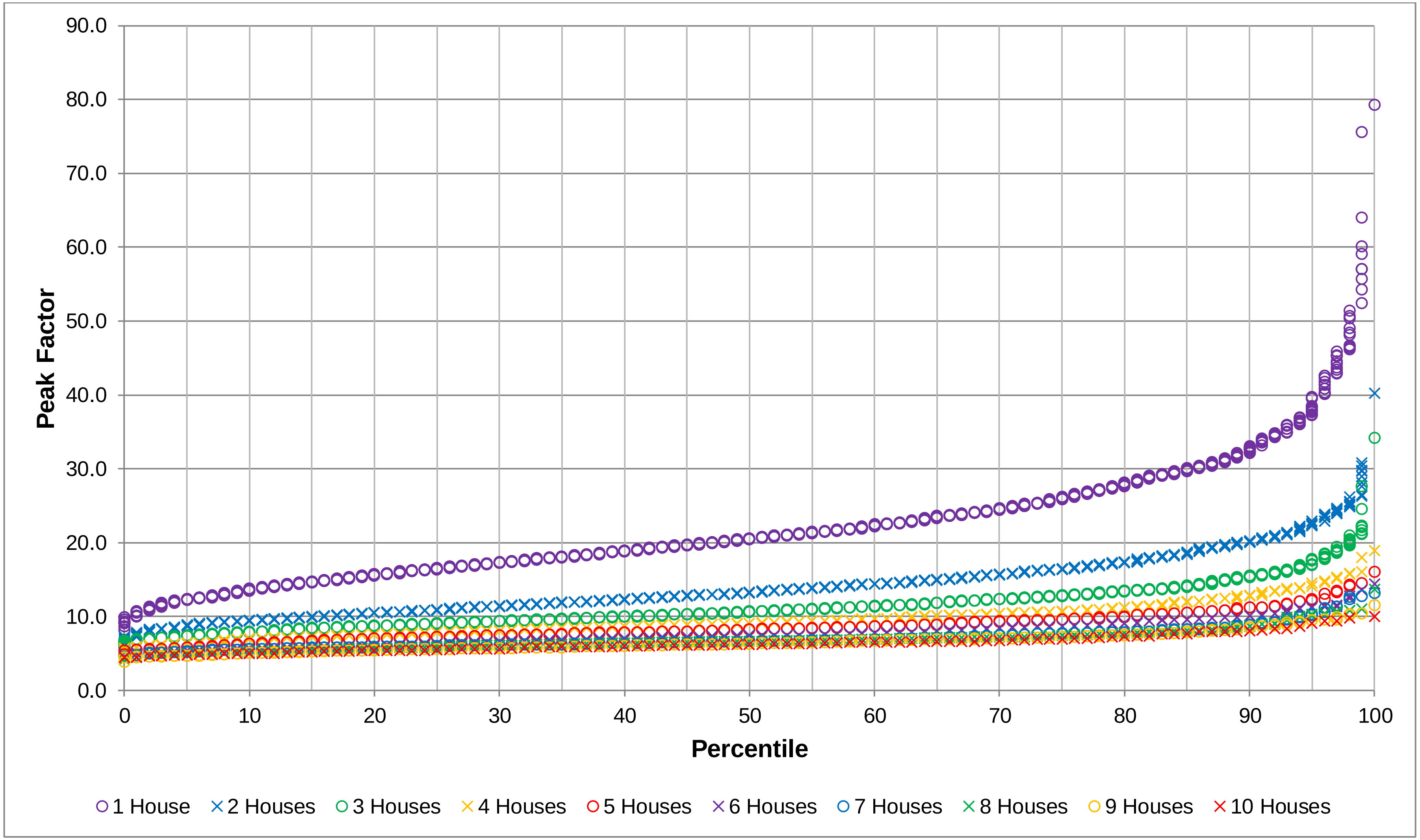


Figure C1.10: Result set for 10 minute peak factor and household group sizes ranging from 1 to 10

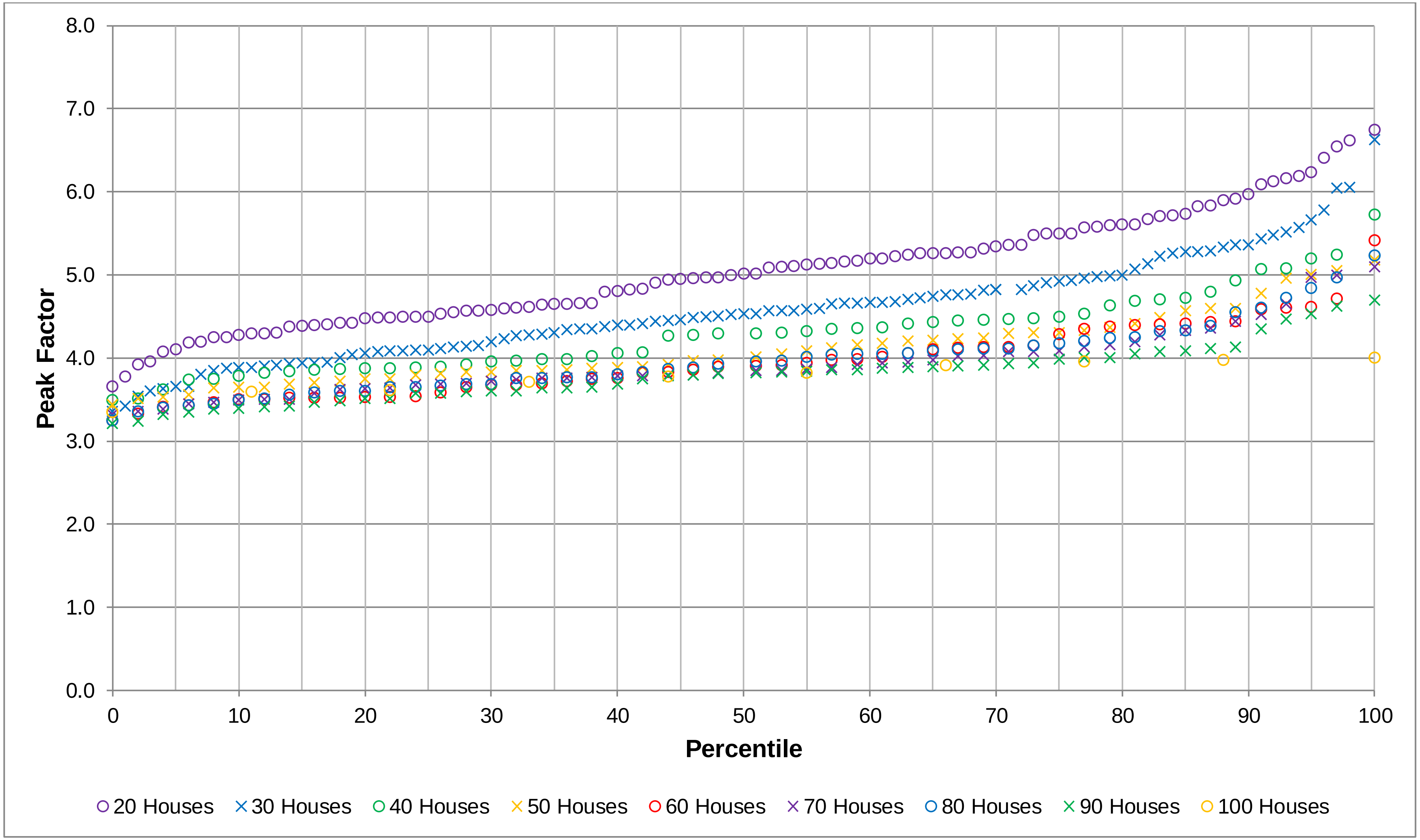


Figure C1.11: Result set for 10 minute peak factor and household group sizes ranging from 20 to 100

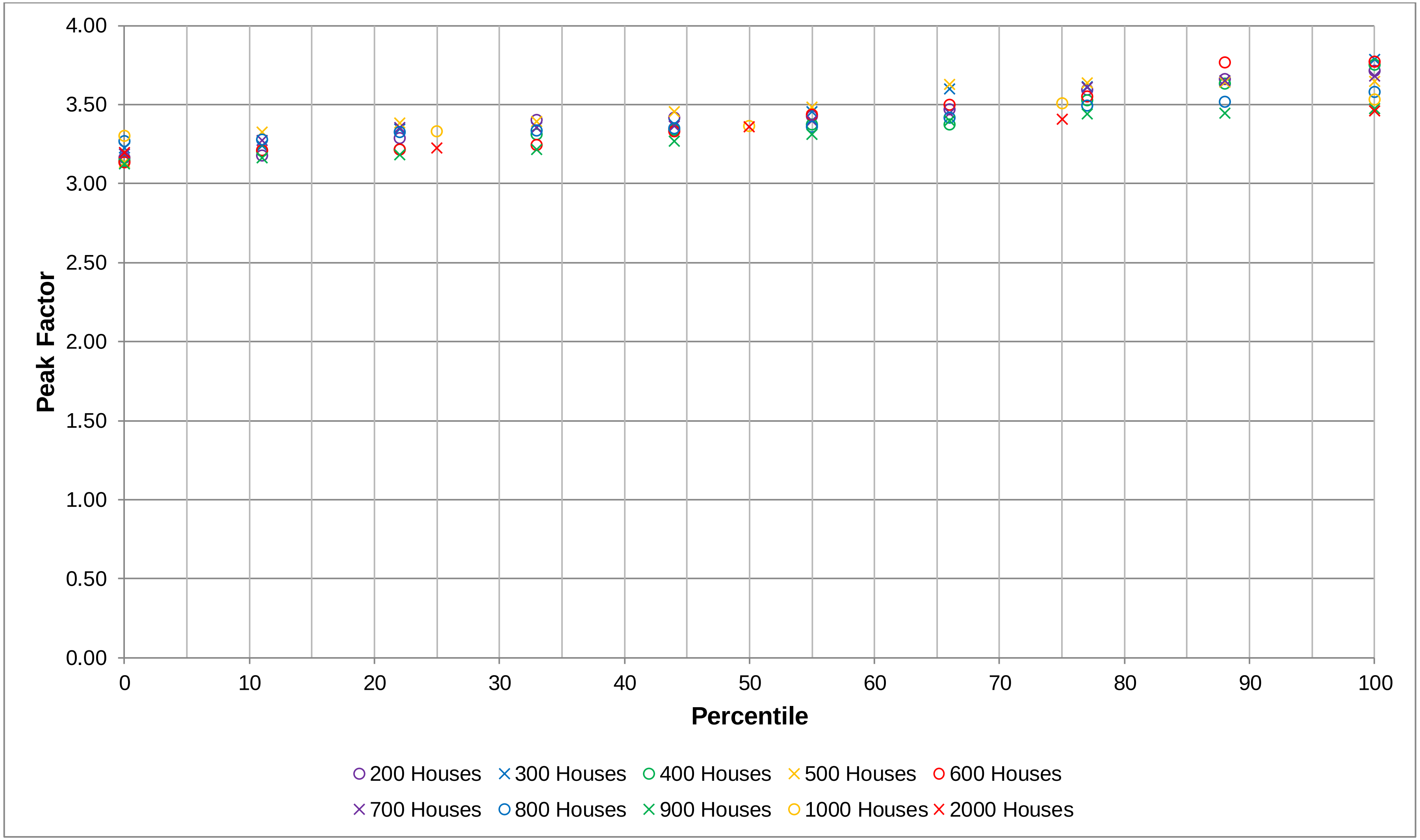


Figure C1.12: Result set for 10 minute peak factor and household group sizes ranging from 200 to 2000

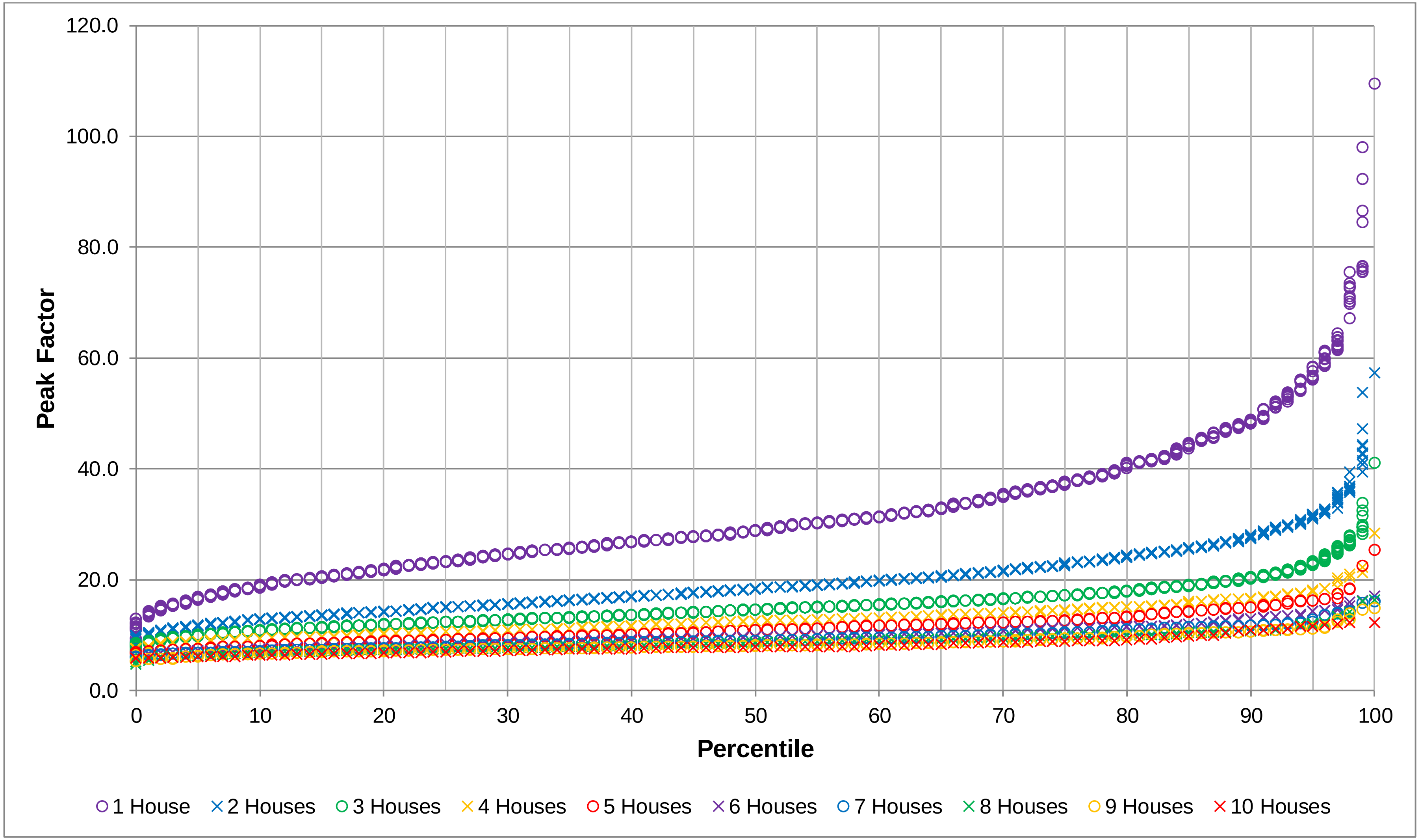


Figure C1.13: Result set for 5 minute peak factor and household group sizes ranging from 1 to 10

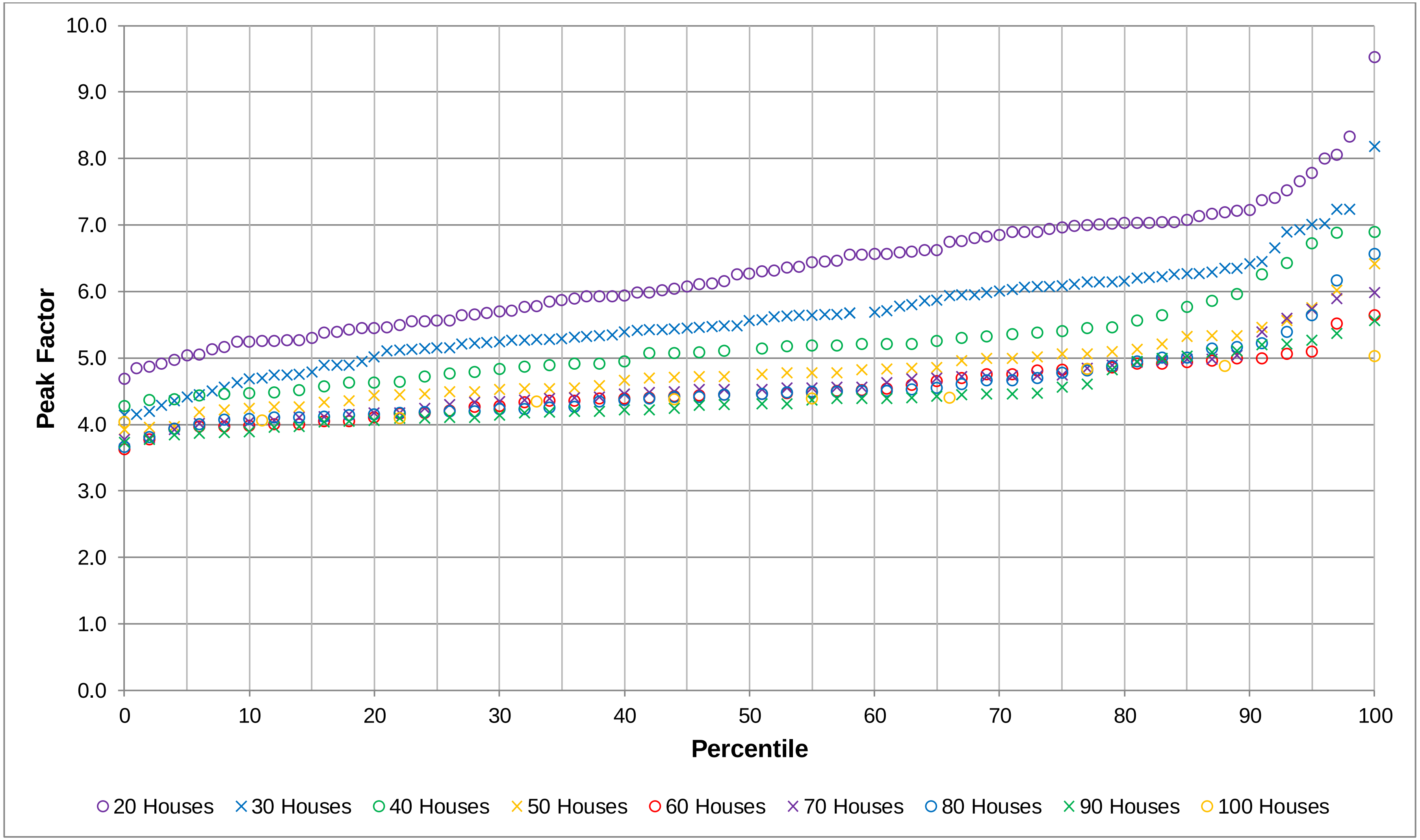


Figure C1.14: Result set for 5 minute peak factor and household group sizes ranging from 20 to 100

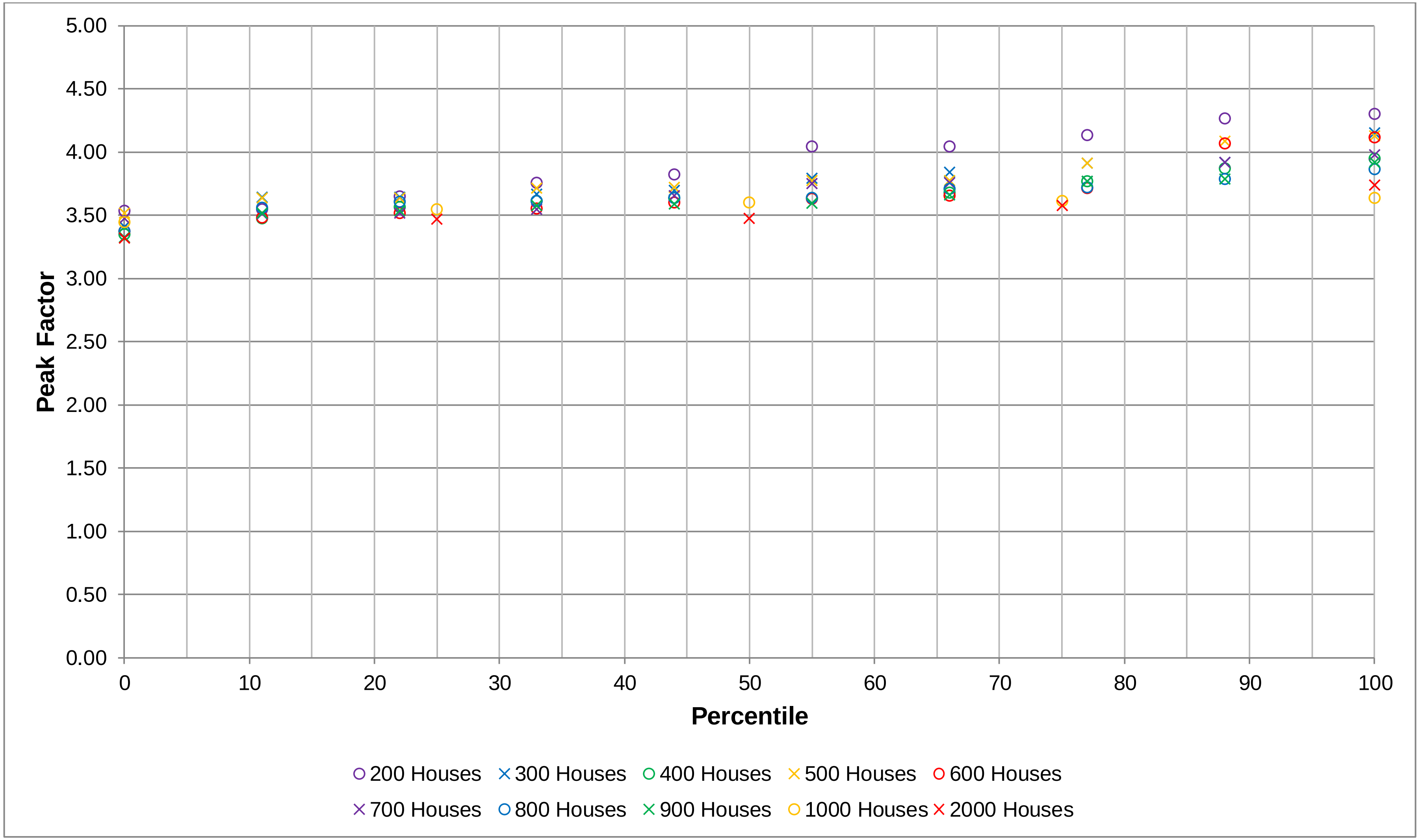


Figure C1.15: Result set for 5 minute peak factor and household group sizes ranging from 200 to 2000

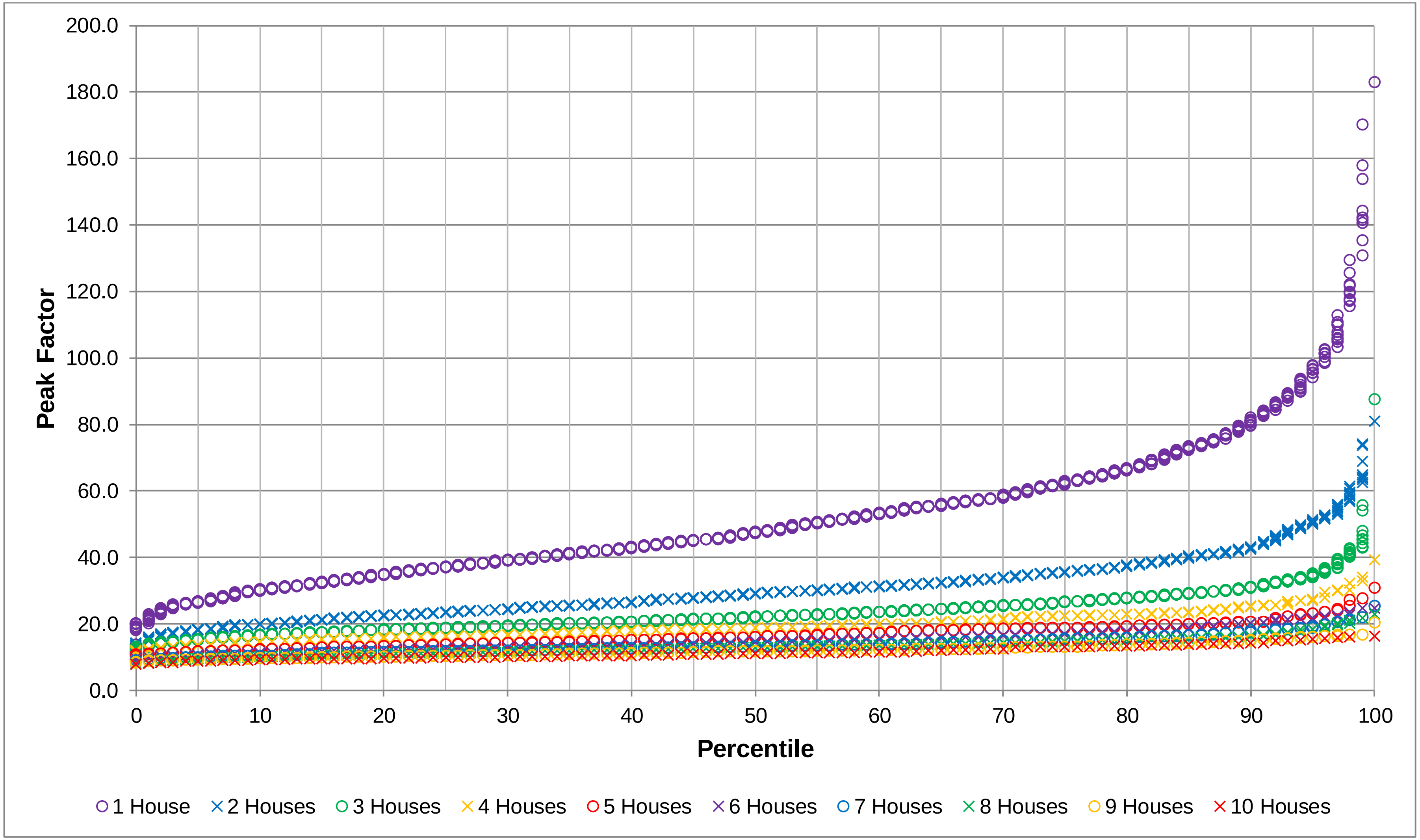


Figure C1.16: Result set for 1 minute peak factor and household group sizes ranging from 1 to 10

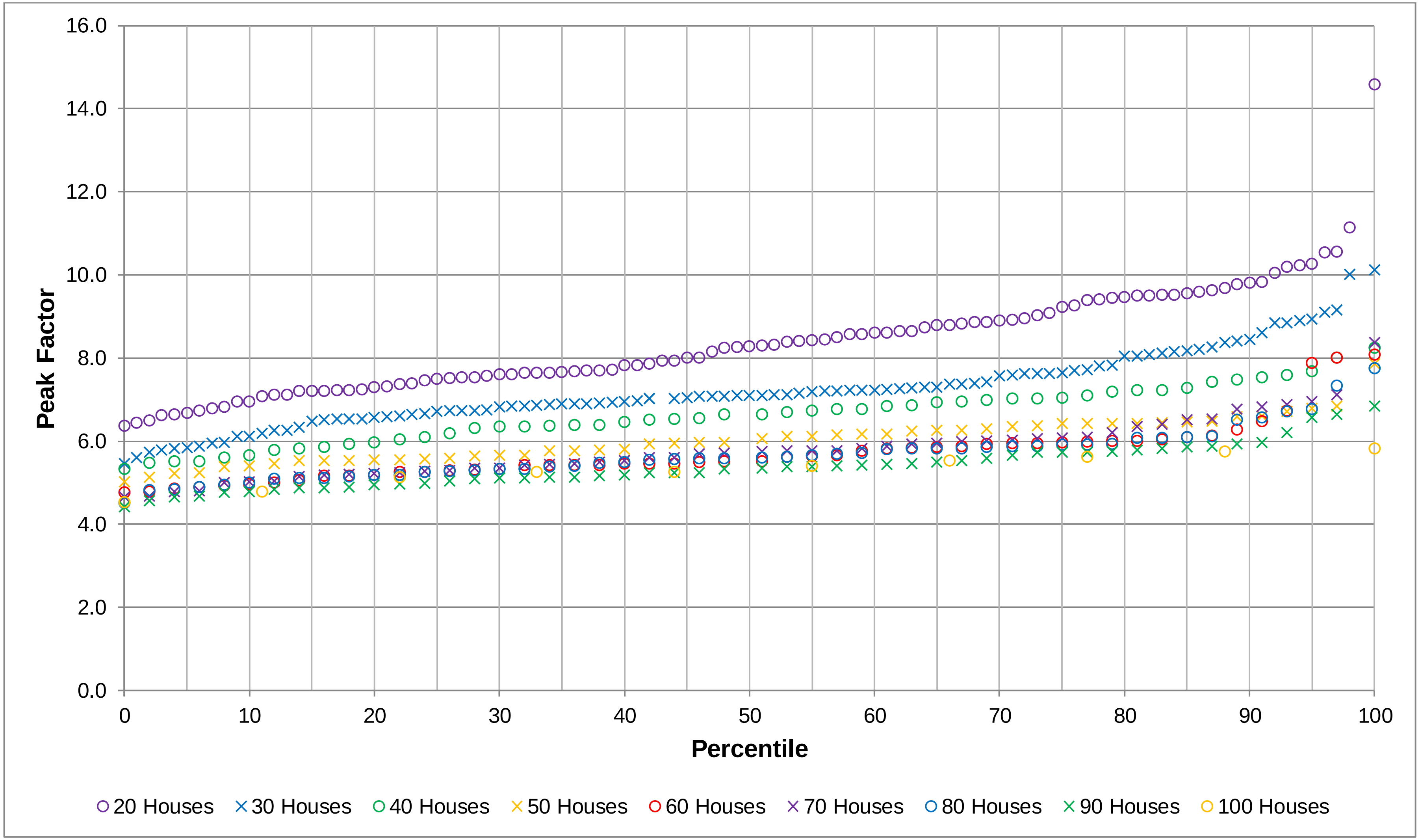


Figure C1.17: Result set for 1 minute peak factor and household group sizes ranging from 20 to 100

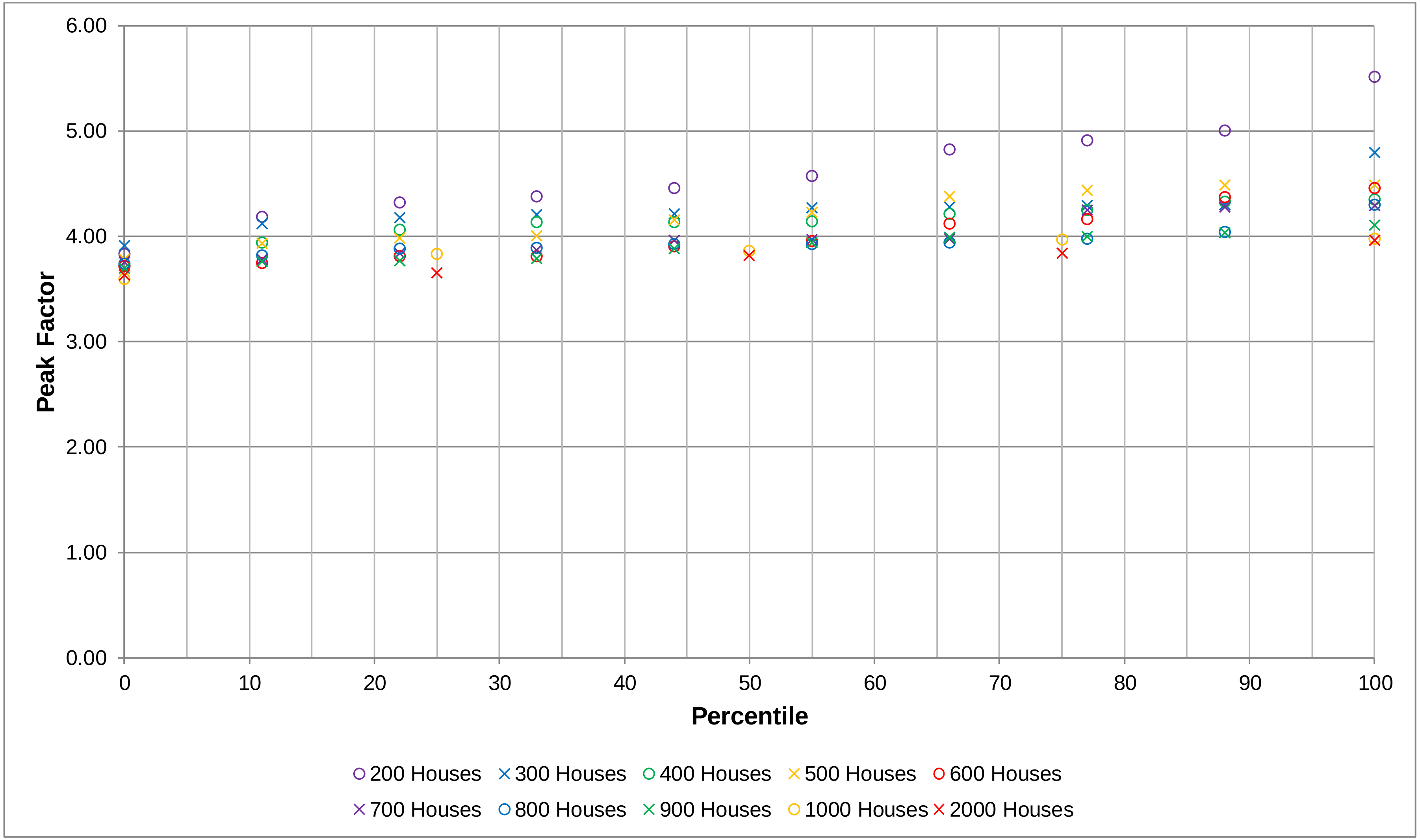


Figure C1.18: Result set for 1 minute peak factor and household group sizes ranging from 200 to 2000

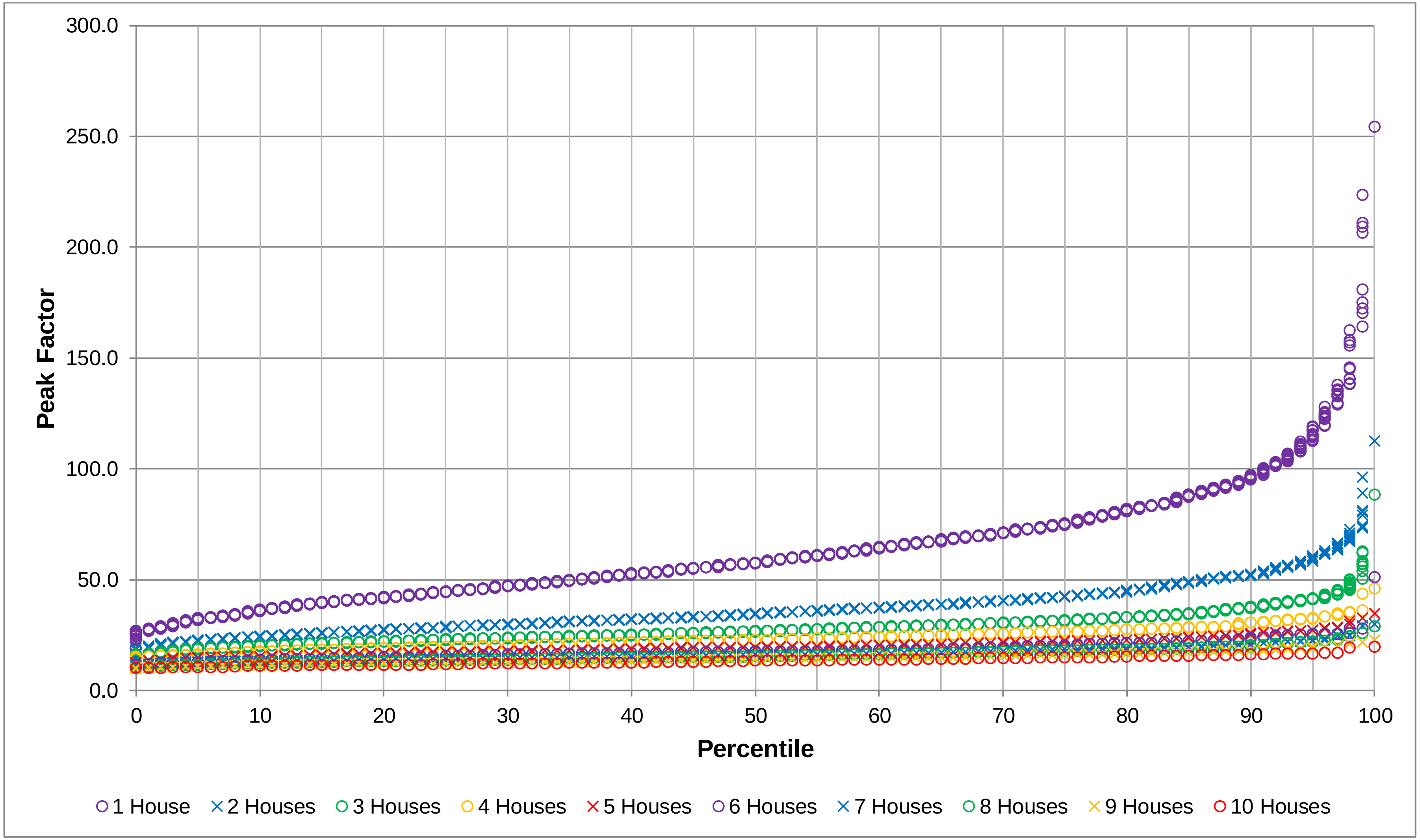


Figure C1.19: Result set for 10 second peak factor and household group sizes ranging from 1 to 10

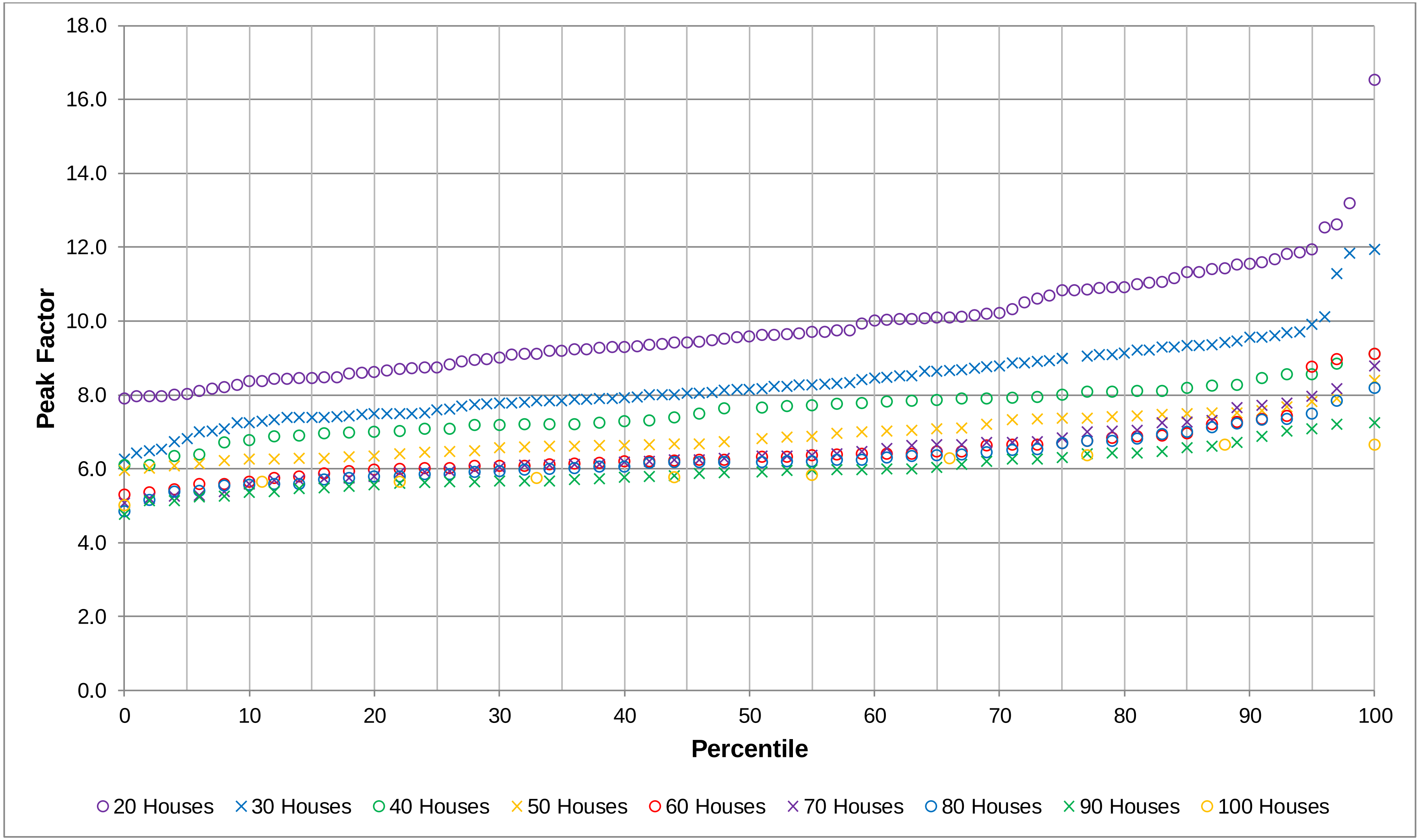


Figure C1.20: Result set for 10 second peak factor and household group sizes ranging from 20 to 100

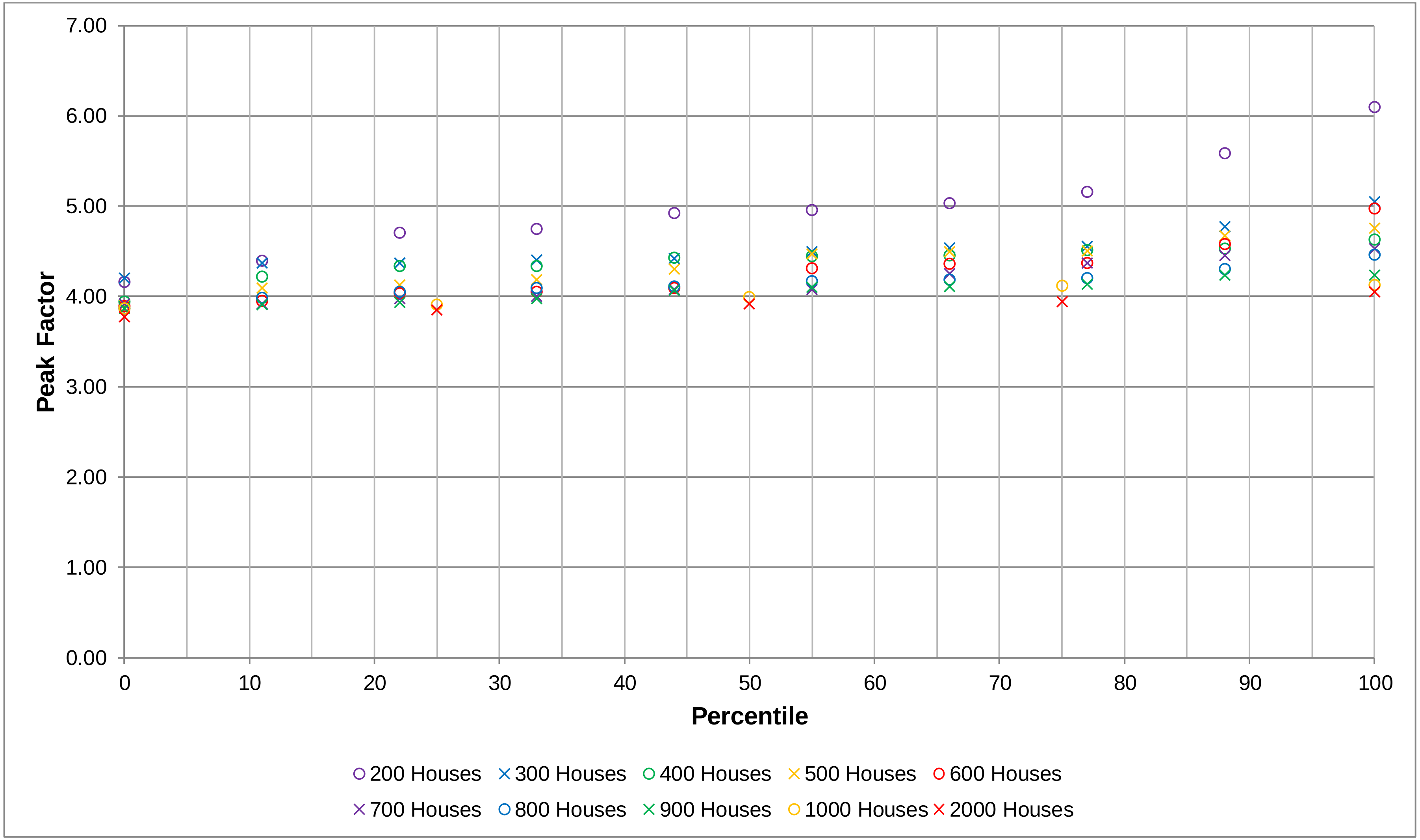


Figure C1.21: Result set for 10 second peak factor and household group sizes ranging from 200 to 2000

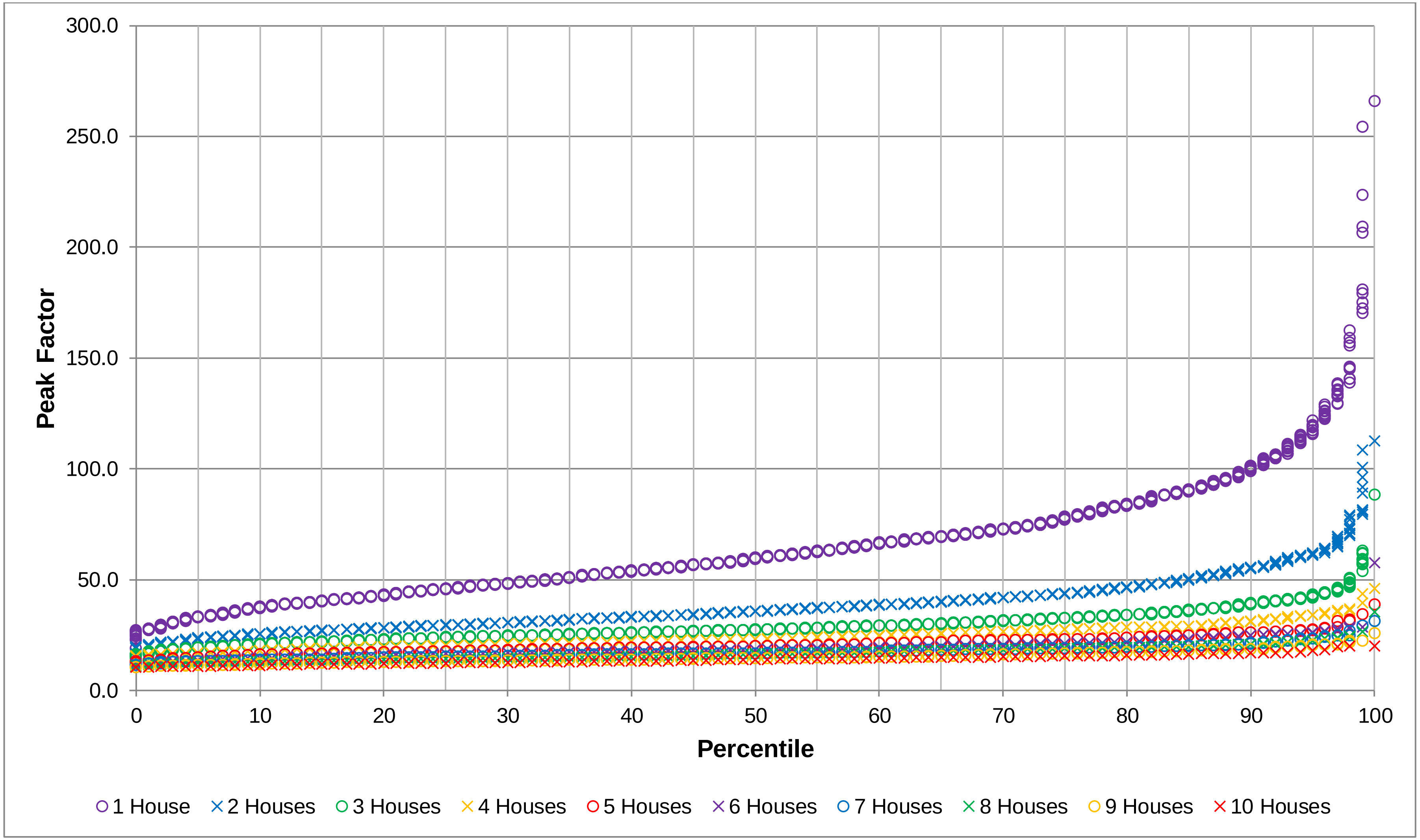


Figure C1.22: Result set for 1 second peak factor and household group sizes ranging from 1 to 10

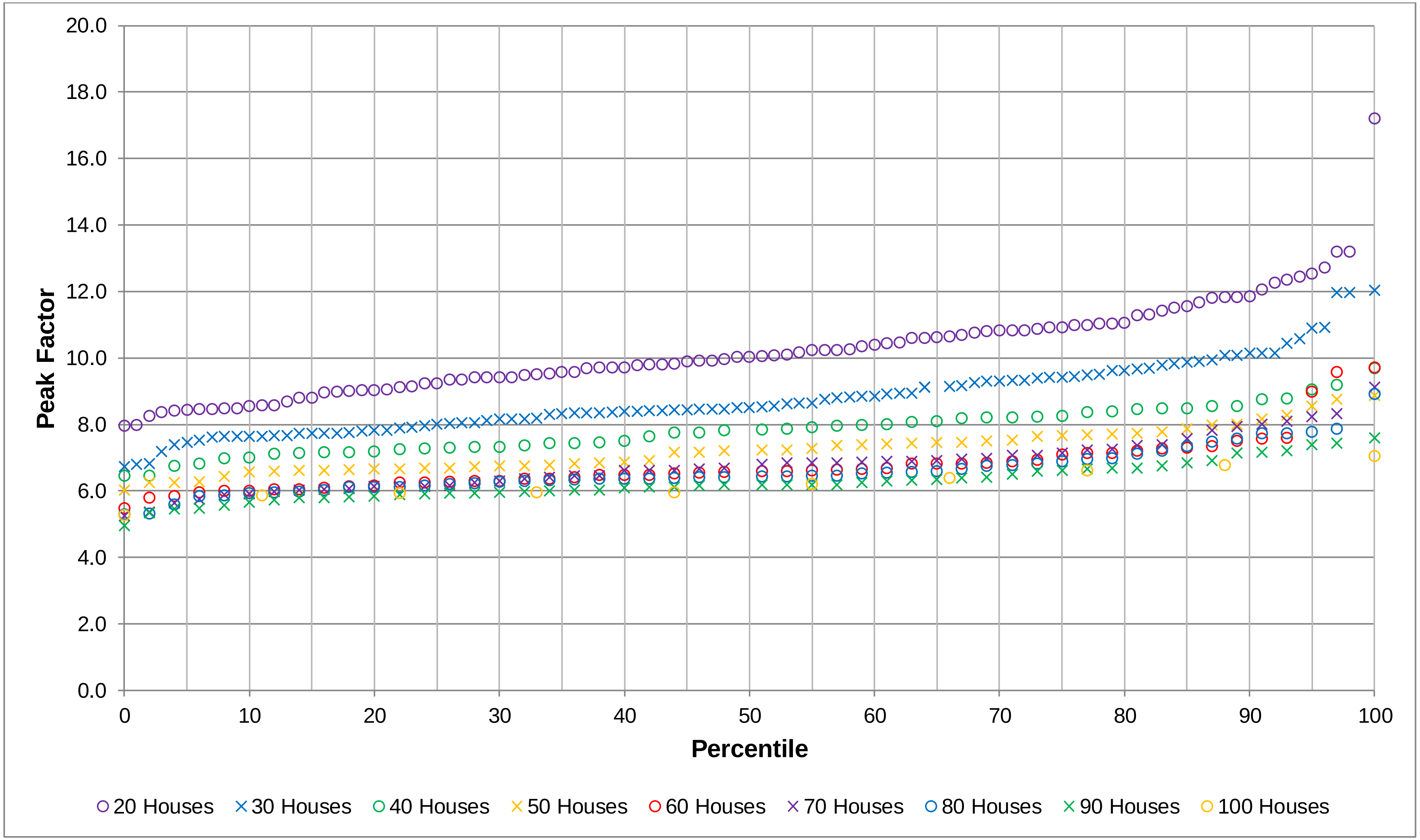


Figure C1.23: Result set for 1 second peak factor and household group sizes ranging from 20 to 100

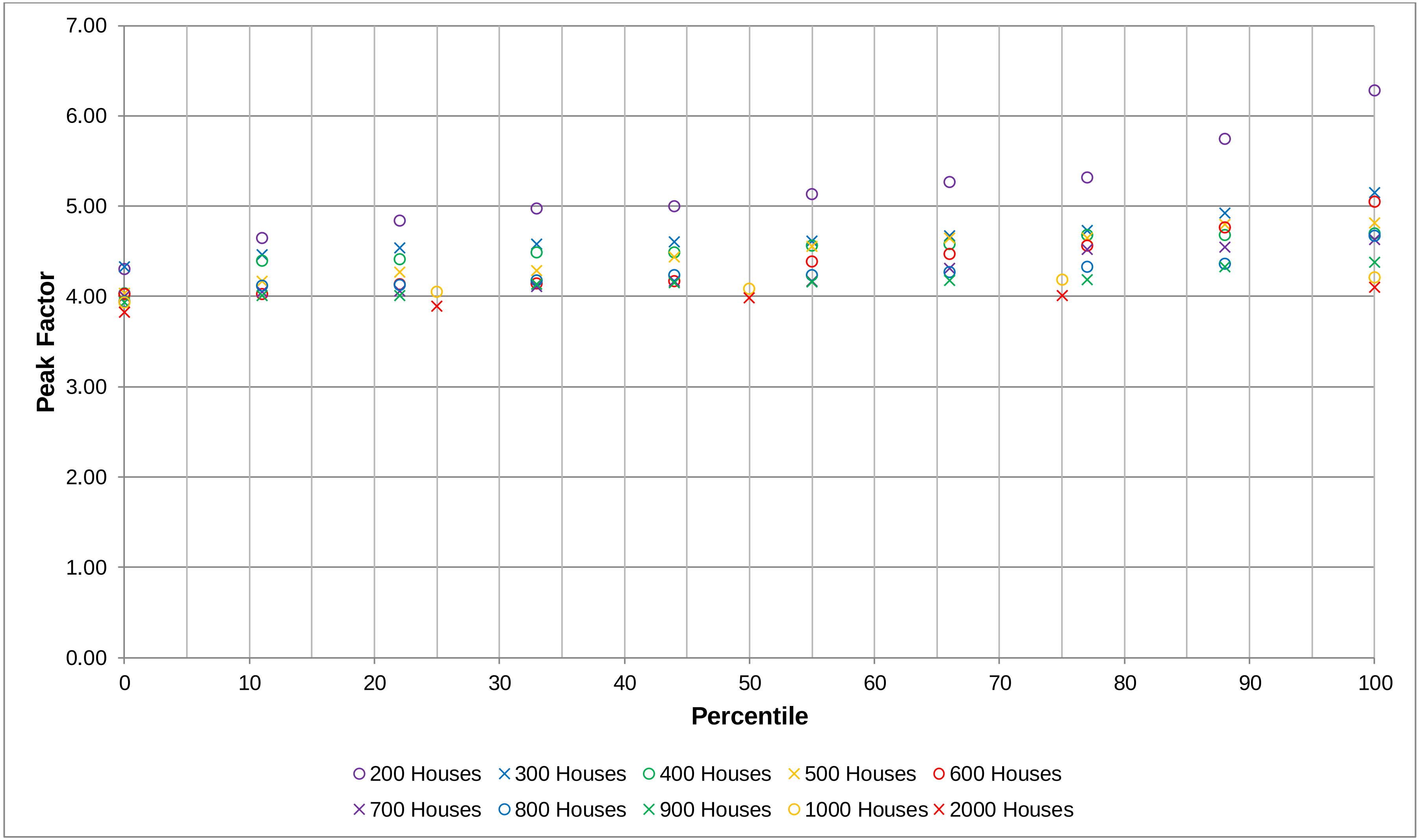


Figure C1.24: Result set for 1 second peak factor and household group sizes ranging from 200 to 2000

C2. Peak Factor Values for Selected Percentiles

The following legend is applicable to Tables C2.1 to C2.8:

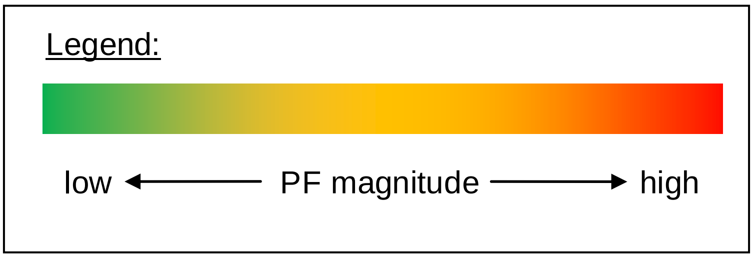


Table C2.7.1: 60 minute peak factor values for selected percentiles

60 min peak factor		Percentiles											
		0	10	20	30	40	50	60	70	80	90	95	100
Group size (Number of combined households)	1	2.23	4.39	5.01	5.53	5.94	6.37	6.94	7.65	8.52	9.85	10.81	17.87
	2	2.20	3.39	3.70	4.02	4.36	4.67	5.03	5.39	5.88	6.77	7.71	11.80
	3	2.23	3.01	3.33	3.52	3.74	3.98	4.26	4.55	4.94	5.66	6.32	8.97
	4	2.23	2.72	2.92	3.14	3.32	3.53	3.78	4.01	4.30	4.88	5.27	7.95
	5	2.16	2.64	2.89	3.06	3.18	3.34	3.50	3.65	3.92	4.32	4.70	6.04
	6	2.08	2.48	2.70	2.89	3.02	3.12	3.29	3.44	3.75	4.15	4.60	6.20
	7	1.91	2.36	2.54	2.64	2.74	2.87	2.99	3.13	3.37	3.81	4.48	6.15
	8	1.92	2.30	2.46	2.56	2.65	2.77	2.92	3.07	3.28	3.55	3.77	4.27
	9	1.85	2.22	2.44	2.54	2.64	2.82	2.96	3.12	3.32	3.68	4.04	5.32
	10	1.97	2.23	2.33	2.42	2.56	2.68	2.84	2.99	3.19	3.54	3.68	4.39
	20	1.72	1.95	2.05	2.12	2.21	2.29	2.37	2.45	2.55	2.77	2.93	3.30
	30	1.70	1.90	2.00	2.05	2.11	2.17	2.24	2.27	2.32	2.44	2.49	2.76
	40	1.59	1.82	1.89	1.91	1.96	2.03	2.06	2.11	2.18	2.30	2.32	2.65
	50	1.71	1.77	1.85	1.93	1.98	2.01	2.04	2.14	2.25	2.35	2.42	2.52
	60	1.62	1.75	1.79	1.83	1.86	1.89	1.94	1.97	2.03	2.15	2.21	2.35
	70	1.67	1.73	1.76	1.81	1.85	1.89	1.93	2.00	2.03	2.30	2.40	2.43
	80	1.63	1.76	1.80	1.84	1.87	1.90	1.95	2.02	2.03	2.15	2.22	2.46
	90	1.61	1.72	1.80	1.83	1.85	1.88	1.95	1.97	2.00	2.06	2.10	2.33
	100	1.58	1.78	1.81	1.81	1.83	1.87	1.91	1.91	1.93	1.99	2.05	2.10
	200	1.63	1.65	1.76	1.79	1.80	1.81	1.85	1.91	1.95	1.98	2.00	2.01
	300	1.66	1.67	1.70	1.72	1.72	1.74	1.77	1.81	1.87	1.88	1.88	1.89
	400	1.67	1.68	1.68	1.72	1.74	1.75	1.76	1.78	1.80	1.87	1.92	1.98
	500	1.63	1.71	1.76	1.77	1.78	1.79	1.81	1.85	1.85	1.87	1.88	1.90
	600	1.58	1.64	1.67	1.69	1.72	1.74	1.75	1.76	1.78	1.85	1.85	1.85
	700	1.70	1.72	1.75	1.76	1.77	1.78	1.78	1.79	1.80	1.81	1.82	1.83
	800	1.68	1.68	1.71	1.73	1.74	1.76	1.80	1.83	1.83	1.84	1.85	1.86
	900	1.60	1.64	1.68	1.71	1.72	1.72	1.73	1.75	1.75	1.76	1.77	1.78
	1000	1.68	1.69	1.69	1.71	1.74	1.76	1.77	1.77	1.78	1.80	1.81	1.82
	2000	1.66	1.66	1.67	1.67	1.69	1.70	1.71	1.72	1.73	1.74	1.74	1.75

Table C2.7.2: 30 minute peak factor values for selected percentiles

30 min peak factor		Percentiles											
		0	10	20	30	40	50	60	70	80	90	95	100
Group size (Number of combined households)	1	4.42	7.18	8.15	8.84	9.63	10.45	11.24	12.24	13.27	15.57	17.48	31.12
	2	3.23	5.40	5.89	6.31	6.78	7.23	7.81	8.45	9.35	10.56	12.21	19.22
	3	2.95	4.63	5.08	5.46	5.77	6.17	6.59	7.05	7.75	8.66	9.43	14.33
	4	3.27	4.36	4.66	4.94	5.24	5.54	5.87	6.18	6.66	7.62	8.19	10.75
	5	3.14	3.89	4.26	4.50	4.79	5.09	5.45	5.76	6.18	6.99	7.24	9.09
	6	2.79	3.69	4.02	4.22	4.48	4.75	5.04	5.28	5.60	6.11	7.02	10.38
	7	3.03	3.64	3.85	4.00	4.20	4.41	4.64	4.90	5.18	5.72	6.13	8.45
	8	2.94	3.47	3.79	3.98	4.16	4.42	4.60	4.87	5.04	5.48	5.81	6.66
	9	2.92	3.38	3.69	3.92	4.12	4.39	4.51	4.71	4.97	5.20	5.71	7.62
	10	2.82	3.52	3.69	3.83	3.94	4.19	4.43	4.61	4.92	5.30	5.74	6.09
	20	2.54	3.21	3.32	3.43	3.53	3.58	3.64	3.83	4.02	4.23	4.40	4.99
	30	2.78	2.97	3.08	3.17	3.26	3.34	3.48	3.60	3.68	3.92	4.02	4.42
	40	2.45	2.92	2.99	3.05	3.14	3.20	3.27	3.35	3.51	3.70	3.74	3.86
	50	2.61	2.91	2.95	3.00	3.08	3.13	3.23	3.32	3.45	3.65	3.90	4.11
	60	2.73	2.78	2.83	2.84	2.96	3.03	3.08	3.19	3.28	3.33	3.48	3.74
	70	2.62	2.78	2.86	2.89	2.95	3.01	3.07	3.16	3.26	3.57	3.73	3.82
	80	2.73	2.83	2.90	2.97	2.99	3.03	3.07	3.13	3.24	3.32	3.47	3.81
	90	2.65	2.79	2.82	2.88	2.93	3.04	3.09	3.13	3.23	3.33	3.48	3.66
	100	2.66	2.83	2.91	2.94	2.94	3.00	3.08	3.15	3.23	3.28	3.36	3.43
	200	2.64	2.64	2.66	2.78	2.87	2.94	2.99	3.01	3.02	3.05	3.10	3.15
	300	2.59	2.65	2.73	2.78	2.84	2.89	2.91	2.94	2.96	2.97	2.97	2.98
	400	2.61	2.65	2.67	2.67	2.70	2.74	2.80	2.86	2.87	2.89	2.90	2.92
	500	2.59	2.73	2.81	2.84	2.86	2.87	2.89	2.92	2.94	2.99	3.00	3.01
	600	2.60	2.61	2.65	2.69	2.75	2.80	2.83	2.84	2.85	2.88	2.91	2.94
	700	2.70	2.75	2.77	2.78	2.78	2.79	2.81	2.83	2.84	2.86	2.86	2.87
	800	2.62	2.74	2.77	2.80	2.81	2.83	2.85	2.88	2.91	2.92	2.95	2.98
	900	2.65	2.68	2.70	2.72	2.72	2.72	2.74	2.79	2.83	2.84	2.86	2.88
	1000	2.66	2.68	2.71	2.74	2.79	2.84	2.85	2.86	2.87	2.90	2.91	2.92
	2000	2.63	2.67	2.71	2.74	2.76	2.77	2.79	2.80	2.82	2.84	2.86	2.87

Table C2.7.3: 15 minute peak factor values for selected percentiles

15 min peak factor		Percentiles											
		0	10	20	30	40	50	60	70	80	90	95	100
Group size (Number of combined households)	1	6.18	10.58	12.05	13.42	14.81	16.17	17.57	19.44	21.82	25.11	28.92	62.24
	2	4.75	7.72	8.51	9.31	9.88	10.60	11.36	12.35	13.58	15.84	17.72	28.04
	3	4.61	6.50	7.13	7.64	8.14	8.66	9.31	9.94	10.91	12.38	13.63	19.78
	4	4.72	5.75	6.16	6.75	7.11	7.55	7.98	8.60	9.50	11.13	12.10	16.37
	5	4.04	5.23	5.70	6.00	6.37	6.75	7.15	7.55	8.26	9.55	10.28	13.87
	6	3.86	4.91	5.34	5.65	5.91	6.14	6.37	6.68	7.22	7.81	8.68	11.97
	7	3.80	4.47	4.86	5.08	5.31	5.57	5.87	6.39	6.79	7.56	8.18	9.85
	8	3.69	4.52	4.91	5.22	5.39	5.63	5.90	6.24	6.67	7.22	7.81	10.87
	9	3.77	4.42	4.79	5.03	5.27	5.50	5.74	6.02	6.46	7.13	7.76	10.97
	10	3.74	4.39	4.58	4.82	5.05	5.28	5.63	5.99	6.27	6.82	7.42	8.04
	20	2.93	3.76	3.94	4.09	4.21	4.30	4.48	4.62	4.77	5.23	5.70	6.04
	30	2.97	3.36	3.66	3.75	3.82	3.94	3.99	4.12	4.35	4.70	4.92	5.38
	40	2.84	3.37	3.44	3.55	3.70	3.74	3.79	3.87	4.01	4.41	4.72	5.07
	50	3.01	3.28	3.31	3.41	3.49	3.61	3.65	3.73	3.90	4.14	4.26	4.46
	60	2.87	3.11	3.17	3.24	3.31	3.37	3.46	3.57	3.70	3.78	3.88	4.18
	70	2.85	3.12	3.20	3.24	3.40	3.45	3.49	3.59	3.65	3.86	3.96	4.98
	80	2.95	3.13	3.18	3.27	3.35	3.40	3.44	3.49	3.64	4.06	4.19	4.55
	90	2.82	2.99	3.10	3.19	3.28	3.32	3.39	3.50	3.58	3.78	4.05	4.14
	100	2.72	3.08	3.18	3.34	3.42	3.47	3.53	3.56	3.60	3.66	3.67	3.69
	200	2.79	2.86	2.87	2.92	3.04	3.15	3.22	3.30	3.38	3.42	3.48	3.53
	300	2.74	2.90	2.93	2.96	3.02	3.06	3.06	3.07	3.10	3.16	3.21	3.25
	400	2.70	2.82	2.87	2.90	2.91	2.94	2.99	3.00	3.01	3.03	3.07	3.11
	500	2.78	2.92	2.94	2.95	2.96	3.06	3.15	3.16	3.17	3.19	3.21	3.23
	600	2.82	2.82	2.83	2.84	2.85	2.90	2.95	3.01	3.10	3.10	3.10	3.10
	700	2.80	2.86	2.96	2.98	2.98	3.02	3.05	3.06	3.07	3.14	3.15	3.16
	800	2.72	2.83	2.90	2.93	2.96	3.00	3.04	3.10	3.14	3.19	3.20	3.21
	900	2.75	2.82	2.87	2.87	2.89	2.92	2.94	2.97	3.02	3.11	3.15	3.18
	1000	2.78	2.80	2.83	2.86	2.89	2.93	2.95	2.97	2.99	2.99	2.99	2.99
	2000	2.72	2.75	2.78	2.83	2.89	2.96	2.98	3.00	3.01	3.02	3.02	3.03

Table C2.7.4: 10 minute peak factor values for selected percentiles

10 min peak factor		Percentiles											
		0	10	20	30	40	50	60	70	80	90	95	100
Group size (Number of combined households)	1	6.76	13.46	15.53	17.21	18.78	20.43	22.16	24.36	27.65	32.09	36.93	79.22
	2	5.98	9.39	10.45	11.35	12.28	13.20	14.36	15.58	17.24	20.00	22.21	40.17
	3	5.75	7.81	8.68	9.32	9.94	10.56	11.35	12.30	13.38	15.29	16.92	34.14
	4	5.42	6.94	7.49	8.04	8.60	9.04	9.65	10.35	11.14	12.75	14.00	18.90
	5	4.88	6.26	6.96	7.44	7.81	8.18	8.59	9.29	9.91	11.10	12.07	15.99
	6	4.43	5.66	6.02	6.58	7.01	7.38	7.70	8.39	9.02	9.94	11.08	14.37
	7	4.53	5.53	5.83	6.11	6.39	6.65	6.85	7.37	7.85	8.79	9.97	13.11
	8	4.11	5.14	5.66	6.07	6.39	6.58	6.87	7.22	7.67	8.62	9.98	13.70
	9	3.89	5.06	5.40	5.68	6.01	6.33	6.69	6.98	7.44	8.28	9.26	11.48
	10	4.41	4.95	5.37	5.61	5.94	6.19	6.49	6.74	7.28	7.99	8.69	9.95
	20	3.65	4.27	4.46	4.57	4.80	5.00	5.18	5.32	5.60	5.92	6.19	6.75
	30	3.35	3.88	4.05	4.18	4.39	4.52	4.66	4.82	4.99	5.36	5.57	6.62
	40	3.49	3.78	3.87	3.95	4.04	4.29	4.36	4.46	4.64	4.94	5.14	5.72
	50	3.44	3.64	3.74	3.83	3.88	4.00	4.16	4.25	4.37	4.61	4.98	5.17
	60	3.25	3.49	3.53	3.66	3.75	3.90	3.99	4.13	4.38	4.46	4.61	5.41
	70	3.32	3.49	3.62	3.71	3.77	3.83	3.93	4.04	4.17	4.45	4.82	5.10
	80	3.25	3.49	3.60	3.68	3.79	3.94	4.04	4.11	4.24	4.56	4.79	5.23
	90	3.21	3.39	3.50	3.60	3.67	3.82	3.86	3.92	4.01	4.16	4.50	4.69
	100	3.34	3.57	3.60	3.68	3.75	3.80	3.86	3.92	3.96	3.97	3.99	4.00
	200	3.16	3.17	3.27	3.37	3.41	3.42	3.44	3.51	3.61	3.67	3.69	3.72
	300	3.22	3.24	3.32	3.35	3.37	3.41	3.51	3.60	3.62	3.65	3.72	3.78
	400	3.14	3.20	3.21	3.28	3.33	3.35	3.37	3.42	3.55	3.64	3.70	3.75
	500	3.13	3.31	3.37	3.39	3.43	3.47	3.54	3.63	3.64	3.64	3.64	3.64
	600	3.13	3.20	3.21	3.24	3.29	3.38	3.46	3.51	3.59	3.77	3.77	3.77
	700	3.19	3.26	3.34	3.35	3.36	3.38	3.42	3.49	3.61	3.65	3.67	3.68
	800	3.27	3.28	3.31	3.33	3.34	3.36	3.38	3.44	3.50	3.52	3.55	3.58
	900	3.12	3.16	3.18	3.21	3.25	3.29	3.34	3.41	3.44	3.45	3.46	3.47
	1000	3.30	3.31	3.32	3.34	3.35	3.37	3.42	3.48	3.51	3.52	3.53	3.53
	2000	3.20	3.21	3.22	3.25	3.30	3.36	3.38	3.40	3.42	3.44	3.45	3.46

Table C2.7.5: 5 minute peak factor values for selected percentiles

5 min peak factor		Percentiles											
		0	10	20	30	40	50	60	70	80	90	95	100
Group size (Number of combined households)	1	8.83	18.53	21.63	24.48	26.67	28.68	31.17	34.74	39.79	48.08	56.04	109.48
	2	7.31	12.74	14.13	15.48	16.86	18.21	19.69	21.38	23.99	27.44	30.77	57.29
	3	6.94	10.69	11.86	12.69	13.57	14.55	15.42	16.42	17.76	20.15	22.52	40.98
	4	7.42	9.31	10.17	10.85	11.58	12.45	13.04	13.92	15.00	16.47	17.70	28.32
	5	6.63	7.95	8.79	9.32	10.00	10.77	11.62	12.22	13.10	14.88	16.16	25.32
	6	5.96	7.47	7.99	8.54	9.13	9.58	10.03	10.69	11.39	12.85	13.76	17.07
	7	5.99	6.99	7.52	8.09	8.56	8.94	9.44	10.00	10.91	11.58	12.25	16.01
	8	4.75	6.78	7.31	7.67	8.07	8.58	9.07	9.51	10.02	10.68	11.91	16.29
	9	5.19	6.61	7.08	7.47	7.75	8.03	8.33	8.81	9.62	10.51	11.05	14.80
	10	5.85	6.33	6.69	7.10	7.47	7.79	8.01	8.60	8.99	10.69	11.55	12.26
	20	4.68	5.24	5.45	5.69	5.94	6.26	6.55	6.83	7.02	7.21	7.66	9.52
	30	4.14	4.68	5.00	5.24	5.37	5.52	5.68	5.99	6.15	6.35	6.93	8.18
	40	4.28	4.47	4.63	4.82	4.93	5.12	5.21	5.33	5.48	5.98	6.59	6.89
	50	3.92	4.24	4.42	4.52	4.63	4.74	4.83	4.99	5.10	5.34	5.67	6.41
	60	3.63	3.98	4.09	4.27	4.38	4.45	4.52	4.75	4.87	4.99	5.08	5.64
	70	3.77	4.04	4.17	4.34	4.44	4.53	4.59	4.72	4.90	5.07	5.67	5.98
	80	3.65	4.08	4.15	4.22	4.36	4.45	4.50	4.66	4.89	5.17	5.53	6.56
	90	3.73	3.89	4.06	4.13	4.21	4.30	4.39	4.46	4.85	5.12	5.24	5.56
	100	4.03	4.05	4.09	4.27	4.36	4.38	4.39	4.53	4.84	4.89	4.96	5.03
	200	3.53	3.55	3.63	3.72	3.79	3.93	4.04	4.07	4.16	4.27	4.28	4.30
	300	3.41	3.62	3.64	3.66	3.68	3.74	3.81	3.86	3.91	3.94	4.04	4.15
	400	3.35	3.46	3.54	3.60	3.62	3.64	3.65	3.70	3.79	3.88	3.91	3.95
	500	3.52	3.62	3.64	3.69	3.72	3.74	3.77	3.82	3.94	4.09	4.10	4.12
	600	3.37	3.47	3.51	3.54	3.58	3.61	3.64	3.67	3.78	4.07	4.09	4.11
	700	3.47	3.51	3.52	3.54	3.61	3.70	3.75	3.76	3.80	3.92	3.95	3.98
	800	3.37	3.54	3.59	3.61	3.62	3.63	3.66	3.71	3.73	3.79	3.83	3.86
	900	3.32	3.50	3.53	3.56	3.58	3.59	3.62	3.69	3.77	3.80	3.86	3.92
	1000	3.44	3.48	3.52	3.56	3.58	3.60	3.60	3.61	3.61	3.63	3.63	3.64
	2000	3.32	3.38	3.44	3.47	3.47	3.47	3.52	3.56	3.61	3.67	3.71	3.74

Table C2.7.6: 1 minute peak factor values for selected percentiles

1 min peak factor		Percentiles											
		0	10	20	30	40	50	60	70	80	90	95	100
Group size (Number of combined households)	1	13.69	29.97	34.61	39.02	42.61	47.22	53.00	57.81	66.07	79.48	93.64	182.95
	2	11.39	19.86	22.39	24.31	26.37	28.97	31.12	33.67	37.08	42.42	49.81	81.04
	3	11.23	16.45	18.10	19.32	20.54	21.98	23.50	25.32	27.55	30.69	34.08	87.47
	4	11.58	14.55	15.67	16.55	17.90	18.65	19.83	21.25	22.76	25.13	27.04	39.36
	5	10.43	12.21	13.21	14.17	15.02	16.02	17.16	18.48	19.13	20.52	22.82	30.73
	6	8.74	11.14	12.22	12.81	13.54	14.46	15.41	16.47	17.56	20.01	21.46	24.78
	7	8.83	10.34	11.41	11.91	12.58	13.13	13.81	14.85	15.77	17.34	18.73	25.45
	8	7.73	10.06	10.88	11.49	12.20	12.82	13.46	14.16	15.17	16.07	18.24	22.93
	9	8.22	9.62	10.13	10.54	11.11	11.74	12.14	12.67	13.62	14.92	15.97	20.42
	10	7.93	9.17	9.64	10.03	10.37	11.01	11.44	12.32	13.35	14.10	15.07	16.24
	20	6.37	6.95	7.28	7.59	7.78	8.27	8.58	8.88	9.46	9.77	10.23	14.58
	30	5.46	6.11	6.56	6.79	6.94	7.09	7.22	7.47	7.87	8.40	8.90	10.11
	40	5.33	5.65	5.96	6.33	6.43	6.64	6.80	7.00	7.19	7.48	7.63	8.25
	50	5.02	5.40	5.54	5.65	5.79	6.01	6.16	6.31	6.42	6.58	6.75	7.86
	60	4.76	4.97	5.17	5.33	5.44	5.51	5.79	5.93	6.01	6.30	7.36	8.08
	70	4.66	5.01	5.22	5.34	5.52	5.73	5.84	6.01	6.23	6.77	6.92	8.37
	80	4.51	4.99	5.17	5.32	5.48	5.59	5.75	5.86	5.96	6.52	6.75	7.75
	90	4.42	4.77	4.94	5.10	5.17	5.34	5.42	5.61	5.75	5.94	6.40	6.84
	100	4.52	4.75	5.06	5.22	5.25	5.32	5.44	5.55	5.65	5.76	5.78	5.81
	200	3.83	4.15	4.29	4.36	4.43	4.51	4.67	4.85	4.93	5.05	5.28	5.51
	300	3.91	4.10	4.17	4.20	4.21	4.24	4.27	4.28	4.29	4.34	4.57	4.80
	400	3.71	3.92	4.04	4.11	4.13	4.14	4.17	4.22	4.26	4.33	4.34	4.34
	500	3.78	3.91	3.97	3.99	4.09	4.19	4.29	4.40	4.44	4.48	4.49	4.49
	600	3.72	3.74	3.80	3.81	3.86	3.93	4.02	4.13	4.20	4.38	4.41	4.45
	700	3.75	3.78	3.82	3.86	3.92	3.96	3.97	4.05	4.25	4.28	4.29	4.29
	800	3.74	3.81	3.87	3.89	3.91	3.92	3.93	3.95	3.99	4.06	4.18	4.29
	900	3.69	3.75	3.76	3.78	3.85	3.91	3.96	3.99	4.00	4.04	4.07	4.10
	1000	3.59	3.69	3.78	3.84	3.85	3.86	3.90	3.94	3.97	3.97	3.97	3.97
	2000	3.63	3.64	3.65	3.69	3.75	3.82	3.83	3.83	3.86	3.91	3.94	3.96

Table C2.7.7: 10 second peak factor values for selected percentiles

10 s peak factor	Percentiles											
Group size (Number of combined households)	0	10	20	30	40	50	60	70	80	90	95	100
1	15.58	35.56	41.62	46.92	52.19	57.30	63.99	70.62	80.53	94.58	112.18	254.25
2	14.58	24.14	27.24	29.63	31.88	34.24	37.19	40.30	44.15	51.77	58.29	112.59
3	13.61	19.91	21.95	23.49	24.84	26.49	28.39	30.22	32.83	37.13	40.90	88.41
4	14.01	17.37	18.86	20.04	21.52	22.61	24.00	25.52	27.27	30.14	32.24	45.75
5	12.23	15.17	16.66	17.54	18.72	19.37	20.24	21.40	22.64	24.93	26.88	34.52
6	10.79	13.61	14.71	16.01	16.72	17.37	18.12	19.24	21.22	23.13	25.08	51.02
7	10.99	13.20	13.79	14.70	15.26	16.09	16.56	17.62	18.80	20.61	22.25	29.82
8	10.32	12.03	13.15	13.90	14.79	15.32	16.27	17.42	18.24	19.61	21.71	28.95
9	9.82	11.39	12.01	12.47	13.09	13.94	14.65	15.38	16.54	17.93	18.97	22.60
10	9.96	10.92	11.44	12.08	12.59	13.31	13.83	14.37	15.27	15.89	16.49	19.68
20	7.89	8.35	8.61	8.99	9.29	9.57	9.96	10.20	10.91	11.53	11.86	16.52
30	6.26	7.24	7.48	7.78	7.91	8.15	8.43	8.77	9.10	9.47	9.71	11.94
40	6.09	6.77	6.99	7.19	7.27	7.64	7.80	7.90	8.09	8.29	8.55	9.11
50	5.95	6.25	6.33	6.54	6.63	6.77	7.01	7.24	7.41	7.52	7.76	8.40
60	5.29	5.64	5.95	6.07	6.18	6.28	6.40	6.63	6.84	7.28	8.16	9.10
70	5.06	5.57	5.78	6.01	6.15	6.31	6.50	6.71	7.01	7.65	7.88	8.78
80	4.85	5.55	5.79	5.93	6.06	6.18	6.26	6.47	6.77	7.23	7.42	8.19
90	4.77	5.35	5.55	5.65	5.75	5.90	5.98	6.21	6.43	6.73	7.05	7.24
100	5.02	5.57	5.64	5.71	5.76	5.80	6.01	6.30	6.43	6.65	6.65	6.65
200	4.16	4.37	4.64	4.73	4.85	4.94	4.99	5.07	5.24	5.63	5.86	6.09
300	4.20	4.35	4.37	4.39	4.41	4.46	4.51	4.54	4.59	4.80	4.92	5.05
400	3.94	4.19	4.31	4.34	4.39	4.44	4.45	4.47	4.51	4.54	4.58	4.62
500	3.91	4.07	4.12	4.16	4.25	4.38	4.48	4.49	4.53	4.68	4.71	4.75
600	3.89	3.94	4.02	4.04	4.07	4.20	4.33	4.36	4.41	4.62	4.80	4.97
700	3.92	3.92	3.97	3.99	4.04	4.07	4.15	4.29	4.39	4.46	4.49	4.52
800	3.87	3.97	4.04	4.08	4.10	4.14	4.18	4.19	4.22	4.31	4.39	4.46
900	3.87	3.90	3.93	3.96	4.03	4.09	4.10	4.11	4.15	4.23	4.24	4.24
1000	3.87	3.89	3.90	3.92	3.96	3.99	4.04	4.09	4.12	4.12	4.13	4.13
2000	3.78	3.81	3.84	3.86	3.89	3.92	3.92	3.93	3.96	4.01	4.03	4.05

Table C2.7.8: 1 second peak factor values for selected percentiles

1 s peak factor		Percentiles											
		0	10	20	30	40	50	60	70	80	90	95	100
Group size (Number of combined households)	1	15.58	36.95	42.56	48.08	53.49	59.16	65.92	72.54	83.25	98.59	115.36	265.74
	2	14.75	25.33	28.09	30.36	33.01	35.53	38.42	41.79	46.12	54.88	60.80	112.59
	3	14.40	20.63	22.80	24.52	25.93	27.28	29.05	31.25	33.85	38.68	41.65	88.41
	4	15.07	17.75	19.62	20.83	21.92	22.99	24.46	26.42	28.29	30.73	33.68	45.75
	5	12.56	16.03	17.15	18.07	19.21	19.89	21.15	22.59	23.55	26.04	27.09	38.84
	6	10.79	14.41	15.80	16.81	17.54	18.39	19.07	20.03	21.69	24.30	25.58	57.66
	7	11.45	13.57	14.71	15.44	16.19	16.66	17.40	18.52	19.43	20.79	23.23	31.16
	8	10.61	12.87	13.86	14.43	15.03	16.32	17.17	18.36	19.23	20.61	22.18	35.26
	9	10.26	11.77	12.81	13.30	13.91	14.60	15.03	16.10	17.23	18.59	19.94	25.74
	10	10.29	11.23	12.07	12.40	13.20	13.72	14.44	15.00	15.70	16.66	17.13	20.05
	20	7.95	8.55	9.03	9.41	9.71	10.03	10.37	10.81	11.05	11.84	12.44	17.20
	30	6.73	7.64	7.81	8.14	8.38	8.51	8.84	9.30	9.62	10.08	10.60	12.03
	40	6.46	7.00	7.18	7.32	7.48	7.83	7.99	8.21	8.41	8.56	8.92	9.69
	50	6.03	6.55	6.66	6.74	6.86	7.21	7.41	7.50	7.72	8.01	8.43	8.87
	60	5.49	6.00	6.16	6.29	6.48	6.58	6.67	6.85	7.15	7.52	8.35	9.71
	70	5.24	5.90	6.11	6.28	6.55	6.74	6.87	7.00	7.26	7.94	8.17	9.13
	80	5.28	5.92	6.11	6.25	6.36	6.43	6.53	6.76	7.00	7.59	7.76	8.92
	90	4.96	5.65	5.84	5.94	6.07	6.18	6.27	6.44	6.68	7.15	7.30	7.60
	100	5.26	5.81	5.91	5.94	5.95	6.08	6.27	6.44	6.63	6.79	6.91	7.04
	200	4.30	4.61	4.80	4.93	4.99	5.07	5.19	5.28	5.40	5.80	6.04	6.28
	300	4.33	4.45	4.52	4.57	4.60	4.61	4.63	4.69	4.77	4.94	5.04	5.14
	400	4.03	4.35	4.40	4.46	4.49	4.53	4.57	4.61	4.68	4.68	4.69	4.69
	500	4.03	4.15	4.24	4.28	4.38	4.50	4.59	4.64	4.68	4.80	4.81	4.82
	600	4.02	4.03	4.12	4.14	4.15	4.28	4.42	4.50	4.60	4.79	4.92	5.04
	700	4.00	4.05	4.06	4.09	4.14	4.17	4.23	4.37	4.53	4.55	4.59	4.63
	800	3.94	4.10	4.12	4.16	4.21	4.23	4.25	4.29	4.34	4.39	4.53	4.67
	900	3.93	4.00	4.01	4.09	4.14	4.15	4.16	4.18	4.21	4.33	4.36	4.38
	1000	3.93	3.98	4.02	4.06	4.07	4.08	4.12	4.16	4.19	4.20	4.21	4.21
	2000	3.83	3.85	3.87	3.91	3.95	3.99	3.99	4.00	4.03	4.07	4.08	4.10