

Statistical Analysis of Electricity Demand Profiles

by

Siphumlile Mangisa

Submitted in fulfilment of the requirements for the degree of

Magister Scientiae

in the

Faculty of Science

at the

Nelson Mandela Metropolitan University

January 2012

Supervisor: Prof IN Litvine

ACKNOWLEDGEMENTS

I would like to express my appreciation to the following:

My God, who gave me the strength and the courage to complete this project.

My family, especially my parents, for supporting and motivating me.

My supervisor Professor IN Litvine for all his guidance, advice, knowledge and encouragement.

My colleagues at the Centre of Expertise in Forecasting (COEF) and at the Department of Statistics for their advice and support.

Mr Thandekile Nyulu for kindly providing the calendar database.

Eskom for the financial support, collaboration and for kindly providing the data.

The National Research Foundation (NRF) for financial assistance towards this research. Opinions expressed and conclusions arrived at are those of the author and should not necessarily be attributed to the NRF.

The NMMU Research Office for the NMMU Postgraduate Bursary.

Contents

1	Introduction and Research Background	1
1.1	Introduction	1
1.2	Research background	3
1.3	Electricity profiles	5
1.4	Structure of the dissertation	6
2	Data and Methodology	7
2.1	Data	7
2.2	The model	9
2.3	When a day starts in South Africa, shifting the start of the day . . .	13
2.4	ANOVA	15
2.5	Practical significance	17
3	Results for the Parameters	18
3.1	General remarks on the parameters	18
3.2	Results for parameter a - The minimum demand in a day	20
3.3	Results for parameter m_1 - The time of the first peak	26
3.4	Results for parameter m_2 - The time of the second peak	29
3.5	Results for parameter s_1 - The width of the first peak	33
3.6	Results for parameter s_2 - The width of the second peak	36

3.7	Results for parameter b_1 - The total energy used during the first peak minus the minimum demand	40
3.8	Results for parameter b_2 - The total energy used during the second peak minus the minimum demand	44
4	Conclusions and Recommendations	48
A	Mathematica Script	54

List of Figures

1.1	Typical daily electricity demand profile	2
2.1	The daily electricity energy for the period 2000 to 2010	8
2.2	Model and parameter descriptions	10
2.3	A particular daily profile with a fitted model	12
3.1	Estimates for parameter a - The minimum demand in a day	20
3.2	A 100 day moving average for the daily energy	21
3.3	A 100 day moving average for parameter a , the daily minimum demand	21
3.4	Minimum demand for the 5 year period, 2000-2005	22
3.5	Minimum demand for the 5 year period, 2000-2005 deseasonalized . .	22
3.6	Estimates for parameter m_1 - The time of the first peak	26
3.7	A 100-day moving average of parameter m_1 , the time of the first peak	27
3.8	Estimates for parameter m_2 - The time of the second peak	29
3.9	A 100 day moving average of parameter m_2 , the time of the second peak	30
3.10	Estimates for parameter s_1 - The width of the first peak	33
3.11	A 100 day moving average for parameter s_1 , the width of the first peak	34
3.12	Estimates for parameter s_2 - The width of the second peak	36
3.13	Parameter s_2 , the time of the second peak deseasonalized	37
3.14	A 100-day moving average of parameter s_2	37
3.15	Parameter s_2 - Only the weekdays	38

3.16	Parameter s_2 - Only the weekend days	38
3.17	Estimates for parameter b_1 - The total energy used during the first peak minus the minimum demand	40
3.18	Parameter b_1 , the volume of energy for the first peak deseasonalised .	41
3.19	A 100-day moving average of parameter b_1	41
3.20	Parameter b_1 - Only the weekdays	42
3.21	Parameter b_1 - Only the weekend days	42
3.22	Estimates for parameter b_2 - The total energy used during the second peak minus the minimum demand	44
3.23	Parameter b_2 , the volume of energy used during the second peak de- seasonalised	45
3.24	A 100-day moving average of parameter b_2 , the volume of energy used during the second peak	45
3.25	Parameter b_2 - Only the weekend days	45
3.26	Parameter b_2 - Only the weekdays	45
4.1	A 100 day moving average for the daily demand	49
4.2	A 100 day moving average for parameter a	49
4.3	100-day moving averages of parameters b_1 (blue) and b_2 (green) . . .	49
4.4	A 100-day moving average of parameter m_1	50
4.5	A 100-day moving average of parameter m_2	50
4.6	100-day moving averages of parameters s_1 (blue) and s_2 (green) . . .	50
4.7	An illustrative summary of the model	51

List of Tables

2.1	Tests on the best time to start a day	13
2.2	Criteria for <i>Cohen's</i> f^2 and R^2 interpretation	17
3.1	ANOVA results for parameter a - The daily minimum demand	24
3.2	ANOVA results for parameter m_1 - The time of the first peak	28
3.3	ANOVA results for parameter m_2 - The time of the second peak . . .	31
3.4	ANOVA results for parameter s_1 - The width of the first peak	35
3.5	ANOVA results for parameter s_2 - The width of the second peak . . .	39
3.6	ANOVA results for parameter b_1 - The volume of energy used during the first peak	43
3.7	ANOVA results for parameter b_2 - The volume of energy used during the second peak	47

Abstract

An electricity demand profile is a graph showing the amount of electricity used by customers over a unit of time. It shows the variation in electricity demand versus time. In the demand profiles, the shape of the graph is of utmost importance. The variations in demand profiles are caused by many factors, such as economic and environmental factors. These variations may also be due to changes in the electricity use behaviours of electricity users.

This study seeks to model daily profiles of energy demand in South Africa with a model which is a composition of two de Moivre type models. The model has seven parameters, each with a natural interpretation (one parameter representing minimum demand in a day, two parameters representing the time of morning and afternoon peaks, two parameters representing the shape of each peak, and two parameters representing the total energy per peak). With the help of this model, we trace change in the demand profile over a number of years. The proposed model will be helpful for short to long term electricity demand forecasting.

Keywords:

Forecasting, electricity demand, profiles, modelling

Chapter 1

Introduction and Research Background

1.1 Introduction

An electricity demand profile is a graph showing the amount of electricity used by customers over a period of time. It shows the variation in electricity demand versus time. In the demand profiles, the shape of the graph is of utmost importance. Electricity demand profiles are mostly used in the electricity generation production, transmission and distribution for planning how much power will be needed at a given time. The variations in demand profiles are caused by many factors, such as economic and environmental factors. These variations may also be due to changes in residential, commercial or industrial electricity usage behaviours.

For the purpose of this dissertation, we will look at daily variations, that is, we will look at the daily electricity demand profiles. A profile consists of two time units, a “period” and smaller units, which may be called “grains”. For the daily electricity demand profile, the period is one (1) day and a grain is one (1) hour. The graphic in

Figure 1.1 below shows a typical daily profile (note the two peaks, one in the morning, and another one in the afternoon).

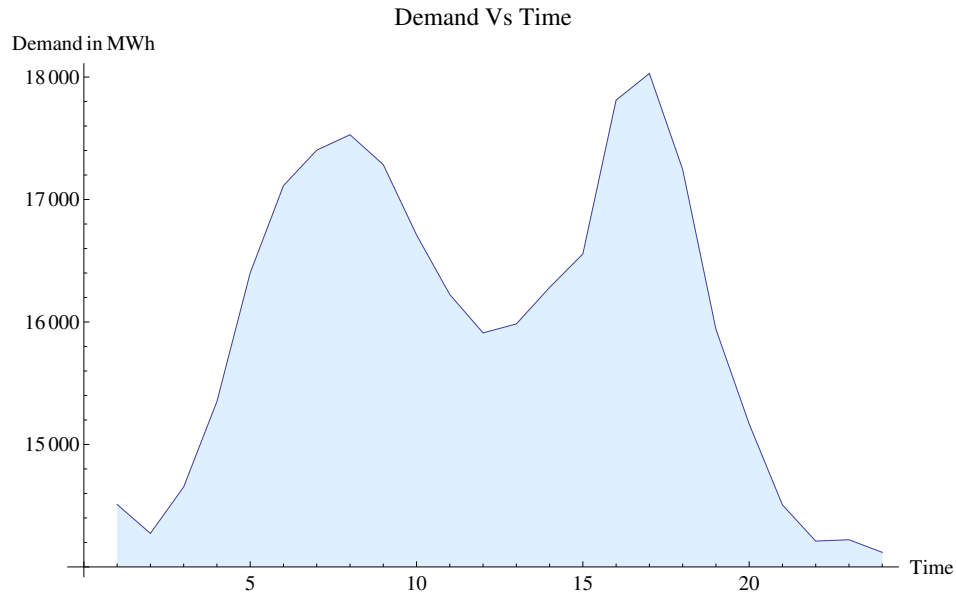


Figure 1.1: Typical daily electricity demand profile

Like all other economic phenomena, electricity demand has to be accurately forecasted in order to save costs, especially generation costs. An electricity utility may need to have their forecasting error less than a certain level per period for total energy to consider their forecasting models as being accurate. For instance, for the South African electricity utility Eskom, their forecasting error must be less than 2% per annum for total energy.

Also, electricity providers need to understand the behaviours of electricity peaks, so that their maximum electricity generating capacity is always more than maximum demand at any time. Also with regard to the time of the occurrence of the peaks, they need to understand their behaviours in order to better forecast when to switch on the additional power generators in anticipation of the increase in electricity demand.

1.2 Research background

The primary objective of the study is to use statistical methods to analyse the trends of change in variation of electricity demand profiles. This will be done by producing a model which best describes the trends in electricity demand profiles. The question now is, why it is important to build these kind of models?

The electricity sector is one of the most important sectors in any economy, especially a developing economy like South Africa. Economic growth is largely dependent on the availability of energy. The industrial (be it the mining, manufacturing, or agricultural industries), traction commercial and residential sectors all depend on electricity to function properly and efficiently.

The importance of accurate forecasting cannot be understated. Those responsible for balancing electricity generation and demand required must get it right if a power utility is to be efficient and effective.

Accurate forecasting can avoid wasting energy and prevent system failure. The former occurs when there is no need to generate electricity over a certain level and the latter occurs when normal operation is unable to withstand a heavy load (Soares and Souza, 2006). Over-forecasting of electricity demand will lead to energy wastage and monetary loss for the power utility. Bunn and Farmer (1985) in Soares and Souza (2006) demonstrated the importance of good forecasts, by noting that, in 1984, a 1% increase in forecasting error would cause an increase of £10 million in the operating costs per year in the United Kingdom.

Under-forecasting of electricity demand will lead to lost income or even power outages, which may affect industrial, commercial and residential areas. This may lead

to loss of revenue for commercial enterprises, which will translate into a loss for the national economy.

This dissertation will contribute to a better understanding of the trends in electricity demand and could help decision makers to better understand the different components within electricity demand. It could also contribute to the better understanding of the behaviours of electricity users, and towards the behaviours of the different trends in electricity demand. It could also contribute to the science of electricity forecasting in South Africa and beyond.

It must be stated at the outset that there is a general lack of literature on the analysis of electricity demand profiles. This dissertation therefore does not include an extensive analysis of literature.

1.3 Electricity profiles

According to mpoweruk.com, overall demand follows regular patterns with fairly large variations in demand during the course of the day. Superimposed on these daily patterns are smaller, long term seasonal variations with a greater demand for heating and lighting during the winter months or for airconditioning during summer months. Electricity demand from some major industrial users may also follow cyclic variations which could also affect the aggregate demand.

Soares and Souza (2006) and Taylor (2010) also mention the existence of these kinds of seasonality in what they express as intraday, intraweek and intrayear seasonal cycles.

These multiple trends or seasonalities in electricity profiles also exist in our data, and these will be discussed in detail in chapter 3.

1.4 Structure of the dissertation

In this chapter, the first chapter, we have the introduction and the research background. We define in detail what electricity demand profiles are, and also explain the importance of this project.

Chapter 2 describes the methodology used in our analysis.

Chapter 3 presents the results of the model. The estimates of the parameters and the trends in these parameters are discussed in detail.

Chapter 4 contains the conclusion and recommendations.

In Appendix A we attach the whole Mathematica script.

Chapter 2

Data and Methodology

2.1 Data

The data were kindly provided by the national power utility Eskom. They are a time series, with demand given hourly for all the days in the study. The data is for all the electricity that was used in that particular hours nationally. The data are for the 11 year period, from January 2000 to July 2010. They consist of 92 712 data points, which translates into 3863 days.

This being a fairly large data set, the computer program Mathematica was used for all the analysis.

In figure 2.1, we show the data for the daily demand over the time period under consideration. The graph exhibits a strong positive trend for the years 2000 to 2007. This corresponds with the fact that the South African economy was developing rapidly at the time. This is also the time when many households (especially rural households) which previously did not have access to electricity were being electrified and when the energy intensive construction industry was booming, from the thousands of sub-

sidized government housing (the so-called RDP houses) to large scale commercial and residential buildings.

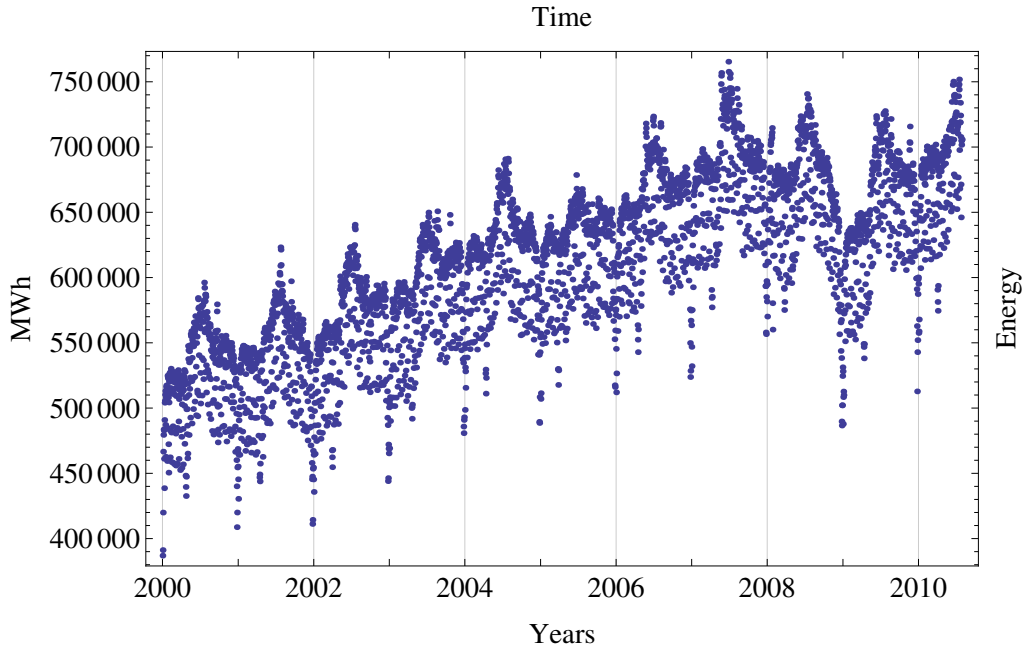


Figure 2.1: The daily electricity energy for the period 2000 to 2010

It can also be seen from the graph how the global economic crisis affected electricity demand in South Africa. One of the parameters in the model (representing the daily minimum demand) has a graph of similar shape to the daily energy graph. Therefore more discussions on the trends of this data will follow in the next chapter (section 3.2) when we discuss the trends of this parameter.

2.2 The model

A statistical model is a formalization of relationships between variables in the form of mathematical equations. It describes how one or more variables are related to one or more variables. Statistical models are useful and are used to estimate probabilistic future behaviour of a system based on past statistical information (wikipedia.org).

Meade (2000) made the following distinction between a model and a method. A model is an equation or set of equations representing the stochastic structure of the time series while a method is the combination of an estimation procedure and a model. He also says that statistically sophisticated and complex methods do not necessarily produce more accurate results than simpler ones.

The notion that a model must make sense and that a parameter must have a well defined meaning is deeply ingrained in applied statistical work (McCullagh, 2002). McCullagh (2002) further argues that inference is meaningful only for natural parameters.

We propose to model the daily profiles of electricity demand with a model which is a composition of two de Moivre type models, or rather a de Moivre dual model. The model has seven parameters, each one with a natural interpretation.

The proposed model is as follows:

$a + b_1 * \text{Normal}(m_1, s_1^2) + b_2 * \text{Normal}(m_2, s_2^2)$, where

- a is the minimum demand in a day (also known as the base load),
- m_1 is the time of the morning peak,

- m_2 is the time of the afternoon peak,
- s_1 is the width (shape) of the first peak,
- s_2 is the width (shape) of the second peak,
- b_1 is the total (scale) energy used during the first peak minus the minimum demand,
- b_2 is the total (scale) energy used during the second peak minus the minimum demand.

Also, the $f(x|m, s)$ is a de Moivre closed form model expressed as:

$$f(x|m, s) = \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(x-m)^2}{2s^2}}$$

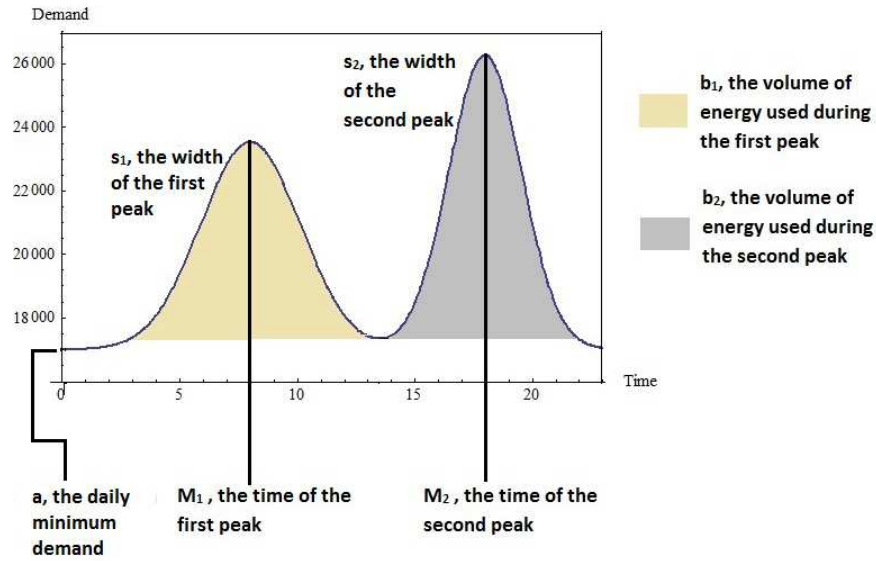


Figure 2.2: Model and parameter descriptions

The graph above shows a description of the model and the location of the different parameters.

Alternatively, we can define the volume of energy parameters b_1 and b_2 in the following manner:

Let b'_1 and b'_2 are the weights of the total energy used during the first peak and the second peak respectively. That is,

let t_0 and t_1 be the beginning and end of the first peak respectively, and t_1 and t_2 be the beginning and end of the second peak respectively, then

$$b'_1 = \frac{\int_{t_0}^{t_1} f(x|m_1, s_1) dx}{\int_{t_0}^{t_1} f(x|m_1, s_1) dx + \int_{t_1}^{t_2} f(x|m_2, s_2) dx} \quad \text{and}$$

$$b'_2 = \frac{\int_{t_1}^{t_2} f(x|m_2, s_2) dx}{\int_{t_0}^{t_1} f(x|m_1, s_1) dx + \int_{t_1}^{t_2} f(x|m_2, s_2) dx} ,$$

Where $dM(m_1, s_1) = f(x|m_1, s_1)dx$ and $dM(m_2, s_2) = f(x|m_2, s_2)dx$.

Thus $b_1 = (F(t_1) - F(t_0))b'_1$ and $b_2 = (F(t_2) - F(t_1))b'_2$

The parameters can be classified into two categories, volume and shape parameters. The volume parameters are a , b_1 and b_2 . And the shape parameters are m_1 , m_2 , s_1 and s_2 . This is one of the most exciting aspects of this model, as we can distinguish between volume and shape parameters.

Below, in Figure 2.3, is an example of a random day in our data with the model fitted. The red graph represents a typical daily profile for a particular day, given hourly. The green graph represents the fitted model. For all the more than 3800 days, the model is fitted on the daily data and the parameter estimates are recorded. This is all done in Mathematica successively using a “for loop”. In Chapter 3, the parameters will be presented graphical, and we will then trace and explain the changes in the demand profiles for the years 2000 to 2010.

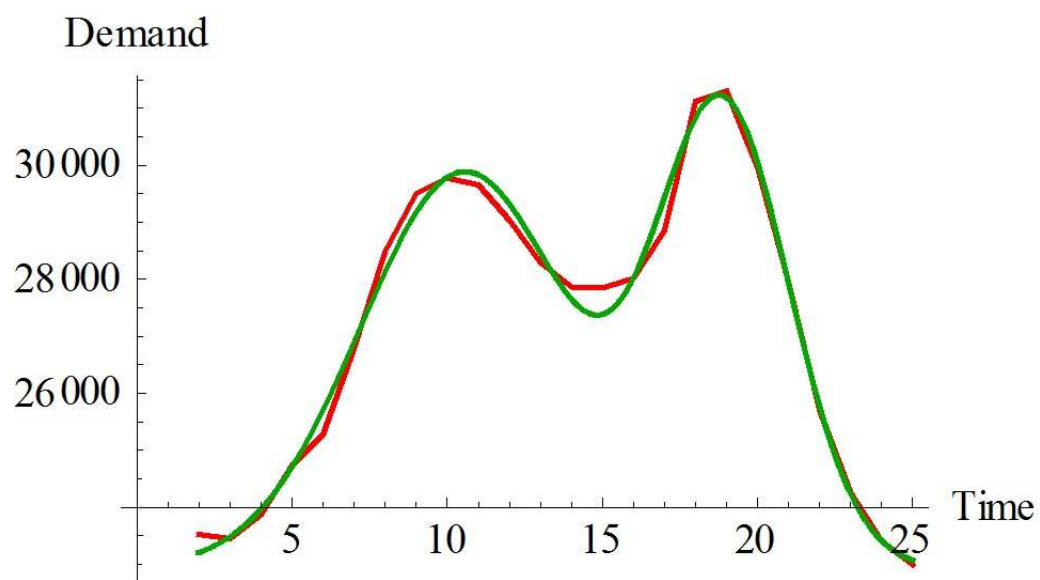


Figure 2.3: A particular daily profile with a fitted model

2.3 When a day starts in South Africa, shifting the start of the day

At the beginning of the study we found, through graphical observations, that it may be best to fit the model in a 24 hour period that is different from the traditional day. That is, instead of the day starting at 12 o'clock midnight and ending 24 hours later, our day would start at perhaps 2 AM and end at 2 AM the following day. We found that the residuals were smaller when the day started at 1 AM or 2 AM. Also through observing the graphics, it could be clearly seen that the model fitted on a day starting an hour or two after midnight had a better fit than when the day started at 12 midnight.

Table 2.1: Tests on the best time to start a day

Shift	Sum of Residuals	Maximum of Residuals
0	4.6199×10^6	3535.92
1	4.3054×10^6	2392.83
2	4.1545×10^6	2342.79
3	4.3130×10^6	2481.75
4	4.7500×10^6	5369.50
5	4.7391×10^6	5786.80

The sums of the absolute residuals are given in table 2.1 above. A residual is the difference between the data and the estimated model, with small residuals indicating a tight fit of the model to the data. The days starting at 2 AM and ending at 2 AM the following day had the least sum of absolute maximum residuals, and thus were used. The absolute daily maximum residuals were also calculated and the days starting at 2 AM and ending at 2 AM the following day were the best for this statistic too.

This system of shifting the start is used in some industries as well. For instance, in sales or demand forecasting, the start of the month may be shifted to the second week (ending during the second week of the next month). This then produces more accurate forecasts.

2.4 ANOVA

Analysis of variance (ANOVA) is a collection of statistical models, where the observed variance of a particular variable is partitioned into components attributable to different sources of variation (wikipedia.org). In its simplest form, ANOVA tests whether the means of several groups are all equal or not. It therefore generalizes the t-test to more than one group. Doing multiple two-sample t-tests would result in an increased chance of committing a type I error. Hence, ANOVA tests are useful in comparing two or more means.

ANOVA makes the following assumptions:

- Even if a statistical model is nonlinear, it can be approximated by a linear model
- The errors are independent
- The errors follow a Normal distribution
- The variances of all errors are equal to each other
- The expected values of the errors are zero

ANOVA is mainly used in the social and health sciences and particularly in psychology and medicine.

We used ANOVA in this study to check the effects that certain factors have on the parameters. For this study the factors were:

- which month of the year it is,
- which day of the week it is,

- is it a public holiday
- and is it a school holiday?

We also test if there are any interactions between the terms. At the end we also give regression models for each of the parameters.

If a result is significant (for this project all results were significant), a Bonferroni post-hoc test was used to check exactly which months of the year have significant different means and which days of the week have significantly different means. The Bonferroni post-hoc test is probably the most commonly used post-hoc test, because it is highly flexible and very simple to compute.

2.5 Practical significance

Since we were dealing with a large sample size in our ANOVA tests, it was necessary to use practical significance. Effect size or practical significance is a statistic used to determine the magnitude of a research result. Effect size quantifies the size of the difference between two groups, and may therefore be said to be a true measure of the significance of the difference (Coe, 2002). With this in mind, we resolved that our statistical tests will be significant if and only if they are both statistically and practically significant.

For ANOVA and multiple regression, the coefficient of determination R^2 and *Cohen's* f^2 are recommended. *Cohen's* f^2 represents the square root of variance explained over variance not explained. These statistics are defined as:

$$R^2 = 1 - \frac{SS_{\text{err}}}{SS_{\text{tot}}} \quad \text{and} \quad f^2 = \frac{R^2}{1-R^2}$$

For *Cohen's* f^2 , an effect size of 0.02 to 0.1 can be regarded as a “small” effect, around 0.15 a “medium” effect and around 0.35 to infinity a “large” effect. *Cohen's* f^2 can be larger than one. The table below shows how to interpret these statistics.

Table 2.2: Criteria for *Cohen's* f^2 and R^2 interpretation

	Small	Medium	Large
<i>Cohen's</i> f^2	$0.02 < f^2 < 0.15$	$0.15 < f^2 < 0.35$	$f^2 > 0.35$
R^2	$0.01 < R^2 < 0.09$	$0.09 < R^2 < 0.25$	$R^2 > 0.25$

Chapter 3

Results for the Parameters

3.1 General remarks on the parameters

All the seven parameters of the model were estimated for all the 3863 days in the dataset. The results for these parameters are given in graphical form in the next sections with discussion.

All the parameters contained seasonal trends as will be seen in the 100 day moving averages included in the next sections. The ANOVA tests conducted on the parameters confirmed that the means of the different months and the different days of the week are not equal. Thus not only were there yearly seasonal variations, there were also weekly and monthly variations. The moving averages clearly showed the different trends in the data. A 100 day moving average was selected after trial and error, because it best showed the trends (especially the yearly trends) and at the same time had some degree of smoothness.

A few observations which we feel are important are listed below:

- Only parameter a , the minimum daily demand, catches the economic conditions

and trends of the day. The parameter was the only one to really catch the global economic crisis of the late 2000's. The other parameters kept their trends constant or became flat, rather than decrease.

- All the other parameters had quite consistent behaviour. Except for the seasonal variations, the trends of the other parameters did not change that much. One would have expected parameters b_1 and b_2 (the volume of energy used during the first and second peak respectively) to follow the trends of the daily demand, but only parameter a follows these trends.
- Some of the parameters had what we call two parallel datasets. One dataset contains the weekend days of Saturday and Sunday. The other dataset contains the weekdays Monday, Tuesday, Wednesday, Thursday and Friday. This does confirm a view held by most people that electricity is used differently in the weekend days than during the rest of the week, possible because the weekend days are public holidays.
- Parameter s_2 , the width of the second peak, had what we call a v-shape, and this special effect was only found on this parameter.

3.2 Results for parameter a - The minimum demand in a day

Figure 3.2 below represents the minimum demand (parameter a) over the period under consideration.

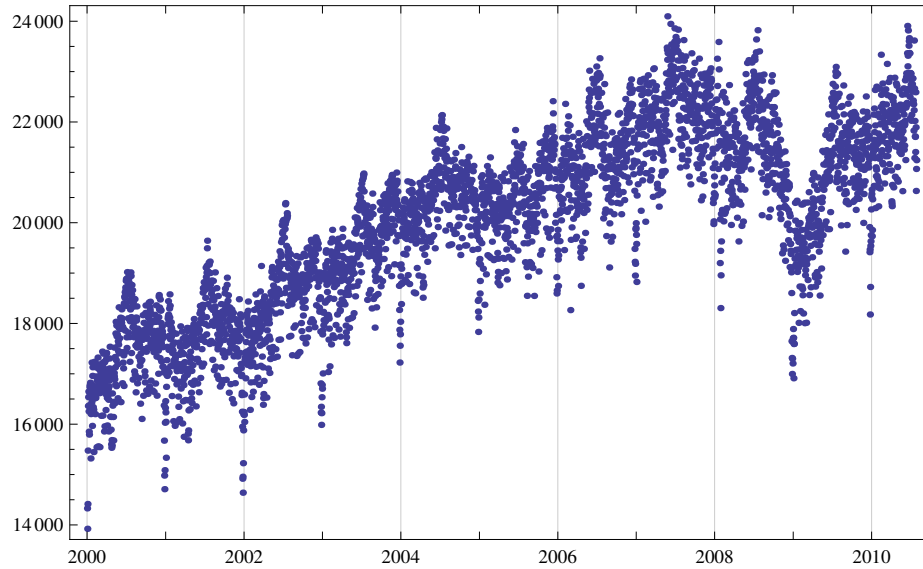


Figure 3.1: Estimates for parameter a - The minimum demand in a day

The daily electricity minimum demand has a graph of similar shape to the daily energy. As will be seen in the following sections of this chapter, and as mentioned in the previous section, of the seven parameters in the model, this parameter is the only one that catches the economic realities and trends of electricity use in South Africa. Figures 3.3 and 3.4 show 100 day moving averages of the daily demand and the minimum demand respectively; they are of similar shape, with a few slight differences. There is an upward moving trend between the years 2000 to 2007. A minor downward trend occurred during the end of 2005, but the overall upward

trend for this period did not change. This corresponds with the fact that the South African economy was developing rapidly at the time. And this is also the time when many rural households which previously did not have access to electricity were being electrified. Figure 3.6 in the next page shows the minimum demand for the first five years (2000-2005) deseasonalized, as opposed to figure 3.5 with seasonality. Figure 3.6 clearly shows the upward trend for those years.

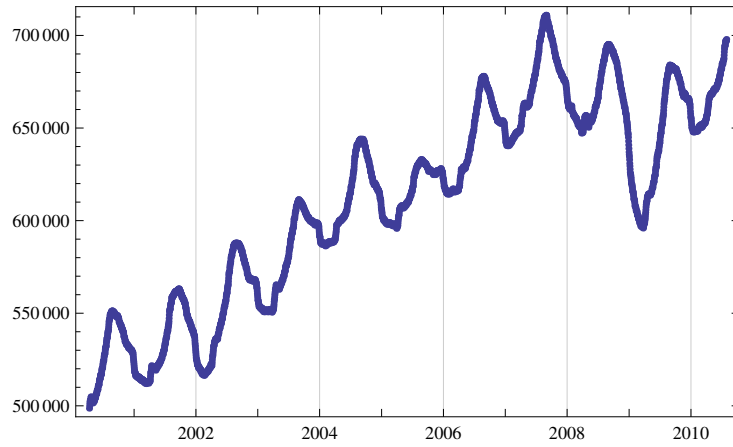


Figure 3.2: A 100 day moving average for the daily energy

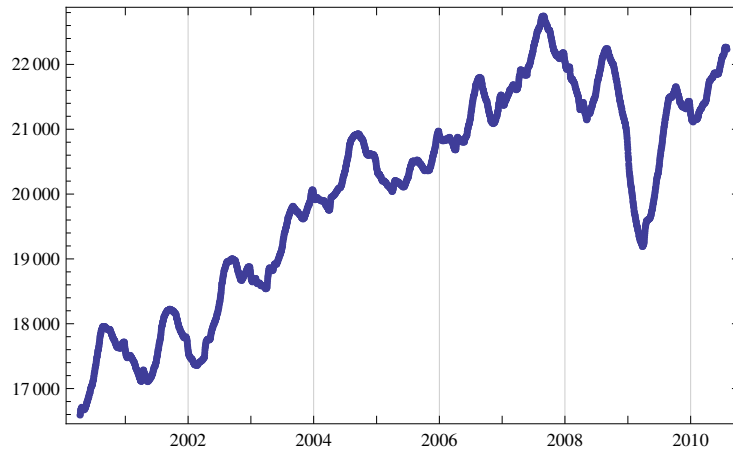


Figure 3.3: A 100 day moving average for parameter a , the daily minimum demand

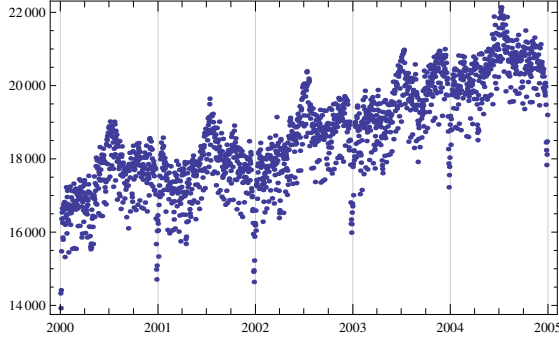


Figure 3.4: Minimum demand for the 5 year period, 2000-2005

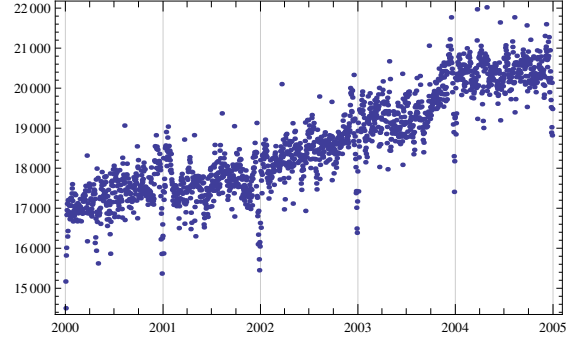


Figure 3.5: Minimum demand for the 5 year period, 2000-2005 deseasonalized

The seasonal component of the trends is the most obvious and the most glaring. The 100 day moving average, in Figure 3.4, does catch the trends and the elements of seasonality very well. Other moving averages were tried and the 100 day moving average was chosen as the most appropriate.

The daily minimum demand is at its highest during the winter months of May, June and July. These are the months when South Africa is cold, especially during the evenings and the mornings, hence the extra need for heating which implies extra need for electricity. The minimum demand is quite low during the summer months of November, December and January. There are, during the month of December of every year, what may be considered as outliers, a few days when the minimum electricity demand is very low. The only plausible explanation is that these are the days when there is a total shutdown by industry (i.e. mining, manufacturing, construction industries). As can be seen on Figure 2.1, the same observations can be made from the daily demand.

The daily minimum demand was also affected by the global financial crisis of the late 2000's. South Africa, being an integral part of the global economy, did not es-

cape unscathed from that economic crisis. There was a sharp fall in the demand for South African export products, and a significant decrease in the prices of key export commodities (info.gov). This meant that many of the big users of electricity (i.e. large industries), had to at the very least produce less than usual, and in some cases shut down all production. This in effect means less electricity was used in that period. Also, we can assume from the isolated dip from November 2009 to February 2010 that there was a total shutdown in big industry during those four months.

An ANOVA test is used to test whether there is any significant difference in the minimum electricity demand between the different months of the year, the different days of the week, during holiday and non-holiday days of the year or during school holidays and non-school holiday days. A hypothesis for checking if there is significant difference between the different months would be,

$$H_0 : u_{January} = u_{February} = \dots = u_{December}$$

with the alternate hypothesis being that at least one pair of the u 's (i.e. the monthly means) are not equal.

The ANOVA test was done on Mathematica, and the results are presented in Table 3.1. The second row gives the results for the test on whether the monthly means are equal. With the p-value being a value very close to 0 (zero), the null hypothesis is rejected. This means that there is significant difference between the monthly means. This further confirms the seasonality of the minimum demand during this period. A Bonferroni post-hoc test also shows that there are major variations between the means of the different months, in particular between the month of June and the other months of the year.

Table 3.1: ANOVA results for parameter a - The daily minimum demand

	DF	SumOfSq	MeanSq	FRatio	PValue
Month	11	7.74482×10^8	7.04075×10^7	24.1045	1.21823×10^{-48}
WeekDay	6	5.65403×10^8	9.42339×10^7	32.2616	4.49847×10^{-38}
Holiday	1	3.78728×10^7	3.78728×10^7	12.966	0.00032
SchoolHol	1	1.15439×10^7	1.15439×10^7	3.95213	0.04688
Month WeekDay	66	4.74096×10^7	718327.	0.245924	1.
Holiday Month	7	7.72192×10^7	1.10313×10^7	3.77665	0.000434015
Month SchoolHol	9	1.57744×10^8	1.75271×10^7	6.00054	2.18167×10^{-8}
Holiday WeekDay	6	1.76551×10^7	2.94251×10^6	1.00739	0.418425
SchoolHol WeekDay	6	3.28157×10^7	5.46928×10^6	1.87245	0.0816997
Holiday SchoolHol	1	4.52045×10^7	4.52045×10^7	15.4761	0.000085
Error	3747	1.09447×10^{10}	2.92093×10^6		
Total	3861	1.27121×10^{10}			

Similarly, on the hypothesis on whether the means of the different days of the week significantly differ, we reject the null hypothesis. There is also an element of seasonality between the different days of the week. From the test there is also a significant difference between holiday and non-holiday days, and also between the school holidays and school days. The p-value of 0.04688 for the school holidays test indicates that the means are quite close.

An ANOVA test was also done to test the effect that these factors have on each other, that is if there is any interaction between any two factors. The hypothesis for checking whether the type of month it is has an effect on the minimum demand of a certain day of the week (or rather if there is an interaction between the variables

month and weekday) is,

H_0 : There is no interaction between the factors Month and Weekday

and the alternate hypothesis is that there is an interaction between the factors.

The results for this test are also shown in Table 3.1. With a p-value of 1, the null hypothesis is not rejected. The type of month has no effect on the minimum demand of any particular day of the week. Similarly, for the hypothesis about the interaction between the variables Holiday and WeekDay, and SchoolHol and WeekDay, the null hypothesis is not rejected.

For the hypothesis of whether month has an effect on variable school holiday, the p-value is 0.00043 and the null hypothesis is rejected. This is understandable as school holidays always occur during certain months. The null hypothesis is also rejected for whether there is an interaction between the variables holidays and school holidays.

In section 2.5, we said that these statistical tests are significant if they are both statistically and practically significant. This test produced a co-efficient of determination(r^2) of 0.139, and Cohen's f^2 of 0.161. This effect size, as discussed in section 2.5, means that these variables have a moderate effect. Hence the results are significant.

3.3 Results for parameter m_1 - The time of the first peak

Figure 3.7 below represents the time of the first peak (parameter m_1) over the period under consideration.

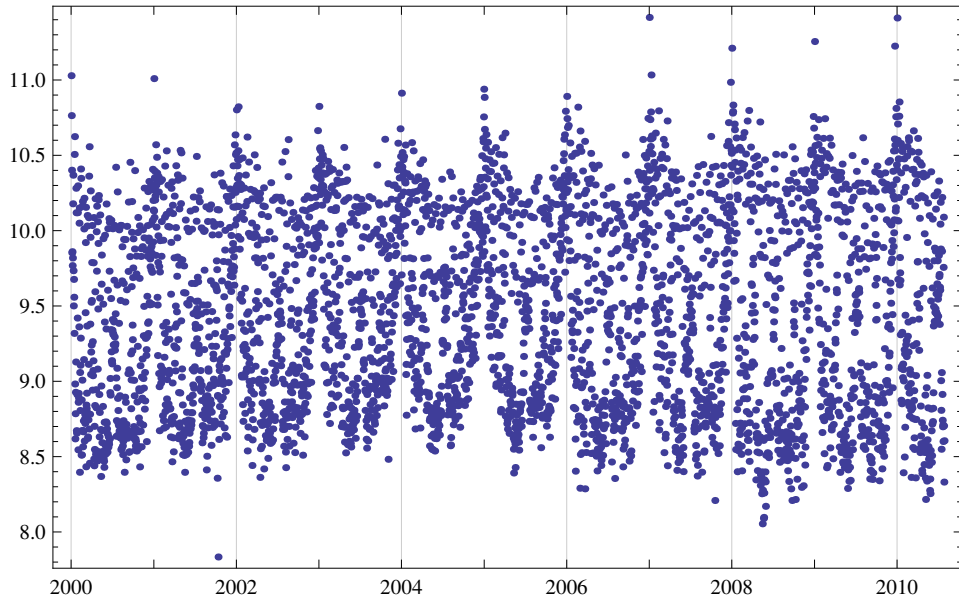


Figure 3.6: Estimates for parameter m_1 - The time of the first peak

The trends for the time of the first peak remain flat, oscillating between 08:30 AM and 10:30 AM, with the mean at 09:44 AM. The time of the first peak seems relatively stationary for the period under consideration. The global economic crisis seems not to have affected this parameter. The behaviours of electricity users, at least when it comes to the time of the first peak, remained the same.

The data have a subtle yearly seasonal component. It seems that, during the months of December and January, the school holiday months, people generally wake up later

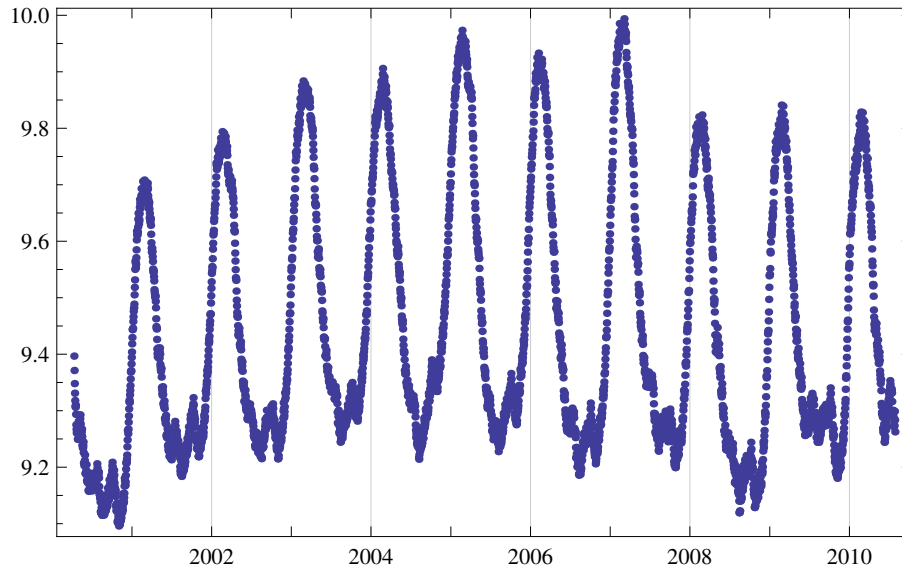


Figure 3.7: A 100-day moving average of parameter m_1 , the time of the first peak

than they usually do. The 100 day moving average in Figure 3.8 catches the seasonal variations very well. The moving average confirms that there may have been a few subtle changes in the trends, but the overall trends remained the same.

An ANOVA test was done for this parameter too, with the results in Table 3.2. All the p-values were small values very close to zero. Thus the null hypotheses were rejected in all four tests. We, for instance, found that there is sufficient evidence to conclude that the monthly means are not equal. Similarly, we reached the same conclusions in relation to the means of the days of the week, the holiday and non-holiday days and between school holiday and school days. The Bonferroni post-hoc test shows that there is a difference between most of the months. It also shows there is a difference between the weekdays Friday, Saturday and Sunday and the the other days of the week.

The tests on whether there is an interaction between the variables also produced

Table 3.2: ANOVA results for parameter m_1 - The time of the first peak

	DF	SumOfSq	MeanSq	FRatio	PValue
Month	11	319.28	29.0255	431.015	3.44308×10^{-654}
WeekDay	6	752.449	125.408	1862.25	$5.07893 \times 10^{-1119}$
Holiday	1	88.1889	88.1889	1309.56	3.39739×10^{-246}
SchoolHol	1	67.6728	67.6728	1004.91	1.36737×10^{-195}
Month WeekDay	66	67.4442	1.02188	15.1745	1.23417×10^{-145}
Holiday Month	7	19.184	2.74057	40.6962	1.48200×10^{-55}
Month SchoolHol	9	9.40142	1.0446	15.5119	3.90848×10^{-25}
Holiday WeekDay	6	39.9506	6.65843	98.8746	8.55635×10^{-116}
SchoolHol WeekDay	6	23.9908	3.99847	59.3755	1.68920×10^{-70}
Holiday SchoolHol	1	3.48105	3.48105	51.692	7.80018×10^{-13}
Error	3747	252.331	0.0673421		
Total	3861	1643.37			

small p-values. Therefore, the null hypotheses were rejected for all four cases. There are interactions between the four factors.

In section 2.5, we said that these statistical tests are significant if they are both statistically and practically significant. This test produced a co-efficient of determination(r^2) of 0.846, and Cohen's f^2 of 5.513. This effect size, as discussed in section 2.5, means that these variables have a large effect. Hence the results are significant.

3.4 Results for parameter m_2 - The time of the second peak

Figure 3.9 represents the time of the second peak (parameter m_2) over the period under consideration.

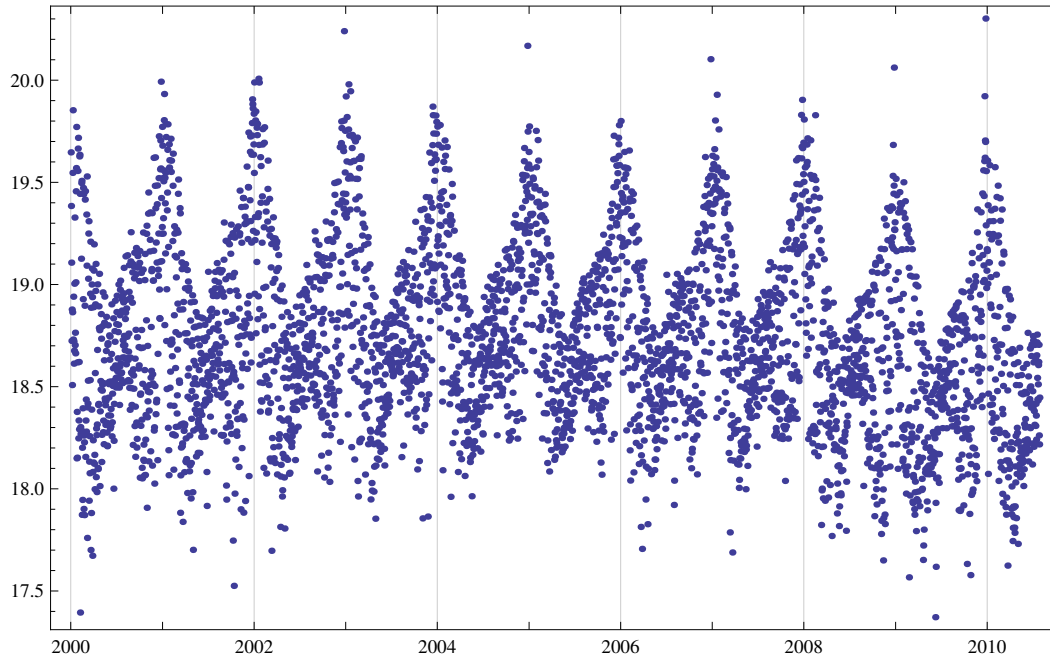


Figure 3.8: Estimates for parameter m_2 - The time of the second peak

The general trend of the graph is flat, with a slight decreasing trend after the year 2007. We suspect that the global economic crisis played a role, however minor that role may have been, in this decreasing pattern. The 100 day moving average in Figure 3.10 clearly shows these patterns.

The seasonal component of the trends is quite apparent. It is more pronounced for

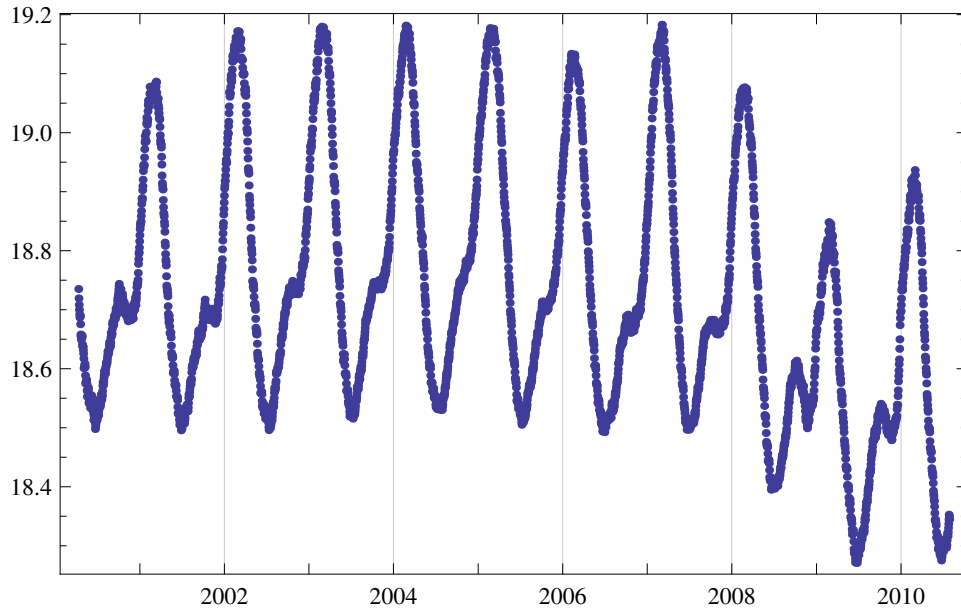


Figure 3.9: A 100 day moving average of parameter m_2 , the time of the second peak the time of the second peak than for the time of the first peak. During the summer months of December, January and February, the time of the second peak occurs far later in the evening than the average time of 18:45, while during the winter months around June the second peak occurs earlier than that. The time of the second peak oscillates between 17:30 and 20:00.

An ANOVA test was done on the factors we were testing, with the results shown in Table 3.3 below. The tests produced p-values which are close to zero. Thus, the null hypotheses were rejected in all four cases, and there is enough evidence to suggest that the monthly means are not the same and the means of the different days of the week are also not the same. Also there is a significant difference between the means of holiday and non-holiday days and also between school holidays and school days. A Bonferroni post-hoc test confirmed the difference between most of the months and between most of the days of the week.

Table 3.3: ANOVA results for parameter m_2 - The time of the second peak

	DF	SumOfSq	MeanSq	FRatio	PValue
Month	11	232.663	21.1512	427.373	2.39748×10^{-650}
WeekDay	6	261.805	43.6341	881.653	2.96906×10^{-711}
Holiday	1	17.3245	17.3245	350.051	9.53613×10^{-75}
SchoolHol	1	3.4441	3.4441	69.59	1.01378×10^{-16}
Month WeekDay	66	33.8377	0.512692	10.3592	1.33756×10^{-93}
Holiday Month	7	3.10278	0.443255	8.95622	5.44090×10^{-11}
Month SchoolHol	9	4.839	0.537667	10.8639	7.61190×10^{-17}
Holiday WeekDay	6	8.62665	1.43777	29.0511	3.58527×10^{-34}
SchoolHol WeekDay	6	1.67212	0.278687	5.63103	7.84991×10^{-6}
Holiday SchoolHol	1	0.000766916	0.000766916	0.015496	0.90094
Error	3747	185.444	0.0494913		
Total	3861	752.759			

The ANOVA test of whether there were interactions between the factors produced p-values very close to zero for all but one of the tests. Thus, there was an interaction between almost all the variables. The test on whether there was an interaction between variables holiday and school holidays produced a p-value of 0.901 and thus the null hypothesis was accepted. Hence there was no interaction between the variables holiday and school holiday.

In section 2.5, we said that these statistical tests are significant only if they are both statistically and practically significant. This test produced a co-efficient of

determination(r^2) of 0.754, and Cohen's f^2 of 3.059. This effect size, as discussed in section 2.5, means that these variables have a large effect. Hence the results are considered significant.

3.5 Results for parameter s_1 - The width of the first peak

Figure 3.11 represents the width of the first peak (parameter s_1) over the period under consideration.

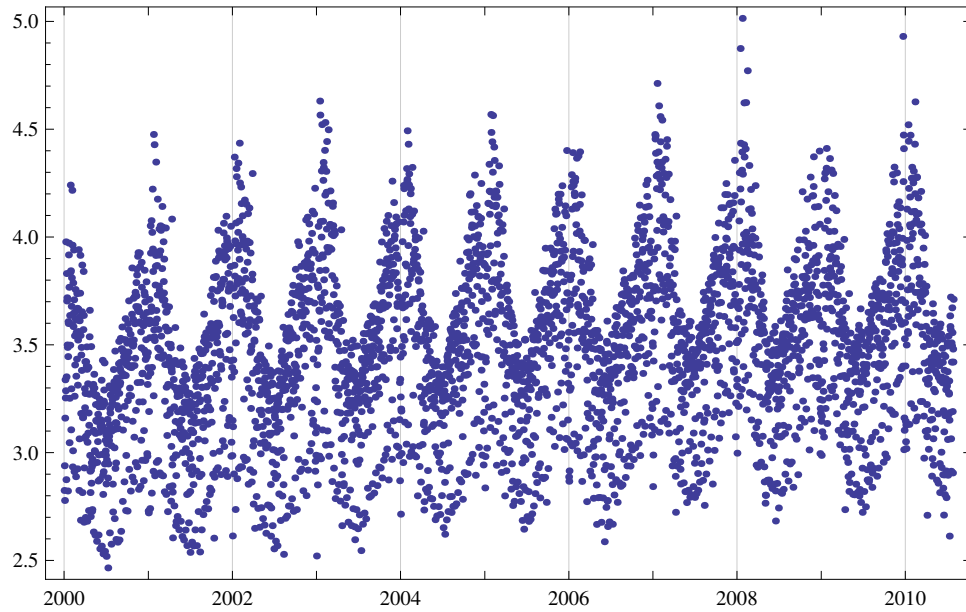


Figure 3.10: Estimates for parameter s_1 - The width of the first peak

The width of a peaks is a measure which gives us an indication of how spread out the data is during that peak. The time series of the width of the first peak shows a generally flat trend. As can be seen from the 100 day moving average in Figure 3.12, the trend looks flat, with a minor increasing trend during the period under study.

This parameter also shows a very noticable seasonal component. The width of the first peak widens during the summer months of November, December and January.

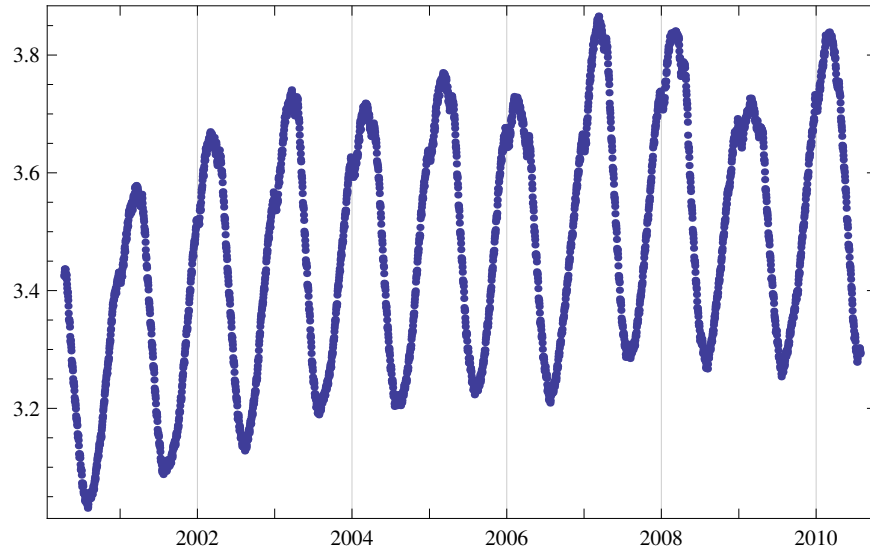


Figure 3.11: A 100 day moving average for parameter s_1 , the width of the first peak

The economic crisis that occurred during the late 2000's does not seem to have affected this parameter.

An ANOVA test was undertaken on the factors, with the results provided in Table 3.4. As can be seen, the p-values were almost equal to zero. Thus, the null hypotheses were rejected in all the four tests. Therefore, we can conclude that the means of the different months, weekdays, holiday and non-holiday days and, school holiday and school days are not equal. A Bonferroni post-hoc test confirmed the difference between most of the months and weekdays.

In an ANOVA test for the interaction between two factors, the null hypothesis was rejected for all but one of the tests. Thus, as in the previous section, there was no interaction between parameters holiday and school holiday.

In section 2.5, we said that these statistical tests are significant if they are both statis-

Table 3.4: ANOVA results for parameter s_1 - The width of the first peak

	DF	SumOfSq	MeanSq	FRatio	PValue
Month	11	172.564	15.6876	427.515	1.69704×10^{-650}
WeekDay	6	251.188	41.8647	1140.89	2.19874×10^{-840}
Holiday	1	24.6581	24.6581	671.975	2.04156×10^{-136}
SchoolHol	1	3.49824	3.49824	95.333	2.95786×10^{-22}
Month WeekDay	66	16.3922	0.248367	6.76843	5.3464910^{-54}
Holiday Month	7	3.62599	0.517999	14.1164	3.37660×10^{-18}
Month SchoolHol	9	4.51443	0.501603	13.6695	7.72556×10^{-22}
Holiday WeekDay	6	5.52147	0.920245	25.0782	2.48217×10^{-29}
SchoolHol WeekDay	6	0.986332	0.164389	4.47987	0.000158013
Holiday SchoolHol	1	0.0288407	0.0288407	0.785958	0.375382
Error	3747	137.496	0.036695		
Total	3861	620.474			

tically and practically significant. This test produced a co-efficient of determination(r^2) of 0.778, and Cohen's f^2 of 3.513. This effect size, as discussed in section 2.5, means that these variables have a large effect. Hence the results are significant.

3.6 Results for parameter s_2 - The width of the second peak

Figure 3.13 represents the width of the second peak (parameter s_2) over the period under consideration.

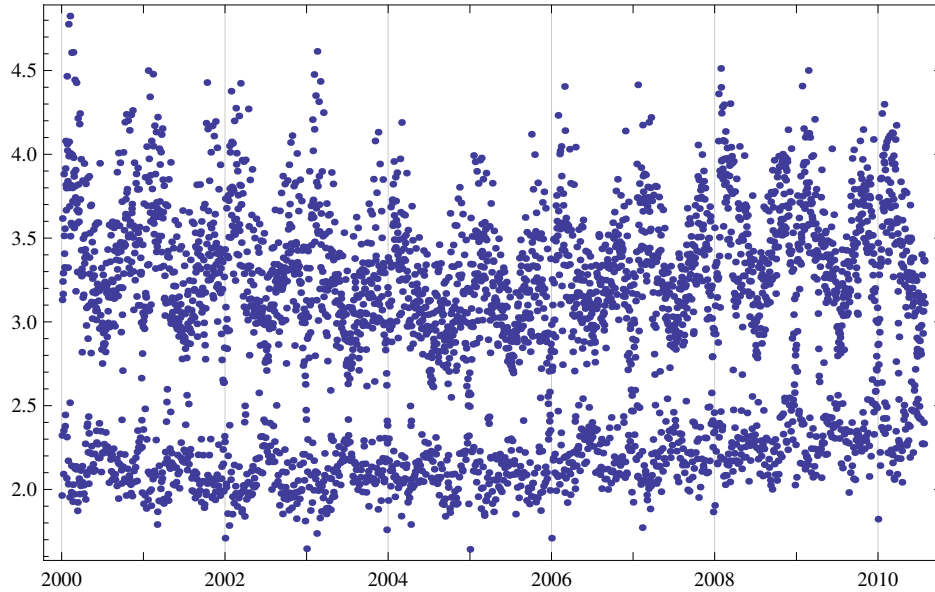


Figure 3.12: Estimates for parameter s_2 - The width of the second peak

The width of the second peak seems to be stationary for most of the time under consideration. There seems to be a slight decreasing trend during the years 2000 to 2005, and a slight increasing trend between the years 2005 to 2009 (a sort of v-shape). This so called v-shape, which can be clearly seen on the 100 day moving average in Figure 3.15, occurs only on this parameter, and it is quite a special effect. In order to understand the reasons behind this trend without speculating, one would have to observe this parameter over a longer period of time.

The data also show elements of seasonality, with the width more spread out during

the summer months when compared to the other months. For illustrative purposes, a graph for this parameter deseasonalized is included as Figure 3.14 below. Also the moving average in Figure 3.15 shows all the different trends in this parameter.

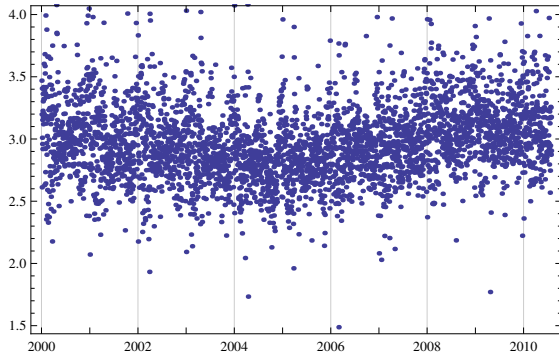


Figure 3.13: Parameter s_2 , the time of the second peak deseasonalized

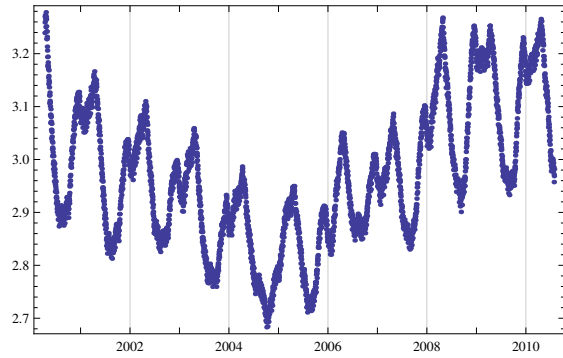


Figure 3.14: A 100-day moving average of parameter s_2

Another noticable trend of this graphic is the existence of two sort of parallel data sets. The upper set of data points consists of the weekdays (i.e. Mondays to Fridays), and the lower set of data points consists of the weekend days (i.e. Saturday and Sunday). Figures 3.16 and 3.17 below show the weekdays and the weekend days respectively. The weekdays have an average of 3.29 and are mostly in the region $[2.5 - 5]$. The weekend days have an average of 2.18 and are mostly in the region $[1.5 - 2.5]$. Although, as seen in Figure 3.16, sometimes the the width of the second peak can be in the same region as the weekend days (i.e. it can be in the region $[1.5 - 2.5]$). This means that the width of the second peak is generally higher or, to be precise, more spread out during weekdays than during weekends.

An ANOVA test was conducted for the four factors under consideration, with the results printed in Table 3.5. For all the tests, the p-values are values very close to

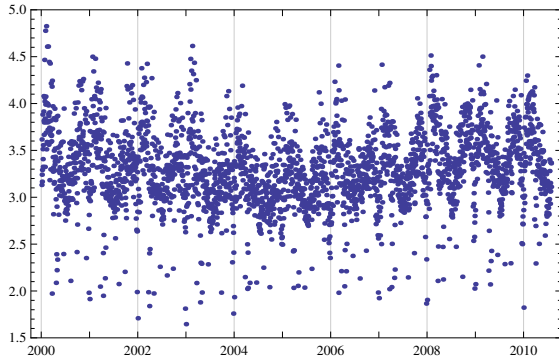


Figure 3.15: Parameter s_2 - Only the weekdays

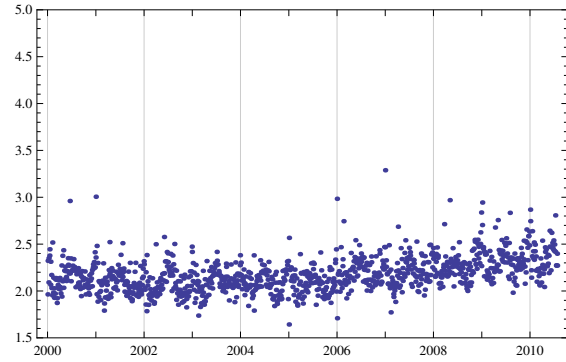


Figure 3.16: Parameter s_2 - Only the weekend days

zero, thus the null hypotheses are rejected. Thus, we conclude that there is significant difference between the means of the different months, of the different days of the week, between holiday and non-holiday days, and between school days and school holidays. A Bonferroni post-hoc test confirmed the difference between most of the months and weekdays.

The test on the interaction between the different factors, also produced p-values that are almost equal to zero. Thus, for all cases the null hypotheses are rejected. Therefore, we can conclude that there are two-way interactions between the all factors.

In section 2.5, we said that these statistical tests are significant if they are both statistically and practically significant. This test produced a co-efficient of determination(r^2) of 0.866, and Cohen's f^2 of 6.443. This effect size, as discussed in section 2.5, means that these variables have a large effect. Hence the tests are significant.

Table 3.5: ANOVA results for parameter s_2 - The width of the second peak

	DF	SumOfSq	MeanSq	FRatio	PValue
Month	11	67.4733	6.13394	113.919	3.70556×10^{-225}
WeekDay	6	1025.22	170.87	3173.37	$1.78018 \times 10^{-1463}$
Holiday	1	70.7304	70.7304	1313.59	7.63327×10^{-247}
SchoolHol	1	8.7648	8.7648	162.778	1.59246×10^{-36}
Month WeekDay	66	63.3135	0.959296	17.8159	3.01878×10^{-173}
Holiday Month	7	9.51718	1.3596	25.2502	6.37052×10^{-34}
Month SchoolHol	9	5.14129	0.571255	10.6092	2.14898×10^{-16}
Holiday WeekDay	6	40.258	6.70966	124.611	4.54370×10^{-144}
SchoolHol WeekDay	6	8.3129	1.38548	25.731	3.97006×10^{-30}
Holiday SchoolHol	1	1.1518	1.1518	21.3911	3.87152×10^{-6}
Error	3747	201.757	0.053845		
Total	3861	1501.64			

3.7 Results for parameter b_1 - The total energy used during the first peak minus the minimum demand

Figure 3.18 represents the total energy used during the first peak minus the minimum demand (parameter b_1) over the period under consideration.

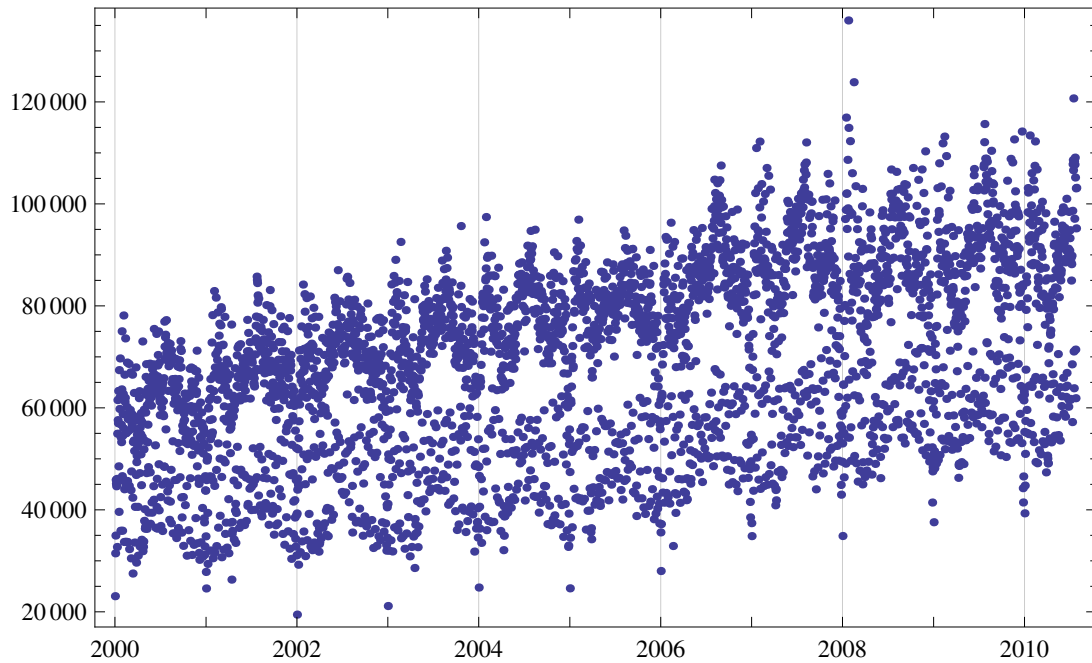


Figure 3.17: Estimates for parameter b_1 - The total energy used during the first peak minus the minimum demand

The graphic for the volume of energy used during the first peak has an upward trend in the years 2000 to 2007. After the year 2005, possibly because of the global economic crisis, the data had neither an upward nor a downward trend. This parameter, like all the previous parameters, has a seasonal component within its trends. These

trends are shown in Figures 3.19 and 3.20 below. As stated previously in section 3.1, this parameter, unlike the minimum demand, was not affected by the global economic crisis.

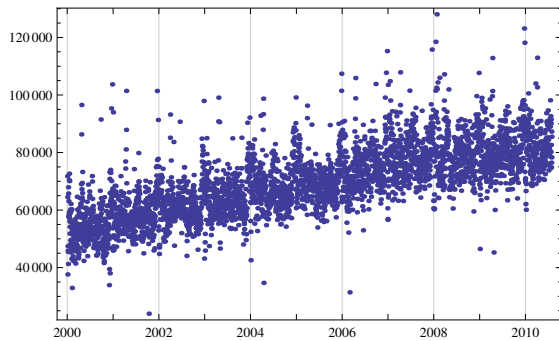


Figure 3.18: Parameter b_1 , the volume of energy for the first peak deseasonalised

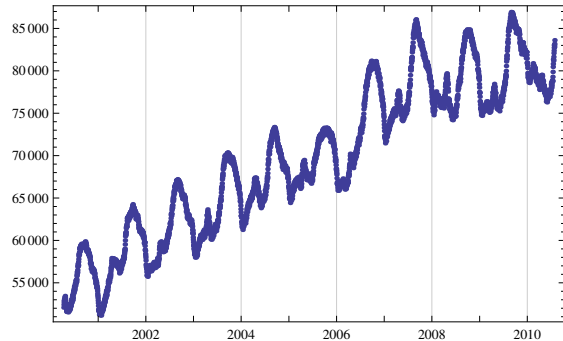


Figure 3.19: A 100-day moving average of parameter b_1

The data also show the two sets of parallel data points, with the lower dataset being the weekend days and the upper dataset being the week days, and the average for the weekend days being 50228.6 and the average for the weekdays being 77319.3. Figures 3.21 and 3.22 clearly show these two data points and their trends.

An ANOVA test was undertaken on the factors, with the results provided in Table 3.6. As can be seen, the p-values were almost equal to zero. Thus, the null hypotheses were rejected for all four tests. Therefore, we can conclude that the means of the different months, weekdays, holiday and non-holiday days and, school holiday and school days are not equal. A Bonferroni post-hoc test confirmed the difference between most of the months and weekdays.

An ANOVA test for interaction between the factors was also undertaken, and the null hypotheses were rejected for all but one of the tests. Thus, there is no interac-

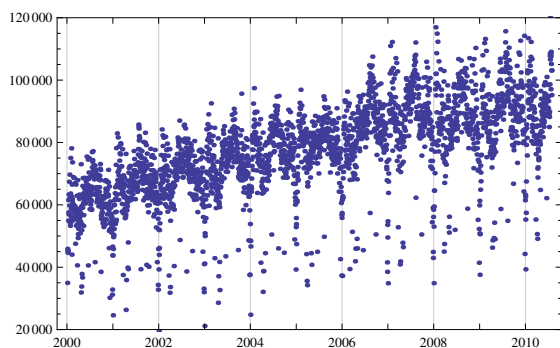


Figure 3.20: Parameter b_1 - Only the weekdays

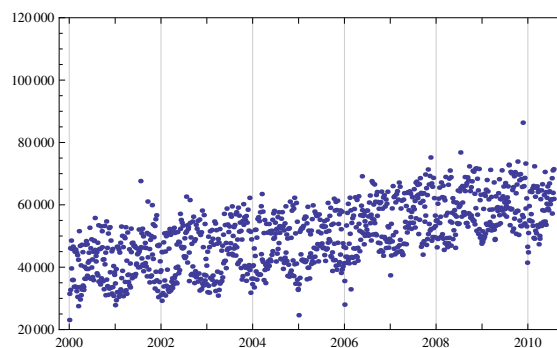


Figure 3.21: Parameter b_1 - Only the weekend days

tion between parameters holiday and school holiday.

In section 2.5, we said that these statistical tests are significant if they are both statistically and practically significant. This test produced a co-efficient of determination(r^2) of 0.670, and Cohen's f^2 of 2.026. This effect size, as discussed in section 2.5, means that these variables have a large effect. Hence the tests are significant.

Table 3.6: ANOVA results for parameter b_1 - The volume of energy used during the first peak

	DF	SumOfSq	MeanSq	FRatio	PValue
Month	11	7.87149×10^{10}	7.1559×10^9	62.6897	8.48785×10^{-129}
WeekDay	6	6.27251×10^{11}	1.04542×10^{11}	915.844	1.66688×10^{-729}
Holiday	1	8.75762×10^{10}	8.75762×10^{10}	767.216	8.62357×10^{-154}
SchoolHol	1	1.10436×10^{10}	1.10436×10^{10}	96.7484	1.47311×10^{-22}
Month WeekDay	66	2.17172×10^{10}	3.29049×10^8	2.88265	1.56643×10^{-13}
Holiday Month	7	4.94542×10^9	7.06488×10^8	6.18923	3.18383×10^{-7}
Month SchoolHol	9	1.18729×10^{10}	1.31921×10^9	11.557	4.48753×10^{-18}
Holiday WeekDay	6	2.06855×10^{10}	3.44759×10^9	30.2028	1.42490×10^{-35}
SchoolHol WeekDay	6	2.52859×10^9	4.21432×10^8	3.69197	0.00116215
Holiday SchoolHol	1	1.1947×10^8	1.1947×10^8	1.04663	0.306352
Error	3747	4.27713×10^{11}	1.14148×10^8		
Total	3861	1.29417×10^{12}			

3.8 Results for parameter b_2 - The total energy used during the second peak minus the minimum demand

Figure 3.23 represents the total energy used during the second peak (parameter b_2) over the period under consideration.

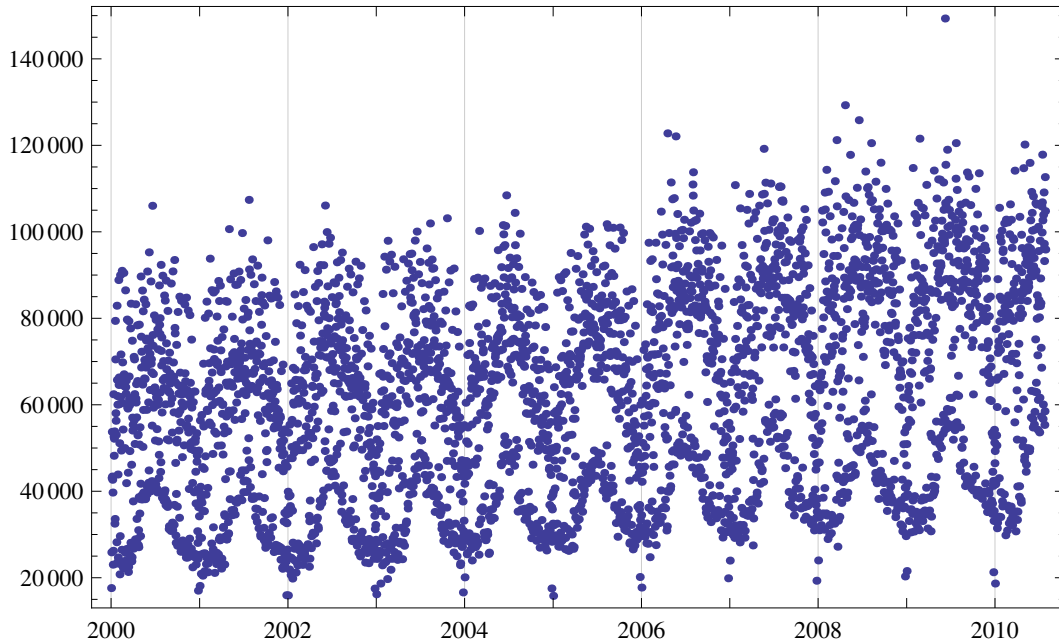


Figure 3.22: Estimates for parameter b_2 - The total energy used during the second peak minus the minimum demand

The graphic for the volume of energy used during the second peak has a slight upward trend. The data also contain elements of seasonality, with less energy being used during the summer months of November, December and January, and more energy being used during the other months. These trends can be seen on Figures 3.24 and 3.25.

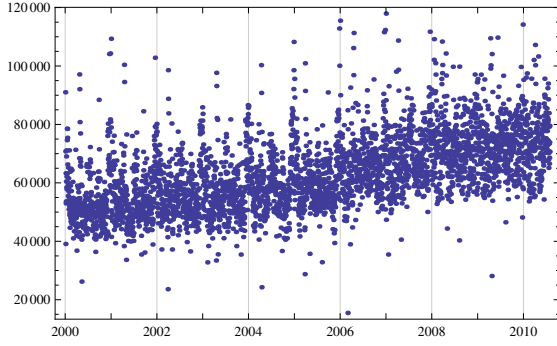


Figure 3.23: Parameter b_2 , the volume of energy used during the second peak deseasonalised

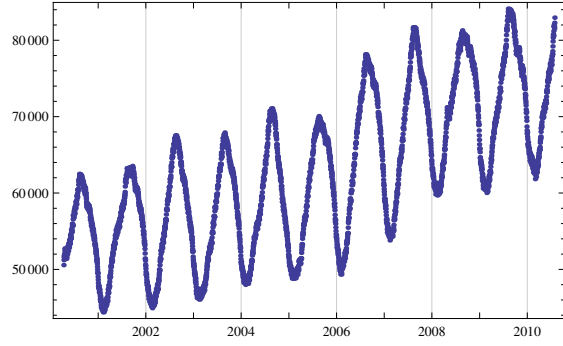


Figure 3.24: A 100-day moving average of parameter b_2 , the volume of energy used during the second peak

There is also some element of two data sets within this data, with the lower dataset being the weekend days, and the upper dataset being the other days of the week. The lower dataset is shown in Figure 3.26, and is in the region $[20\,000 - 60\,000]$ with the mean at 37038.5. The upper dataset is shown in Figure 3.27 and is in the region $[40\,000 - 120\,000]$ with the mean at 73125.8.

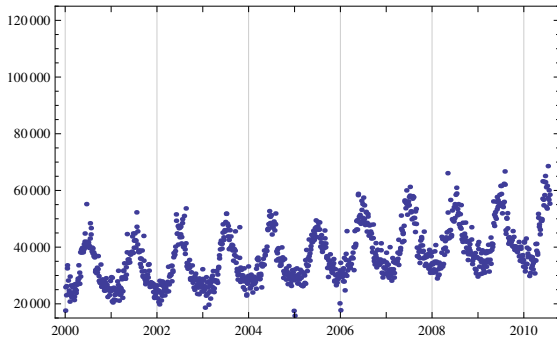


Figure 3.25: Parameter b_2 - Only the weekend days

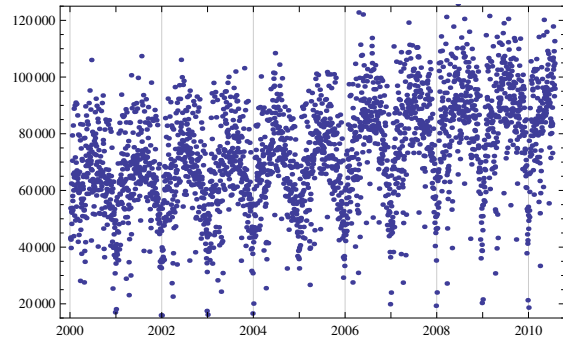


Figure 3.26: Parameter b_2 - Only the weekdays

An ANOVA test was also done for this parameter, with the results shown in Table 3.7. All the p-values are values very close to zero. Thus the null hypotheses were rejected in all four tests. Thus, the means of the different months of the year, of the different days of the week, the holiday and non-holiday days, and between school holiday and non-school holidays are all not equal. The Bonferroni post-hoc test shows that there is a difference between most of the months and also between most of the weekdays.

The tests on whether there is an interaction between the variables also produced small p-values. Therefore, the null hypotheses were rejected for all four cases and therefore, there are two-way interactions between the four variables.

In section 2.5, we said that these statistical tests are significant if they are both statistically and practically significant. This test produced a co-efficient of determination(r^2) of 0.813, and Cohen's f^2 of 4.316. This effect size, as discussed in section 2.5, means that these variables have a large effect. Hence the tests are significant.

Table 3.7: ANOVA results for parameter b_2 - The volume of energy used during the second peak

	DF	SumOfSq	MeanSq	FRatio	PValue
Month	11	2.79785×10^{11}	2.5435×10^{10}	243.675	1.74562×10^{-428}
WeekDay	6	1.19668×10^{12}	1.99447×10^{11}	1910.76	$9.82873 \times 10^{-1135}$
Holiday	1	9.24515×10^{10}	9.24515×10^{10}	885.713	6.72485×10^{-175}
SchoolHol	1	2.19385×10^{10}	2.19385×10^{10}	210.177	2.21359×10^{-46}
Month WeekDay	66	3.64061×10^{10}	5.51608×10^8	5.28457	7.31281×10^{-38}
Holiday Month	7	6.29715×10^9	8.99592×10^8	8.61836	1.58938×10^{-10}
Month SchoolHol	9	8.58406×10^9	9.53784×10^8	9.13754	8.42066×10^{-14}
Holiday WeekDay	6	3.50094×10^{10}	5.8349×10^9	55.9001	2.11567×10^{-66}
SchoolHol WeekDay	6	1.0383×10^{10}	1.73049×10^9	16.5786	5.88016×10^{-19}
Holiday SchoolHol	1	6.62324×10^8	6.62324×10^8	6.34526	0.0118108
Error	3747	3.91115×10^{11}	1.04381×10^8		
Total	3861	2.07932×10^{12}			

Chapter 4

Conclusions and Recommendations

The de Moivre dual model used to model the daily electricity profile has seven (7) parameters. The most exciting part of this model is that, unlike for instance a polynomial model, all the parameters have natural interpretations. Also, the parameters can be divided into two categories, volume of energy and shape parameters. With parameters a , b_1 and b_2 being the volume of energy parameters and parameters m_1 , m_2 , s_1 and s_2 being the shape parameters. An illustrative summary of the parameters is included at the end of this section as Figure 4.7.

This simplification of the daily electricity demand profile into “components”, if we may call them that, will greatly enhance electricity demand forecasting accuracy. With more accurate forecasts, the power utility will be more efficient, and greater efficiency means fewer costs for both the the power utility and the electricity consumer.

Of the three volume of energy parameters, only the shape of parameter a was similar to the shape of the overall daily demand. Figures 4.1 and 4.2 show the 100-day moving averages of the daily demand and the daily minimum demand (parameter a) respectively, and these graphics confirm what is said in the previous sentence. Also, only parameter a clearly responded to the global economic crisis.

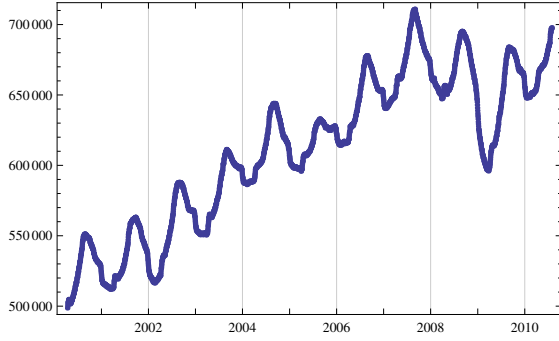


Figure 4.1: A 100 day moving average for the daily demand

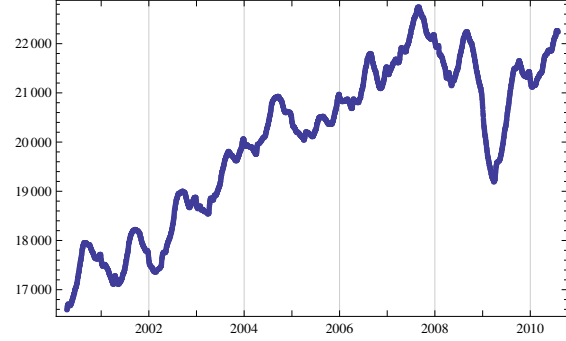


Figure 4.2: A 100 day moving average for parameter a

The volume of energy parameters b_1 and b_2 were quite consistent in their trends. Their 100-day moving averages are shown in the same plot in Figure 4.3, with parameter b_1 being the blue graph and parameter b_2 being the green graph. During summer, the volume of energy used during the second peak was substantially less than the one used during the first peak.

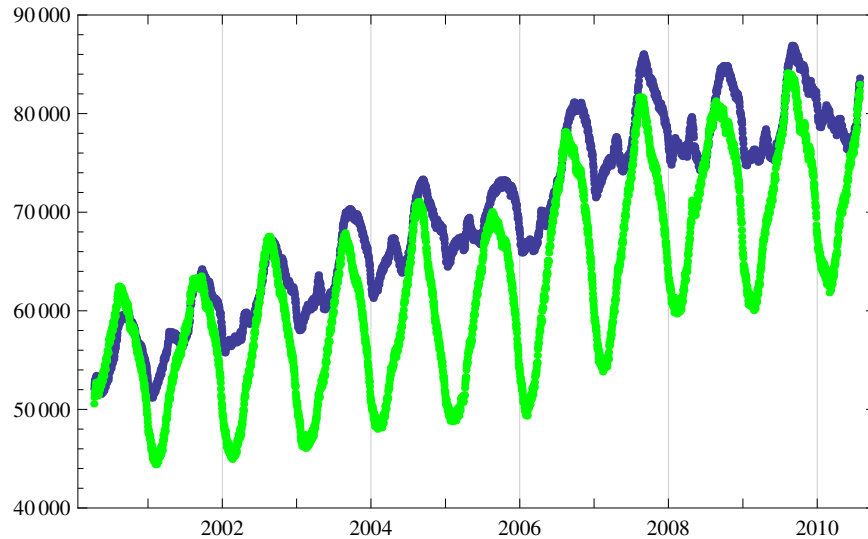


Figure 4.3: 100-day moving averages of parameters b_1 (blue) and b_2 (green)

The 100-day moving averages of the times of the first and second peaks are shown in Figures 4.4 and 4.5 respectively. The trends in the graphics are for the most part of similar shapes. The time of the second peak seems to have been affected by the global economic crisis. The trends of these parameters seem stationary and consistent, with the times of the peaks occurring later in the day during summer than during winter.

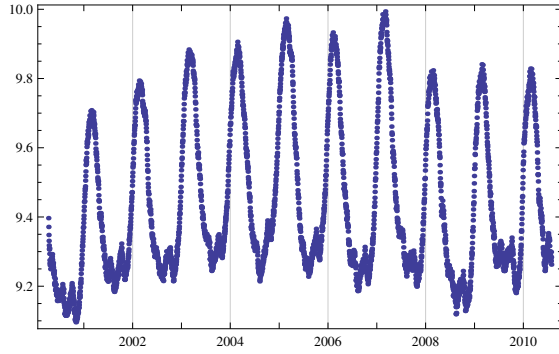


Figure 4.4: A 100-day moving average of parameter m_1

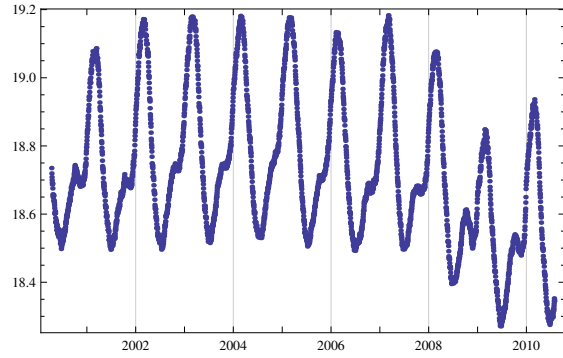


Figure 4.5: A 100-day moving average of parameter m_2

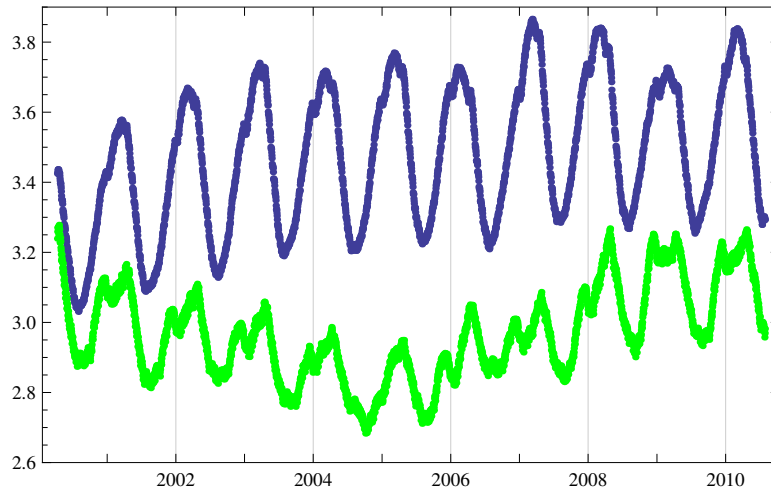


Figure 4.6: 100-day moving averages of parameters s_1 (blue) and s_2 (green)

The shapes of the widths of the peaks differed in the overall shape, with param-

eter s_1 having consistent and stationary trends, while parameter s_2 had that special v shape, which occurred only in this parameter. Their 100-day moving averages are shown in Figure 4.6, with parameter s_1 being the blue graph and parameter s_2 being the green graph. Both widths of the peaks widened during summer and became thinner during winter. In general, the width of the first peak was always thinner than the width of the first peak.

Another important noticable point was the existence of what we called two data sets within the parameters. One dataset consisted of the weekdays (i.e. Monday, Tuesday, Wednesday, Thursday and Friday) and the other data set consisted of the weekend days (i.e. Saturday and Sunday).

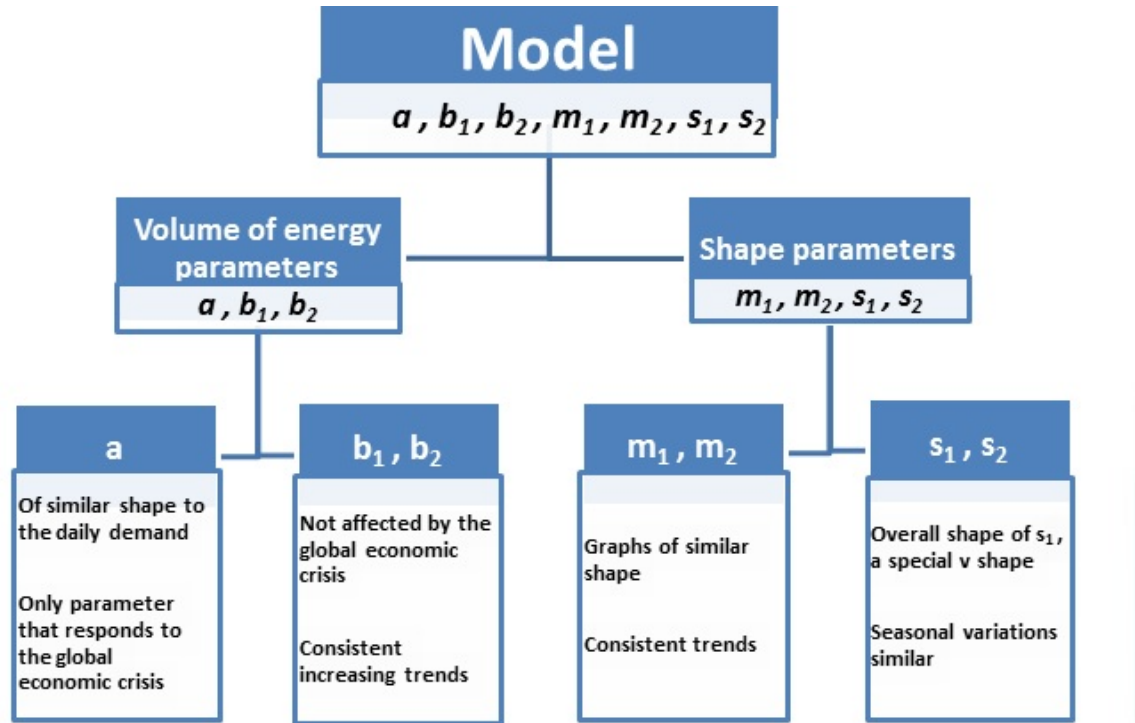


Figure 4.7: An illustrative summary of the model

Future Studies

This study has been done using the national data. It might be of interest to see how the model will perform on provincial and regional data. The model might also be applied on data from the big industrial companies. The differences in the profiles might be of interest to the power supplier.

The de Moivre (dM) models were used in building an integrated model and preliminary studies show that for the first peak a skew dM (log-type) model might be better (as this peak is not symmetric).

Studies on the weekly and monthly profiles will also be of interest. The models produced might have answers on the performance of the electricity demand profiles in the short to long term.

Bibliography

- [1] Coe R. (2002). What effect size is, and why it is important. Accessed on <http://www.leeds.ac.uk/educol/documents/00002182.htm>.
- [2] Soares L. and Souza L. (2006). Forecasting electricity demand using generalized long memory. *International Journal of Forecasting*, 22, 17–28.
- [3] Taylor J. W. (2010). Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research*, 204, 139–152.
- [4] McCullagh P. (2002) What is a statistical model? *The Annals of Statistics*, Volume 30, No. 5, 1225–1310.
- [5] Meade N. (2000). Evidence for the selection of forecasting methods. *Journal of Forecasting*, 19, 515–535.
- [6] Bunn D. W. and Farmer E. D. (1985). Review of short-term forecasting methods in the electric power industry. *New York, Wiley*, 13–30.
- [7] www.mpoweruk.com/electricity_demand.htm.
- [8] [www.http://www.info.gov.za/view/DownloadFileAction?id=96381](http://www.info.gov.za/view/DownloadFileAction?id=96381). Framework for South Africa’s response to the Global Economic Crisis.
- [9] <http://en.wikipedia.org>.

Appendix A

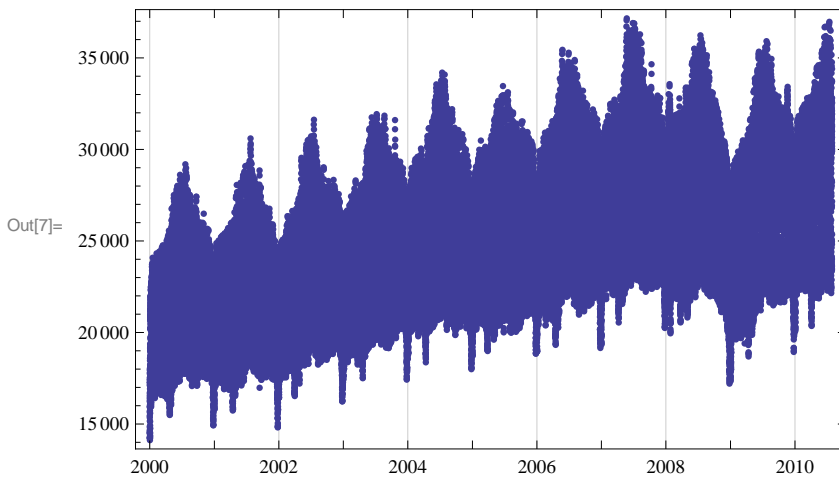
Mathematica Script

Attached in the following pages is the entire Mathematica script.

```

In[1]:= SetDirectory["e://Analysis"];
In[2]:= data = Import["data.csv"];
In[3]:= td = Table[{data[[i, 4]], data[[i, 2]]}, {i, 1, Length[data]}];
In[4]:= Clear[data]
In[5]:= (* td - array of time and demand only *)
In[6]:= Length[td]
Out[6]= 92 712
In[7]:= DateListPlot[td]

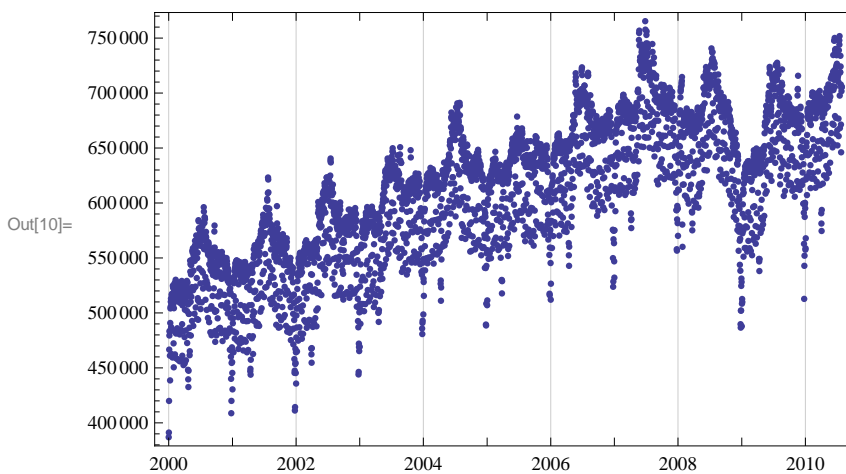
```



```

In[8]:= Length[dailydata]
Out[8]= 0
In[9]:= dailydata = Table[{td[[24 * (k - 1) + 1, 1]],
    Sum[td[[i, 2]], {i, 24 * (k - 1) + 1, 24 * (k - 1) + 24}]}, {k, 1, Length[td] / 24}];
In[10]:= DateListPlot[dailydata]

```



```

In[11]:= Take[td, 5]
Out[11]= {{3 155 673 600, 15 660}, {3 155 677 200, 15 449},
    {3 155 680 800, 14 852}, {3 155 684 400, 14 577}, {3 155 688 000, 14 510}}

```

```

In[12]:= Needs["Calendar`"]

In[13]:= (* select 24 hours of the day defined by "day" (DateList), "shift"
          (number of hours: positive or negative) starts the day later or earlier *)
          (* returns also the day of week and if all 24 values available *)
select24[d_, day_, shift_] := Module[{at, wd, dd, ddd, begin, N},
  at = AbsoluteTime[day];
  begin = at + shift * 3600;
  wd = DayOfWeek[day];
  ddd = Select[d, (#[[1]] ≥ begin) && (#[[1]] ≤ begin + 23 * 3600) &];
  dd = ddd;
  N = Length[dd];
  dd[[All, 1]] = Table[DateList[ddd[[1, 1]]][[4]] + i - 1, {i, 1, N}];
  {dd, ddd, N, wd}];

In[14]:= f[x_, a_, b1_, m1_, s1_, b2_, m2_, s2_] := a + b1 * nd[m1, s1, x] + b2 * nd[m2, s2, x]

In[15]:= nd[m_, s_, x_] := PDF[NormalDistribution[m, s]][x]

In[16]:= result = {};
AbsMaxResiduals = {};
Residuals = {};
MeanResiduals = {};
For[t = Min[td[[All, 1]]], t ≤ Max[td[[All, 1]]] - 86400,
  t = t + 86400, (* 86400 is one day in seconds *)
  shift = 2;
  test = select24[td, DateList[t], shift];

  aa = Min[test[[1, All]][[All, 2]]];
  i = Range[5, 12]; em1m = Max[test[[1]][[All, 2]][[i]]];
  em1m1 = Select[test[[1]], #[[2]] == em1m &]; em1 = em1m1[[1, 1]];
  k = Range[15, 23]; em2m = Max[test[[1]][[All, 2]][[k]]];
  em2m2 = Select[test[[1]], #[[2]] == em2m &]; em2 = em2m2[[1, 1]];

  b1b2 = (Sum[test[[1]][[All, 2]][[j]], {j, 1, 24}] - aa * 24) / 2;

  fit1 = NonlinearModelFit[test[[1]], a + b1 * nd[m1, s1, x] + b2 * nd[m2, s2, x],
    {{a, aa}, {b1, b1b2}, {m1, em1}, {s1, 3}, {b2, b1b2}, {m2, em2}, {s2, 3}}, x];
  fit = FindFit[test[[1]], a + b1 * nd[m1, s1, x] + b2 * nd[m2, s2, x],
    {{a, 18000}, {b1, b1b2}, {m1, 10}, {s1, 3}, {b2, b1b2}, {m2, 20}, {s2, 3}}, x];
  result = Append[result, {DateList[t], shift,
    a, b1, m1, s1, b2, m2, s2} /. fit];
  AbsMaxResiduals = Append[AbsMaxResiduals,
    {DateList[t], Max[Abs[fit1["FitResiduals"]]]}];
  Residuals = Append[Residuals, {DateList[t], fit1["FitResiduals"]};
  MeanResiduals =
    Append[MeanResiduals, {DateList[t], Mean[fit1["FitResiduals"]}]];

```

```
In[21]:= result[[1]]
Residuals[[1]]
```

```
Out[21]= {{2000, 1, 1, 0, 0, 0.}, 2, 14328.6,
          23064.8, 11.0284, 2.82449, 17607.2, 19.6468, 1.96372}
```

```
Out[22]= {{2000, 1, 1, 0, 0, 0.},
          {503.697, 191.041, 34.0442, -388.595, -343.467, -155.764, 238.879, 267.136,
           25.5478, -57.4404, -116.066, -182.642, -35.03, 152.58, 325.571, 162.577,
           -444.082, 34.5029, 159.768, 89.9739, -133.911, 5.53966, -129.172, -204.688}}
```

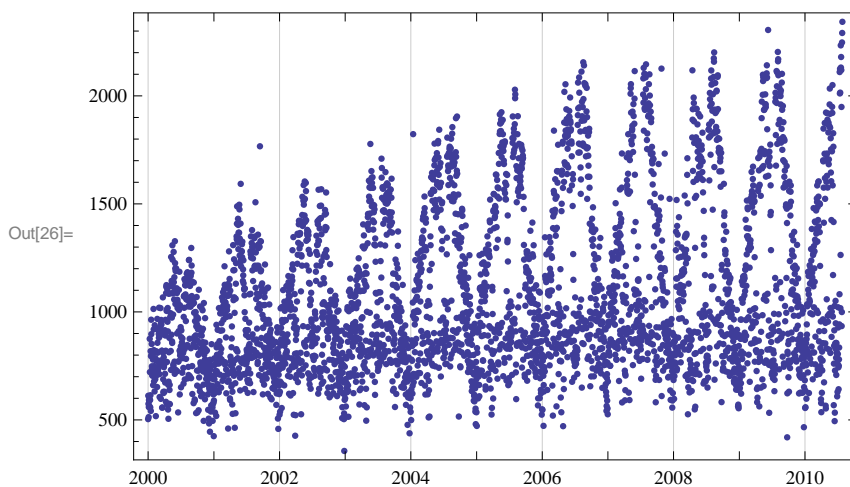
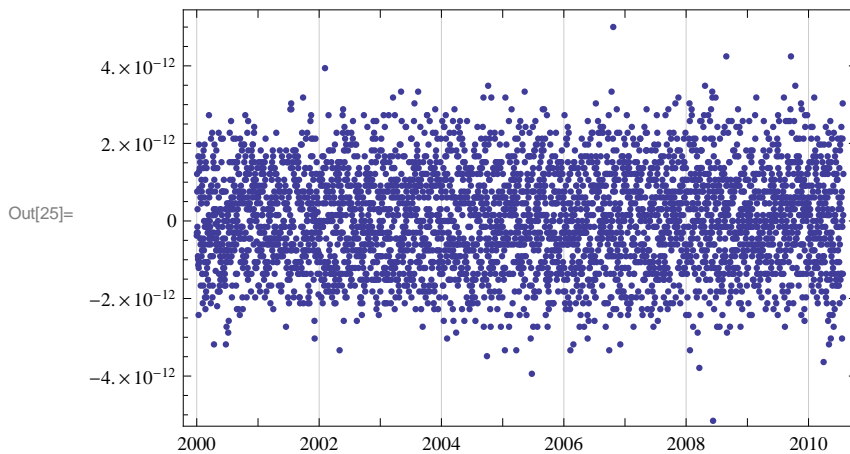
```
In[23]:= Integrate[
          result[[1, 3]] + result[[1, 4]] * nd[result[[1, 5]], Sqrt[result[[1, 6]]], x] +
          result[[1, 7]] * nd[result[[1, 8]], Sqrt[result[[1, 9]]], x], {x, 0, 24}]
```

```
Out[23]= 384542.
```

```
In[24]:= Length[result]
```

```
Out[24]= 3862
```

```
In[25]:= DateListPlot[MeanResiduals]
DateListPlot[AbsMaxResiduals]
```



```
In[27]:= Max[Residuals[[All, 2]]]
```

```
Out[27]= 2342.79
```

```

In[28]:= varA0 = Table[{result[[i, 1]], result[[i, 3]], 0}, {i, 1, Length[result]}];
varA = Table[{result[[i, 1]], result[[i, 3]]}, {i, 1, Length[result]}];
DateListPlot[varA];

In[31]:= varM10 = Table[{result[[i, 1]], result[[i, 5]], 0}, {i, 1, Length[result]}];
varM1 = Table[{result[[i, 1]], result[[i, 5]]}, {i, 1, Length[result]}];
DateListPlot[varM1];

In[34]:= varM20 = Table[{result[[i, 1]], result[[i, 8]], 0}, {i, 1, Length[result]}];
varM2 = Table[{result[[i, 1]], result[[i, 8]]}, {i, 1, Length[result]}];
DateListPlot[varM2];

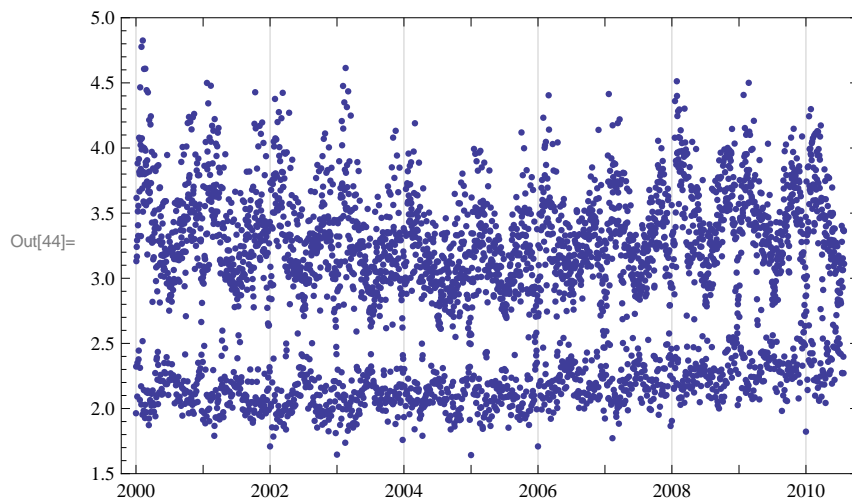
In[37]:= Mean[Table[varM2[[i, 2]], {i, 1, Length[varM2]}]]

Out[37]= 18.7234

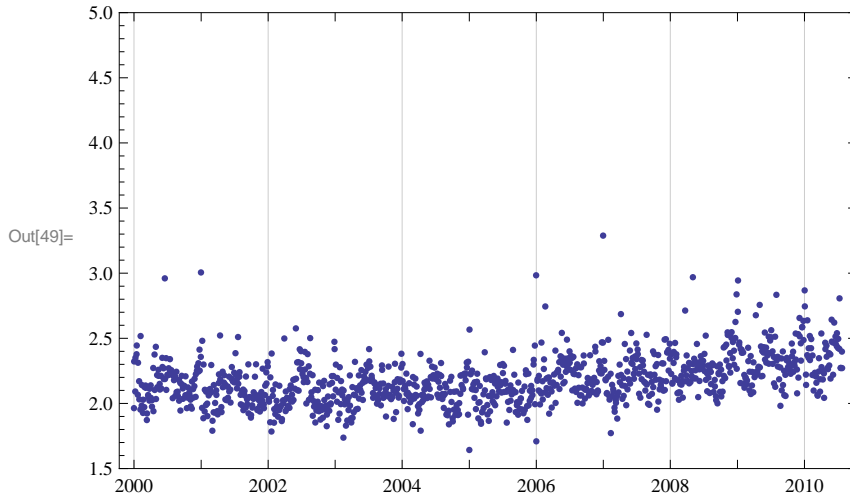
In[38]:= varS1 = Table[{result[[i, 1]], result[[i, 6]]}, {i, 1, Length[result]}];
varS10 = Table[{result[[i, 1]], result[[i, 6]], 0}, {i, 1, Length[result]}];
DateListPlot[varS1];

In[41]:= varS2 = Table[{result[[i, 1]], result[[i, 9]]}, {i, 1, Length[result]}];
varS20 = Table[{result[[i, 1]], result[[i, 9]], 0}, {i, 1, Length[result]}];
DateListPlot[varS2];
DateListPlot[varS2, PlotRange -> {1.5, 5}]

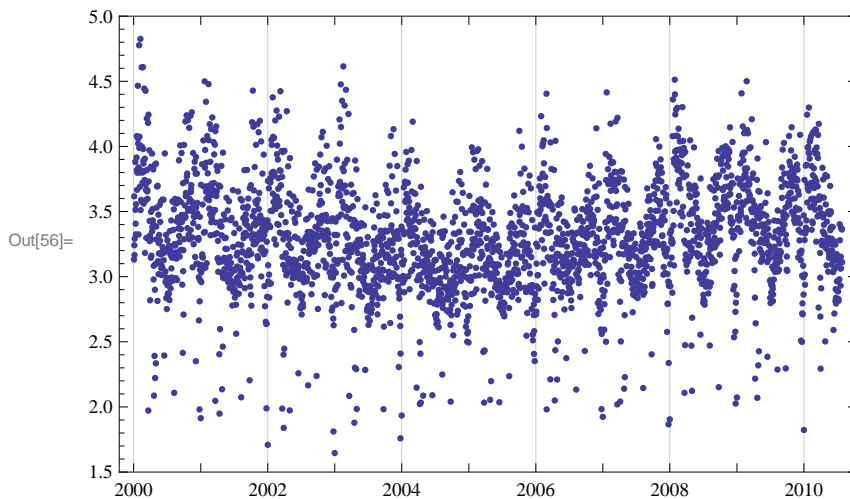
```



```
In[45]:= << Calendar`
vars2sa = Select[vars2, DayOfWeek[#[[1]]] == Saturday &];
vars2su = Select[vars2, DayOfWeek[#[[1]]] == Sunday &];
wk = Union[vars2sa, vars2su];
DateListPlot[wk, PlotRange -> {1.5, 5}]
```



```
In[50]:= vars2mo = Select[vars2, DayOfWeek[#[[1]]] == Monday &];
vars2tu = Select[vars2, DayOfWeek[#[[1]]] == Tuesday &];
vars2we = Select[vars2, DayOfWeek[#[[1]]] == Wednesday &];
vars2th = Select[vars2, DayOfWeek[#[[1]]] == Thursday &];
vars2fr = Select[vars2, DayOfWeek[#[[1]]] == Friday &];
wkd = Union[vars2mo, vars2tu, vars2we, vars2th, vars2fr];
DateListPlot[wkd, PlotRange -> {1.5, 5}]
```



```
In[57]:= Mean[Table[wk[[i, 2]], {i, 1, Length[wk]}]]
Mean[Table[wkd[[i, 2]], {i, 1, Length[wkd]}]]
```

Out[57]= 2.17841

Out[58]= 3.29152

```
In[59]:= varB1 = Table[{result[[i, 1]], result[[i, 4]]}, {i, 1, Length[result]}];
varB10 = Table[{result[[i, 1]], result[[i, 4]], 0}, {i, 1, Length[result]}];
DateListPlot[varB1];
```

```

In[62]:= varB2 = Table[{result[[i, 1]], result[[i, 7]]}, {i, 1, Length[result]};
varB20 = Table[{result[[i, 1]], result[[i, 7]], 0}, {i, 1, Length[result]};
DateListPlot[varB2];

In[210]:= d1 = Import["Database.csv"];

In[211]:= d1[[2]]

Out[211]= {demand, year, month, day, jan, feb, mar, apr, may, jun, jul, aug, sep,
oct, nov, dec, sun, mon, tue, wed, thu, fri, sat, pub hol, day before,
day after, sch hol, univ hol, univ exmz, relg hol, dm prev, time}

In[212]:= d = Drop[d1, 2];

In[213]:= md =
Table[{varA[[i, 2]], AbsoluteTime[{d[[i, 2]], d[[i, 3]], d[[i, 4]]}], d[[i, 5]],
d[[i, 6]], d[[i, 7]], d[[i, 8]], d[[i, 9]], d[[i, 10]], d[[i, 11]], d[[i, 12]],
d[[i, 13]], d[[i, 14]], d[[i, 15]], d[[i, 16]]}, {i, 1, Length[varA]};
md[[
1]]

Out[214]= {14 328.6, 3 029 443 200, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0}

In[215]:= d[[1]]

Out[215]= {357 330, 1996, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1}

In[216]:= d2 = Table[
Append[d[[i]], DateList[AbsoluteTime[{d[[i, 2]], d[[i, 3]], d[[i, 4]]}]]],
{i, 1, Length[d]};

In[217]:= d2[[1]]

Out[217]= {357 330, 1996, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1, {1996, 1, 1, 0, 0, 0.}}

In[218]:= Dimensions[d2]

Out[218]= {5844, 33}

In[219]:= For[i = 1, i ≤ Length[varA], i++,
w = Round[AbsoluteTime[varA[[i, 1]]]];
aF[w] = varA[[i, 2]];
y = Round[AbsoluteTime[varM1[[i, 1]]]];
m1F[y] = varM1[[i, 2]];
m2F[Round[AbsoluteTime[varM1[[i, 1]]]]] = varM2[[i, 2]];
s1F[Round[AbsoluteTime[varS1[[i, 1]]]]] = varS1[[i, 2]];
s2F[Round[AbsoluteTime[varS2[[i, 1]]]]] = varS2[[i, 2]];
b1F[Round[AbsoluteTime[varB1[[i, 1]]]]] = varB1[[i, 2]];
b2F[Round[AbsoluteTime[varB2[[i, 1]]]]] = varB2[[i, 2]]

In[220]:= xx = Round[AbsoluteTime[varA[[1, 1]]]]
aF[xx]

Out[220]= 3 155 673 600

Out[221]= 14 328.6

```



```

In[222]:= d3 = Table[
  Append[d2[[i]], aF[Round[AbsoluteTime[d2[[i, 33]]]]], {i, 1, Length[d2]}];
d4 = Table[Append[d2[[i]], m1F[Round[AbsoluteTime[d2[[i, 33]]]]],
  {i, 1, Length[d2]}];
d5 = Table[Append[d2[[i]], m2F[Round[AbsoluteTime[d2[[i, 33]]]]],
  {i, 1, Length[d2]}];
d6 = Table[Append[d2[[i]], s1F[Round[AbsoluteTime[d2[[i, 33]]]]],
  {i, 1, Length[d2]}];
d7 = Table[Append[d2[[i]], s2F[Round[AbsoluteTime[d2[[i, 33]]]]],
  {i, 1, Length[d2]}];
d8 = Table[Append[d2[[i]], b1F[Round[AbsoluteTime[d2[[i, 33]]]]],
  {i, 1, Length[d2]}];
d9 = Table[Append[d2[[i]], b2F[Round[AbsoluteTime[d2[[i, 33]]]]],
  {i, 1, Length[d2]}];

```

```

In[229]:= d3[[1]]

```

```

Out[229]= {357330, 1996, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0,
  0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 1, {1996, 1, 1, 0, 0, 0.}, aF[3029443200]}

```

```

In[230]:= d3Clean = Select[d3, NumberQ[#[[34]]] &];
d4Clean = Select[d4, NumberQ[#[[34]]] &];
d5Clean = Select[d5, NumberQ[#[[34]]] &];
d6Clean = Select[d6, NumberQ[#[[34]]] &];
d7Clean = Select[d7, NumberQ[#[[34]]] &];
d8Clean = Select[d8, NumberQ[#[[34]]] &];
d9Clean = Select[d9, NumberQ[#[[34]]] &];

```

```

In[237]:= Length[d3Clean]

```

```

Out[237]= 3862

```

```

In[238]:= dataA = Table[{d3Clean[[i, 3]], DayOfWeek[d3Clean[[i, 33]]], d3Clean[[i, 24]],
  d3Clean[[i, 27]], d3Clean[[i, 34]]}, {i, 1, Length[d3Clean]}];
dataM1 = Table[{d4Clean[[i, 3]], DayOfWeek[d4Clean[[i, 33]]], d4Clean[[i, 24]],
  d4Clean[[i, 27]], d4Clean[[i, 34]]}, {i, 1, Length[d4Clean]}];
dataM2 = Table[{d5Clean[[i, 3]], DayOfWeek[d5Clean[[i, 33]]], d5Clean[[i, 24]],
  d5Clean[[i, 27]], d5Clean[[i, 34]]}, {i, 1, Length[d5Clean]}];
datas1 = Table[{d6Clean[[i, 3]], DayOfWeek[d6Clean[[i, 33]]], d6Clean[[i, 24]],
  d6Clean[[i, 27]], d6Clean[[i, 34]]}, {i, 1, Length[d6Clean]}];
datas2 = Table[{d7Clean[[i, 3]], DayOfWeek[d7Clean[[i, 33]]], d7Clean[[i, 24]],
  d7Clean[[i, 27]], d7Clean[[i, 34]]}, {i, 1, Length[d7Clean]}];
datab1 = Table[{d8Clean[[i, 3]], DayOfWeek[d8Clean[[i, 33]]], d8Clean[[i, 24]],
  d8Clean[[i, 27]], d8Clean[[i, 34]]}, {i, 1, Length[d8Clean]}];
datab2 = Table[{d9Clean[[i, 3]], DayOfWeek[d9Clean[[i, 33]]], d9Clean[[i, 24]],
  d9Clean[[i, 27]], d9Clean[[i, 34]]}, {i, 1, Length[d9Clean]}];

```

```

In[245]:= dataA[[2]]

```

```

Out[245]= {1, Sunday, 0, 1, 13923.}

```

```

In[246]:= Length[dataA]

```

```

Out[246]= 3862

```

```

In[247]:= Needs["ANOVA`"]

```

```
In[355]:= anA[[1, 2, 1, 11, 2]]
          anA[[1, 2, 1, 12, 2]]
```

```
Out[355]= 1.09447 × 1010
```

```
Out[356]= 1.27121 × 1010
```

```
In[357]:= rsqr = 1 - (anA[[1, 2, 1, 11, 2]] / anA[[1, 2, 1, 12, 2]])
```

```
Out[357]= 0.139029
```

```
In[358]:= fsq = rsqr / (1 - rsqr)
```

```
Out[358]= 0.16148
```

```
In[252]:= anA = ANOVA[dataA, {Month, WeekDay, Holiday, SchoolHol, Month * WeekDay,
      Month * Holiday, Month * SchoolHol, WeekDay * Holiday, WeekDay * SchoolHol,
      Holiday * SchoolHol}, {Month, WeekDay, Holiday, SchoolHol},
      PostTests → Bonferroni, SignificanceLevel → .01];
anA[[1]]
anA[[2]]
anA[[3]]
```

```
Out[253]= ANOVA →
```

	DF	SumOfSq	MeanSq	FRatio	PValue
Month	11	7.74482×10^8	7.04075×10^7	24.1045	1.21823
WeekDay	6	5.65403×10^8	9.42339×10^7	32.2616	4.49847
Holiday	1	3.78728×10^7	3.78728×10^7	12.966	0.00032
SchoolHol	1	1.15439×10^7	1.15439×10^7	3.95213	0.04688
Month WeekDay	66	4.74096×10^7	718 327.	0.245924	1.
Holiday Month	7	7.72192×10^7	1.10313×10^7	3.77665	0.00043
Month SchoolHol	9	1.57744×10^8	1.75271×10^7	6.00054	2.18167
Holiday WeekDay	6	1.76551×10^7	2.94251×10^6	1.00739	0.41842
SchoolHol WeekDay	6	3.28157×10^7	5.46928×10^6	1.87245	0.08169
Holiday SchoolHol	1	4.52045×10^7	4.52045×10^7	15.4761	0.00008
Error	3747	1.09447×10^{10}	2.92093×10^6		
Total	3861	1.27121×10^{10}			
All			20 013.3		
Month[1]			19 417.5		
Month[2]			19 691.8		
Month[3]			19 671.9		
Month[4]			19 589.4		
Month[5]			20 008.		
Month[6]			20 701.6		
Month[7]			20 940.7		
Month[8]			20 106.		
Month[9]			20 034.3		
Month[10]			20 172.5		
Month[11]			20 222.		
Month[12]			19 605.8		
WeekDay[Friday]			20 108.9		
WeekDay[Monday]			19 167.3		
WeekDay[Saturday]			20 531.3		
WeekDay[Sunday]			19 992.3		
WeekDay[Thursday]			20 157.6		
WeekDay[Tuesday]			20 035.3		

WeekDay[Wednesday]	20 100.9
Holiday[0]	20 051.3
Holiday[1]	19 250.1
SchoolHol[0]	20 010.2
SchoolHol[1]	20 022.5
Month[1] WeekDay[Friday]	19 512.4
Month[1] WeekDay[Monday]	18 596.6
Month[1] WeekDay[Saturday]	19 899.8
Month[1] WeekDay[Sunday]	19 373.8
Month[1] WeekDay[Thursday]	19 550.
Month[1] WeekDay[Tuesday]	19 479.8
Month[1] WeekDay[Wednesday]	19 511.1
Month[2] WeekDay[Friday]	19 774.
Month[2] WeekDay[Monday]	18 691.2
Month[2] WeekDay[Saturday]	20 418.1
Month[2] WeekDay[Sunday]	19 839.2
Month[2] WeekDay[Thursday]	19 809.9
Month[2] WeekDay[Tuesday]	19 568.6
Month[2] WeekDay[Wednesday]	19 739.3
Month[3] WeekDay[Friday]	19 640.1
Month[3] WeekDay[Monday]	18 947.
Month[3] WeekDay[Saturday]	20 187.
Month[3] WeekDay[Sunday]	19 701.7
Month[3] WeekDay[Thursday]	19 673.1
Month[3] WeekDay[Tuesday]	19 764.
Month[3] WeekDay[Wednesday]	19 793.
Month[4] WeekDay[Friday]	19 841.1
Month[4] WeekDay[Monday]	18 813.8
Month[4] WeekDay[Saturday]	19 874.8
Month[4] WeekDay[Sunday]	19 420.4
Month[4] WeekDay[Thursday]	19 813.1
Month[4] WeekDay[Tuesday]	19 571.
Month[4] WeekDay[Wednesday]	19 795.5
Month[5] WeekDay[Friday]	20 007.3
Month[5] WeekDay[Monday]	19 149.8
Month[5] WeekDay[Saturday]	20 675.5
Month[5] WeekDay[Sunday]	20 108.7
Month[5] WeekDay[Thursday]	20 114.6
Month[5] WeekDay[Tuesday]	19 973.4
Month[5] WeekDay[Wednesday]	20 028.8
Month[6] WeekDay[Friday]	20 855.3
Month[6] WeekDay[Monday]	19 838.4
Month[6] WeekDay[Saturday]	21 135.6
Month[6] WeekDay[Sunday]	20 483.1
Month[6] WeekDay[Thursday]	20 896.
Month[6] WeekDay[Tuesday]	20 792.6
Month[6] WeekDay[Wednesday]	20 906.8
Month[7] WeekDay[Friday]	21 246.3
Month[7] WeekDay[Monday]	19 989.5
Month[7] WeekDay[Saturday]	21 457.1
Month[7] WeekDay[Sunday]	20 749.2
Month[7] WeekDay[Thursday]	21 168.7
Month[7] WeekDay[Tuesday]	20 968.5
Month[7] WeekDay[Wednesday]	21 035.4
Month[8] WeekDay[Friday]	20 095.5
Month[8] WeekDay[Monday]	19 261.3
Month[8] WeekDay[Saturday]	20 823.4
Month[8] WeekDay[Sunday]	20 265.1
Month[8] WeekDay[Thursday]	20 128.

```

Month[8] WeekDay[Tuesday]          20 071.7
Month[8] WeekDay[Wednesday]        20 078.
Month[9] WeekDay[Friday]            20 036.2
Month[9] WeekDay[Monday]            19 129.1
Month[9] WeekDay[Saturday]          20 570.3
Month[9] WeekDay[Sunday]            20 052.8
Month[9] WeekDay[Thursday]          20 183.1
Month[9] WeekDay[Tuesday]           20 105.6
Month[9] WeekDay[Wednesday]         20 156.7
Month[10] WeekDay[Friday]            20 334.1
Month[10] WeekDay[Monday]            19 210.4
Month[10] WeekDay[Saturday]          20 859.1
Month[10] WeekDay[Sunday]            20 239.2
Month[10] WeekDay[Thursday]          20 339.9
Month[10] WeekDay[Tuesday]           20 143.4
Month[10] WeekDay[Wednesday]         20 120.5
Month[11] WeekDay[Friday]            20 230.4
Month[11] WeekDay[Monday]            19 295.1
Month[11] WeekDay[Saturday]          20 690.5
Month[11] WeekDay[Sunday]            20 269.1
Month[11] WeekDay[Thursday]          20 378.5
Month[11] WeekDay[Tuesday]           20 337.5
Month[11] WeekDay[Wednesday]         20 330.2
Month[12] WeekDay[Friday]            19 732.2
Month[12] WeekDay[Monday]            19 070.4
Month[12] WeekDay[Saturday]          19 825.6
Month[12] WeekDay[Sunday]            19 434.1
Month[12] WeekDay[Thursday]          19 841.4
Month[12] WeekDay[Tuesday]           19 639.1
Month[12] WeekDay[Wednesday]         19 714.
Holiday[0] Month[1]                  19 456.9
Holiday[0] Month[2]                  19 691.8
Holiday[0] Month[3]                  19 675.9
Holiday[0] Month[4]                  19 688.2
Holiday[0] Month[5]                  20 015.4
Holiday[0] Month[6]                  20 691.3
Holiday[0] Month[7]                  20 940.7
Holiday[0] Month[8]                  20 081.6
Holiday[0] Month[9]                  20 026.9
Holiday[0] Month[10]                 20 172.5
Holiday[0] Month[11]                 20 222.
Holiday[0] Month[12]                 19 845.8
Holiday[1] Month[1]                  18 335.9
Holiday[1] Month[3]                  19 617.1
Holiday[1] Month[4]                  18 979.5
Holiday[1] Month[5]                  19 821.8
Holiday[1] Month[6]                  20 953.3
Holiday[1] Month[8]                  20 770.1
Holiday[1] Month[9]                  20 210.8
Holiday[1] Month[12]                 18 441.6
Month[1] SchoolHol[0]                19 703.5
Month[1] SchoolHol[1]                19 215.8
Month[2] SchoolHol[0]                19 691.8
Month[3] SchoolHol[0]                19 610.5
Month[3] SchoolHol[1]                20 332.9
Month[4] SchoolHol[0]                19 614.7
Month[4] SchoolHol[1]                19 536.7
Month[5] SchoolHol[0]                20 019.4
Month[5] SchoolHol[1]                16 142.8

```

Month[6] SchoolHol[0]	20 440.2
Month[6] SchoolHol[1]	21 492.2
Month[7] SchoolHol[0]	20 793.1
Month[7] SchoolHol[1]	21 061.3
Month[8] SchoolHol[0]	20 106.
Month[9] SchoolHol[0]	19 900.3
Month[9] SchoolHol[1]	20 605.5
Month[10] SchoolHol[0]	20 191.4
Month[10] SchoolHol[1]	20 018.6
Month[11] SchoolHol[0]	20 192.5
Month[11] SchoolHol[1]	20 824.5
Month[12] SchoolHol[0]	20 480.9
Month[12] SchoolHol[1]	19 457.1
Holiday[0] WeekDay[Friday]	20 141.6
Holiday[0] WeekDay[Monday]	19 185.
Holiday[0] WeekDay[Saturday]	20 597.1
Holiday[0] WeekDay[Sunday]	20 049.2
Holiday[0] WeekDay[Thursday]	20 172.1
Holiday[0] WeekDay[Tuesday]	20 068.8
Holiday[0] WeekDay[Wednesday]	20 116.8
Holiday[1] WeekDay[Friday]	19 474.
Holiday[1] WeekDay[Monday]	18 957.4
Holiday[1] WeekDay[Saturday]	19 299.9
Holiday[1] WeekDay[Sunday]	18 967.
Holiday[1] WeekDay[Thursday]	19 729.3
Holiday[1] WeekDay[Tuesday]	19 095.1
Holiday[1] WeekDay[Wednesday]	19 653.8
SchoolHol[0] WeekDay[Friday]	20 056.9
SchoolHol[0] WeekDay[Monday]	19 119.5
SchoolHol[0] WeekDay[Saturday]	20 656.2
SchoolHol[0] WeekDay[Sunday]	20 076.9
SchoolHol[0] WeekDay[Thursday]	20 108.
SchoolHol[0] WeekDay[Tuesday]	20 022.8
SchoolHol[0] WeekDay[Wednesday]	20 051.1
SchoolHol[1] WeekDay[Friday]	20 270.6
SchoolHol[1] WeekDay[Monday]	19 305.3
SchoolHol[1] WeekDay[Saturday]	20 217.
SchoolHol[1] WeekDay[Sunday]	19 786.8
SchoolHol[1] WeekDay[Thursday]	20 325.
SchoolHol[1] WeekDay[Tuesday]	20 075.6
SchoolHol[1] WeekDay[Wednesday]	20 270.8
Holiday[0] SchoolHol[0]	20 002.5
Holiday[0] SchoolHol[1]	20 212.4
Holiday[1] SchoolHol[0]	20 357.9
Holiday[1] SchoolHol[1]	18 682.4

```
Out[255]= PostTests → {Month → Bonferroni {{1, 5}, {1, 6}, {2, 6}, {3, 6}, {4, 6}, {5, 6}, {1, 7}
      Holiday → Bonferroni {0, 1}, SchoolHol → Bonferroni {}}
```

```
In[256]:= meansA = anA[[-2, 2, 1]];
```

```
In[257]:= effectsA = meansA[[All, 2]] - meansA[[1, 2]]
```

```
Out[257]= {0., -595.805, -321.488, -341.371, -423.861, -5.29697, 688.282, 927.426,
  92.6867, 20.9884, 159.169, 208.682, -407.547, 95.5635, -846.004, 518.019,
  -20.9755, 144.318, 21.9616, 87.5527, 37.9644, -763.23, -3.12381, 9.24967,
  -500.911, -1416.69, -113.465, -639.543, -463.306, -533.462, -502.216,
  -239.321, -1322.07, 404.801, -174.112, -203.414, -444.684, -274.028,
  -373.16, -1066.31, 173.75, -311.576, -340.234, -249.294, -220.277,
  -172.223, -1199.52, -138.473, -592.899, -200.236, -442.249, -217.83,
  -6.0061, -863.536, 662.174, 95.4141, 101.352, -39.8947, 15.4579,
  842.023, -174.904, 1122.31, 469.755, 882.719, 779.34, 893.455, 1232.96,
  -23.7494, 1443.79, 735.894, 1155.41, 955.213, 1022.07, 82.225, -751.99,
  810.113, 251.842, 114.675, 58.4011, 64.7177, 22.9282, -884.2, 557.017,
  39.5204, 169.84, 92.2554, 143.4, 320.827, -802.858, 845.757, 225.868,
  326.575, 130.093, 107.229, 217.123, -718.202, 677.192, 255.77, 365.234,
  324.246, 316.906, -281.105, -942.925, -187.738, -579.211, -171.929,
  -374.174, -299.317, -556.355, -321.488, -337.406, -325.07, 2.08307,
  677.958, 927.426, 68.2554, 13.632, 159.169, 208.682, -167.464, -1677.4,
  -396.188, -1033.79, -191.501, 940.01, 756.775, 197.543, -1571.72,
  -309.75, -797.474, -321.488, -402.806, 319.584, -398.554, -476.604,
  6.07137, -3870.53, 426.877, 1478.87, 779.842, 1048.03, 92.6867, -113.01,
  592.243, 178.129, 5.25582, 179.19, 811.162, 467.61, -556.158, 128.278,
  -828.274, 583.822, 35.8759, 158.781, 55.4753, 103.49, -539.344, -1055.88,
  -713.434, -1046.26, -283.959, -918.188, -359.544, 43.6054, -893.788,
  642.938, 63.6332, 94.6936, 9.52469, 37.8122, 257.254, -708.038, 203.732,
  -226.454, 311.701, 62.3337, 257.466, -10.7546, 199.125, 344.569, -1330.86}
```

```
In[258]:= effA = Transpose[{meansA[[All, 1]], effectsA}]
```

```
Out[258]= {{All, 0.}, {Month[1], -595.805}, {Month[2], -321.488}, {Month[3], -341.371},
  {Month[4], -423.861}, {Month[5], -5.29697}, {Month[6], 688.282},
  {Month[7], 927.426}, {Month[8], 92.6867}, {Month[9], 20.9884},
  {Month[10], 159.169}, {Month[11], 208.682}, {Month[12], -407.547},
  {WeekDay[Friday], 95.5635}, {WeekDay[Monday], -846.004},
  {WeekDay[Saturday], 518.019}, {WeekDay[Sunday], -20.9755},
  {WeekDay[Thursday], 144.318}, {WeekDay[Tuesday], 21.9616},
  {WeekDay[Wednesday], 87.5527}, {Holiday[0], 37.9644},
  {Holiday[1], -763.23}, {SchoolHol[0], -3.12381}, {SchoolHol[1], 9.24967},
  {Month[1] WeekDay[Friday], -500.911}, {Month[1] WeekDay[Monday], -1416.69},
  {Month[1] WeekDay[Saturday], -113.465}, {Month[1] WeekDay[Sunday], -639.543},
  {Month[1] WeekDay[Thursday], -463.306}, {Month[1] WeekDay[Tuesday], -533.462},
  {Month[1] WeekDay[Wednesday], -502.216},
  {Month[2] WeekDay[Friday], -239.321}, {Month[2] WeekDay[Monday], -1322.07},
  {Month[2] WeekDay[Saturday], 404.801}, {Month[2] WeekDay[Sunday], -174.112},
  {Month[2] WeekDay[Thursday], -203.414}, {Month[2] WeekDay[Tuesday], -444.684},
  {Month[2] WeekDay[Wednesday], -274.028},
  {Month[3] WeekDay[Friday], -373.16}, {Month[3] WeekDay[Monday], -1066.31},
  {Month[3] WeekDay[Saturday], 173.75}, {Month[3] WeekDay[Sunday], -311.576},
  {Month[3] WeekDay[Thursday], -340.234}, {Month[3] WeekDay[Tuesday], -249.294},
  {Month[3] WeekDay[Wednesday], -220.277},
  {Month[4] WeekDay[Friday], -172.223}, {Month[4] WeekDay[Monday], -1199.52},
  {Month[4] WeekDay[Saturday], -138.473}, {Month[4] WeekDay[Sunday], -592.899},
  {Month[4] WeekDay[Thursday], -200.236}, {Month[4] WeekDay[Tuesday], -442.249},
  {Month[4] WeekDay[Wednesday], -217.83}, {Month[5] WeekDay[Friday], -6.0061},
  {Month[5] WeekDay[Monday], -863.536}, {Month[5] WeekDay[Saturday], 662.174},
  {Month[5] WeekDay[Sunday], 95.4141}, {Month[5] WeekDay[Thursday], 101.352},
```

```

{Month[5] WeekDay[Tuesday], -39.8947}, {Month[5] WeekDay[Wednesday], 15.4579},
{Month[6] WeekDay[Friday], 842.023}, {Month[6] WeekDay[Monday], -174.904},
{Month[6] WeekDay[Saturday], 1122.31}, {Month[6] WeekDay[Sunday], 469.755},
{Month[6] WeekDay[Thursday], 882.719}, {Month[6] WeekDay[Tuesday], 779.34},
{Month[6] WeekDay[Wednesday], 893.455}, {Month[7] WeekDay[Friday], 1232.96},
{Month[7] WeekDay[Monday], -23.7494}, {Month[7] WeekDay[Saturday], 1443.79},
{Month[7] WeekDay[Sunday], 735.894}, {Month[7] WeekDay[Thursday], 1155.41},
{Month[7] WeekDay[Tuesday], 955.213}, {Month[7] WeekDay[Wednesday], 1022.07},
{Month[8] WeekDay[Friday], 82.225}, {Month[8] WeekDay[Monday], -751.99},
{Month[8] WeekDay[Saturday], 810.113}, {Month[8] WeekDay[Sunday], 251.842},
{Month[8] WeekDay[Thursday], 114.675}, {Month[8] WeekDay[Tuesday], 58.4011},
{Month[8] WeekDay[Wednesday], 64.7177}, {Month[9] WeekDay[Friday], 22.9282},
{Month[9] WeekDay[Monday], -884.2}, {Month[9] WeekDay[Saturday], 557.017},
{Month[9] WeekDay[Sunday], 39.5204}, {Month[9] WeekDay[Thursday], 169.84},
{Month[9] WeekDay[Tuesday], 92.2554}, {Month[9] WeekDay[Wednesday], 143.4},
{Month[10] WeekDay[Friday], 320.827}, {Month[10] WeekDay[Monday], -802.858},
{Month[10] WeekDay[Saturday], 845.757}, {Month[10] WeekDay[Sunday], 225.868},
{Month[10] WeekDay[Thursday], 326.575}, {Month[10] WeekDay[Tuesday], 130.093},
{Month[10] WeekDay[Wednesday], 107.229}, {Month[11] WeekDay[Friday], 217.123},
{Month[11] WeekDay[Monday], -718.202}, {Month[11] WeekDay[Saturday], 677.192},
{Month[11] WeekDay[Sunday], 255.77}, {Month[11] WeekDay[Thursday], 365.234},
{Month[11] WeekDay[Tuesday], 324.246}, {Month[11] WeekDay[Wednesday], 316.906},
{Month[12] WeekDay[Friday], -281.105}, {Month[12] WeekDay[Monday], -942.925},
{Month[12] WeekDay[Saturday], -187.738},
{Month[12] WeekDay[Sunday], -579.211},
{Month[12] WeekDay[Thursday], -171.929},
{Month[12] WeekDay[Tuesday], -374.174},
{Month[12] WeekDay[Wednesday], -299.317},
{Holiday[0] Month[1], -556.355}, {Holiday[0] Month[2], -321.488},
{Holiday[0] Month[3], -337.406}, {Holiday[0] Month[4], -325.07},
{Holiday[0] Month[5], 2.08307}, {Holiday[0] Month[6], 677.958},
{Holiday[0] Month[7], 927.426}, {Holiday[0] Month[8], 68.2554},
{Holiday[0] Month[9], 13.632}, {Holiday[0] Month[10], 159.169},
{Holiday[0] Month[11], 208.682}, {Holiday[0] Month[12], -167.464},
{Holiday[1] Month[1], -1677.4}, {Holiday[1] Month[3], -396.188},
{Holiday[1] Month[4], -1033.79}, {Holiday[1] Month[5], -191.501},
{Holiday[1] Month[6], 940.01}, {Holiday[1] Month[8], 756.775},
{Holiday[1] Month[9], 197.543}, {Holiday[1] Month[12], -1571.72},
{Month[1] SchoolHol[0], -309.75}, {Month[1] SchoolHol[1], -797.474},
{Month[2] SchoolHol[0], -321.488}, {Month[3] SchoolHol[0], -402.806},
{Month[3] SchoolHol[1], 319.584}, {Month[4] SchoolHol[0], -398.554},
{Month[4] SchoolHol[1], -476.604}, {Month[5] SchoolHol[0], 6.07137},
{Month[5] SchoolHol[1], -3870.53}, {Month[6] SchoolHol[0], 426.877},
{Month[6] SchoolHol[1], 1478.87}, {Month[7] SchoolHol[0], 779.842},
{Month[7] SchoolHol[1], 1048.03}, {Month[8] SchoolHol[0], 92.6867},
{Month[9] SchoolHol[0], -113.01}, {Month[9] SchoolHol[1], 592.243},
{Month[10] SchoolHol[0], 178.129}, {Month[10] SchoolHol[1], 5.25582},
{Month[11] SchoolHol[0], 179.19}, {Month[11] SchoolHol[1], 811.162},
{Month[12] SchoolHol[0], 467.61}, {Month[12] SchoolHol[1], -556.158},
{Holiday[0] WeekDay[Friday], 128.278}, {Holiday[0] WeekDay[Monday], -828.274},
{Holiday[0] WeekDay[Saturday], 583.822},
{Holiday[0] WeekDay[Sunday], 35.8759}, {Holiday[0] WeekDay[Thursday], 158.781},
{Holiday[0] WeekDay[Tuesday], 55.4753},
{Holiday[0] WeekDay[Wednesday], 103.49},

```

```

{Holiday[1] WeekDay[Friday], -539.344},
{Holiday[1] WeekDay[Monday], -1055.88},
{Holiday[1] WeekDay[Saturday], -713.434},
{Holiday[1] WeekDay[Sunday], -1046.26},
{Holiday[1] WeekDay[Thursday], -283.959},
{Holiday[1] WeekDay[Tuesday], -918.188},
{Holiday[1] WeekDay[Wednesday], -359.544},
{SchoolHol[0] WeekDay[Friday], 43.6054},
{SchoolHol[0] WeekDay[Monday], -893.788},
{SchoolHol[0] WeekDay[Saturday], 642.938},
{SchoolHol[0] WeekDay[Sunday], 63.6332},
{SchoolHol[0] WeekDay[Thursday], 94.6936},
{SchoolHol[0] WeekDay[Tuesday], 9.52469},
{SchoolHol[0] WeekDay[Wednesday], 37.8122},
{SchoolHol[1] WeekDay[Friday], 257.254},
{SchoolHol[1] WeekDay[Monday], -708.038},
{SchoolHol[1] WeekDay[Saturday], 203.732},
{SchoolHol[1] WeekDay[Sunday], -226.454},
{SchoolHol[1] WeekDay[Thursday], 311.701},
{SchoolHol[1] WeekDay[Tuesday], 62.3337},
{SchoolHol[1] WeekDay[Wednesday], 257.466},
{Holiday[0] SchoolHol[0], -10.7546}, {Holiday[0] SchoolHol[1], 199.125},
{Holiday[1] SchoolHol[0], 344.569}, {Holiday[1] SchoolHol[1], -1330.86}

```

```
In[259]:= TableForm[effA];
```

```
In[260]:= argA = Table[ToString[effA[[i, 1]]], {i, 1, Length[effA]}];
```

```
In[261]:= For[i = 1, i ≤ Length[argA], i++,
  f[argA[[i]]] = effA[[i, 2]]]
```

```
In[262]:= f["Month[1]"]
```

```
Out[262]= -595.805
```

```
In[263]:= dA = Table[0, {i, 1, Length[dataA]}];
```

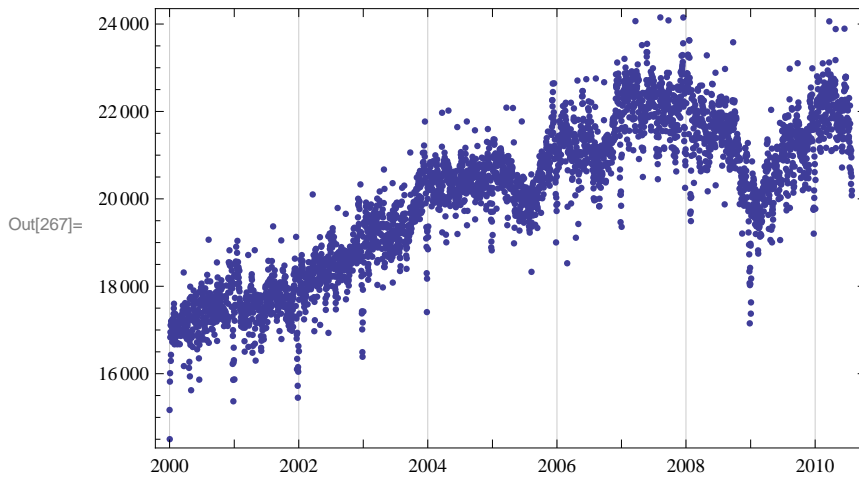
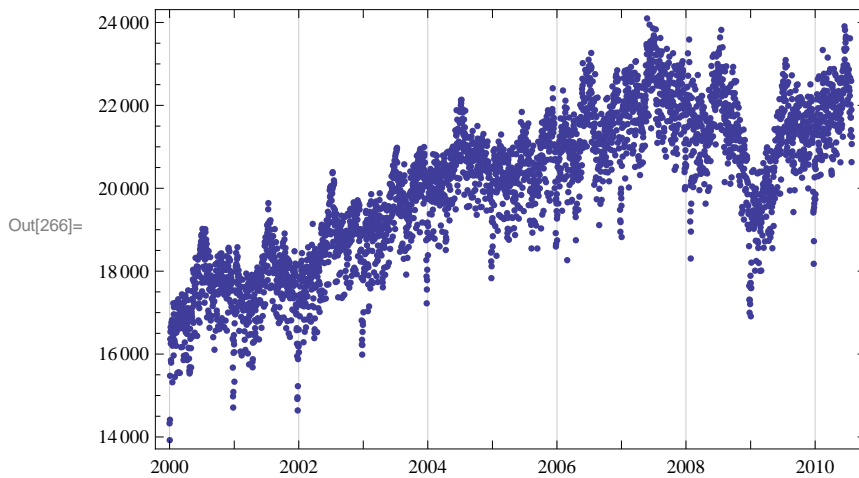
```
For[i = 1, i ≤ Length[dataA], i++,
  varA0[[i, 3]] = varA[[i, 2]] - f["Month[" <> ToString[dataA[[i, 1]]] <> "]" ] -
    f["WeekDay[" <> ToString[dataA[[i, 2]]] <> "]" ] -
    f["Holiday[" <> ToString[dataA[[i, 3]]] <> "]" ]]
```



```

In[265]:= varAmeff = Table[{varA0[[i, 1]], varA0[[i, 3]]}, {i, 1, Length[varA0]};
DateListPlot[varA]
DateListPlot[varAmeff]

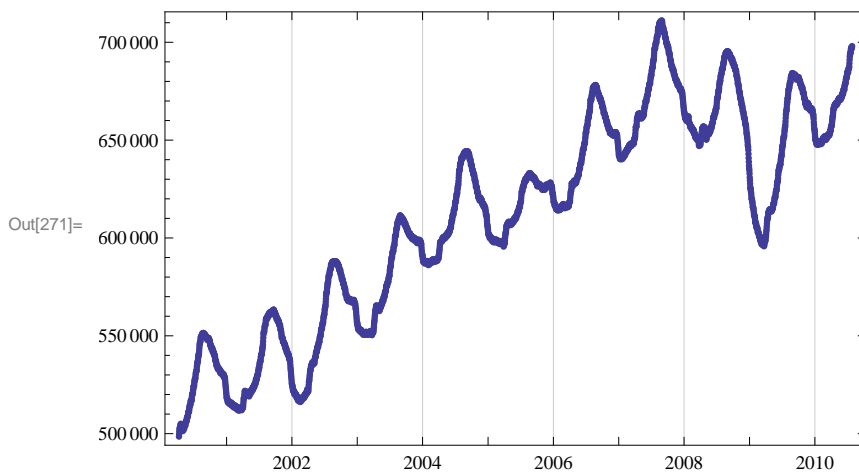
```



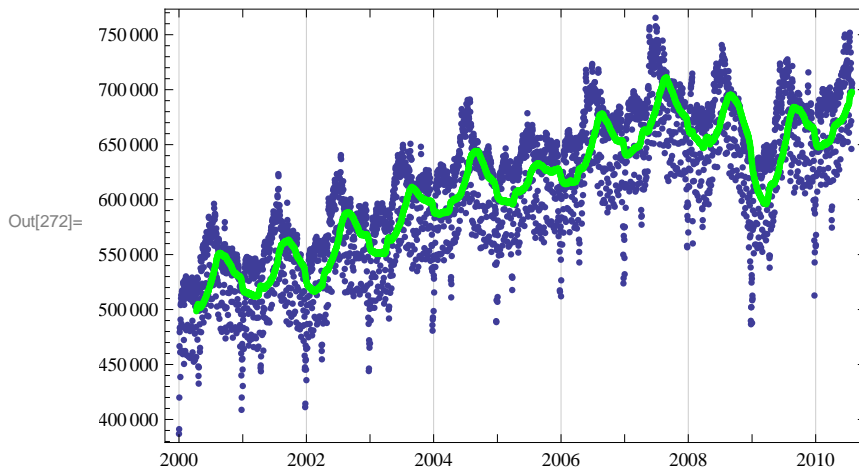
```

In[268]:= m1MA = MovingAverage[dailydata[[All, 2]], 100];
In[269]:= m1MA2 = Table[varA[[i, 1]], {i, 101, Length[varA]}];
m1MA3 = Table[{m1MA2[[i]], m1MA[[i]]}, {i, 1, Length[m1MA2]}];
DateListPlot[m1MA3]

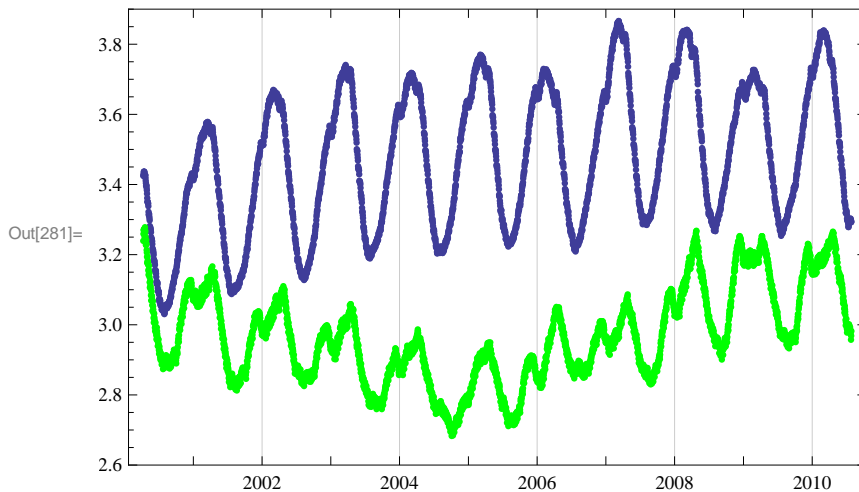
```



```
In[272]:= Show[DateListPlot[dailydata], DateListPlot[m1MA3, PlotStyle -> Green]]
```



```
In[273]:= MAS1 = MovingAverage[vars1[[All, 2]], 100];
MAS2 = MovingAverage[vars2[[All, 2]], 100];
MAS12 = Table[vars1[[i, 1]], {i, 101, Length[vars1]}];
MAS22 = Table[vars2[[i, 1]], {i, 101, Length[vars2]}];
MAS13 = Table[{MAS12[[i]], MAS1[[i]]}, {i, 1, Length[MAS12]}];
MAS23 = Table[{MAS22[[i]], MAS2[[i]]}, {i, 1, Length[MAS22]}];
DateListPlot[MAS13];
DateListPlot[MAS23];
Show[DateListPlot[MAS13, PlotRange -> {2.6, 3.9}],
DateListPlot[MAS23, PlotStyle -> Green]]
```



```
In[282]:= anM1 = ANOVA[dataM1, {Month, WeekDay, Holiday, SchoolHol, Month * WeekDay,
Month * Holiday, Month * SchoolHol, WeekDay * Holiday, WeekDay * SchoolHol,
Holiday * SchoolHol}, {Month, WeekDay, Holiday, SchoolHol}];
```

```
In[283]:= meansM1 = anM1[[-1, 2, 1]];
```

```
In[284]:= effectsM1 = meansM1[[All, 2]] - meansM1[[1, 2]];
```

```
In[285]:= effM1 = Transpose[{meansM1[[All, 1]], effectsM1}];
```

```
In[286]:= argM1 = Table[ToString[effM1[[i, 1]]], {i, 1, Length[effM1]}];
```

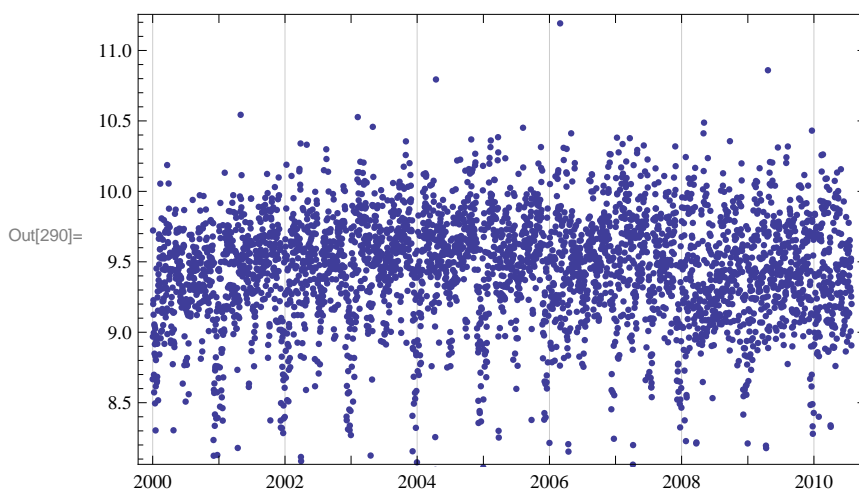
```

In[287]:= For[i = 1, i ≤ Length[argM1], i++,
  f[argM1[[i]]] = effM1[[i, 2]]]

In[288]:= For[i = 1, i ≤ Length[varM1], i++,
  varM10[[i, 3]] = varM1[[i, 2]] - f["Month[" <> ToString[dataM1[[i, 1]]] <> "]" ] -
  f["WeekDay[" <> ToString[dataM1[[i, 2]]] <> "]" ] -
  f["Holiday[" <> ToString[dataM1[[i, 3]]] <> "]" ] -
  f["SchoolHol[" <> ToString[dataM1[[i, 4]]] <> "]" ]]

In[289]:= varM1MinusEff = Table[{varM10[[i, 1]], varM10[[i, 3]]}, {i, 1, Length[varM10]};
DateListPlot[varM1MinusEff]
DateListPlot[varM1];
Variance[varM1MinusEff[[All, 2]]];
Variance[varM1[[All, 2]]];

```



```

In[294]:= anM2 = ANOVA[dataM2, {Month, WeekDay, Holiday, SchoolHol},
  {Month, WeekDay, Holiday, SchoolHol}];

In[295]:= meansM2 = anM2[[-1, 2, 1]];

In[296]:= effectsM2 = meansM2[[All, 2]] - meansM2[[1, 2]];

In[297]:= effM2 = Transpose[{meansM2[[All, 1]], effectsM2}];

In[298]:= argM2 = Table[ToString[effM2[[i, 1]]], {i, 1, Length[effM2]};

In[299]:= For[i = 1, i ≤ Length[argM2], i++,
  f[argM2[[i]]] = effM2[[i, 2]]]

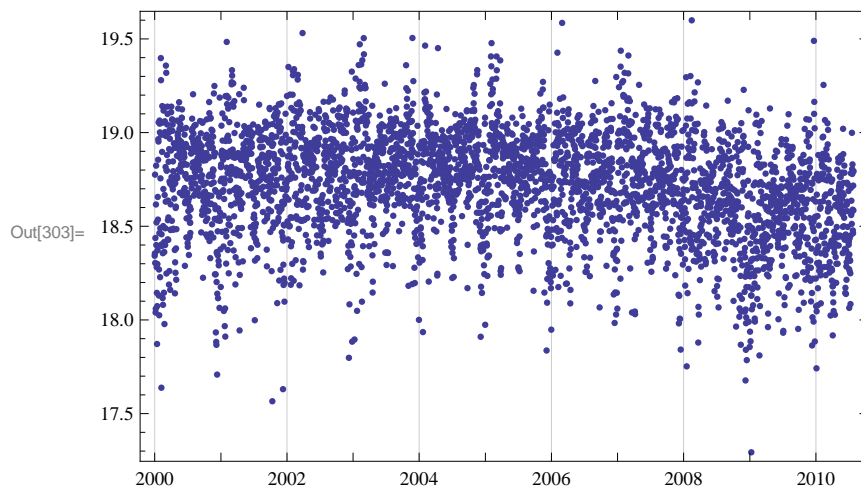
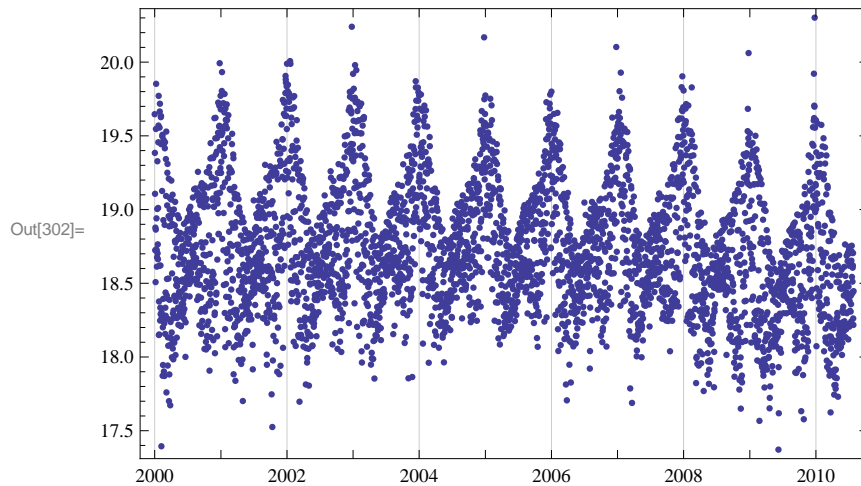
In[300]:= For[i = 1, i ≤ Length[dataM2], i++,
  varM20[[i, 3]] = varM2[[i, 2]] - f["Month[" <> ToString[dataM2[[i, 1]]] <> "]" ] -
  f["WeekDay[" <> ToString[dataM2[[i, 2]]] <> "]" ] -
  f["Holiday[" <> ToString[dataM2[[i, 3]]] <> "]" ] -
  f["SchoolHol[" <> ToString[dataM2[[i, 4]]] <> "]" ]]

```

```

In[301]:= varM2MinusEff = Table[{varM20[[i, 1]], varM20[[i, 3]]}, {i, 1, Length[varM20]}];
DateListPlot[varM2]
DateListPlot[varM2MinusEff]
DateListPlot[Take[varM2MinusEff, 365 * 4]];
DateListPlot[Take[varM2, 365 * 4]];

```



```

In[306]:= Variance[varM2MinusEff[[All, 2]]]
Variance[varM2[[All, 2]]]

```

Out[306]= 0.0772501

Out[307]= 0.194965

```

In[308]:= anS1 = ANOVA[datas1, {Month, WeekDay, Holiday, SchoolHol},
{Month, WeekDay, Holiday, SchoolHol}];

```

```

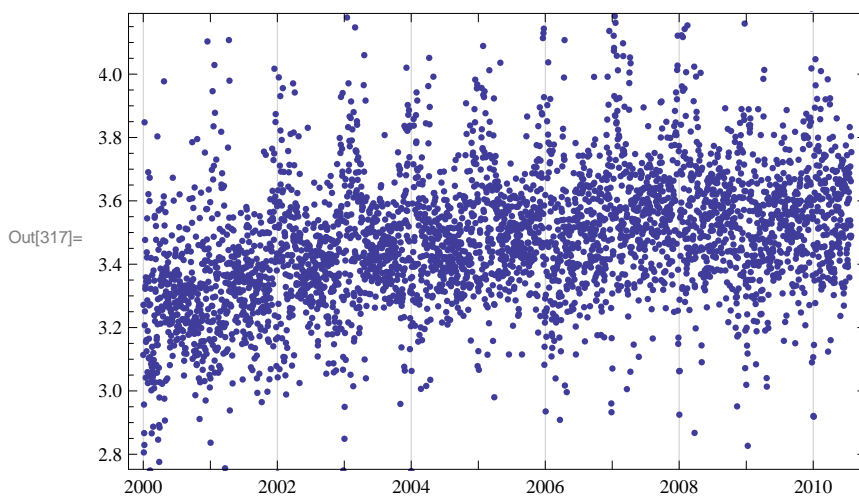
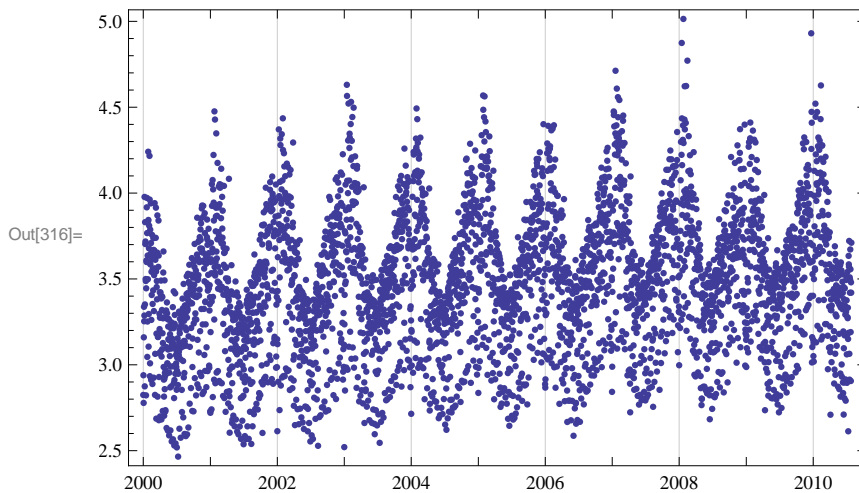
In[309]:= meansS1 = anS1[[-1, 2, 1]];
effectsS1 = meansS1[[All, 2]] - meansS1[[1, 2]];
effs1 = Transpose[{meansS1[[All, 1]], effectsS1}];
argS1 = Table[ToString[effs1[[i, 1]]], {i, 1, Length[effs1]}];

For[i = 1, i ≤ Length[argS1], i++,
  f[argS1[[i]]] = effs1[[i, 2]]]

For[i = 1, i ≤ Length[datas1], i++,
  varS10[[i, 3]] = datas1[[i, 5]] - f["Month[" <> ToString[datas1[[i, 1]] <> "]" ] -
    f["WeekDay[" <> ToString[datas1[[i, 2]] <> "]" ] -
    f["Holiday[" <> ToString[datas1[[i, 3]] <> "]" ] -
    f["SchoolHol[" <> ToString[datas1[[i, 4]] <> "]" ]]

varS1MinusEff = Table[{varS10[[i, 1]], varS10[[i, 3]]}, {i, 1, Length[varS10]}];
DateListPlot[varS1]
DateListPlot[varS1MinusEff]

```



```

In[318]:= anS2 = ANOVA[datas2, {Month, WeekDay, Holiday, SchoolHol},
  {Month, WeekDay, Holiday, SchoolHol}];

```

```

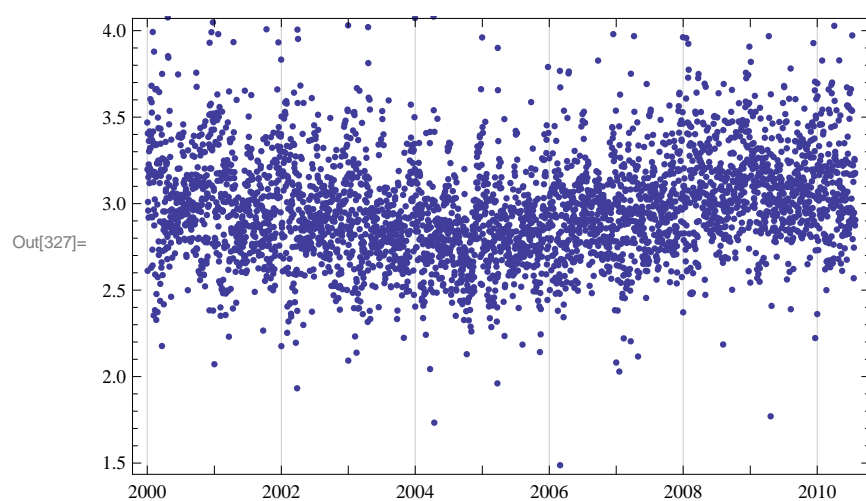
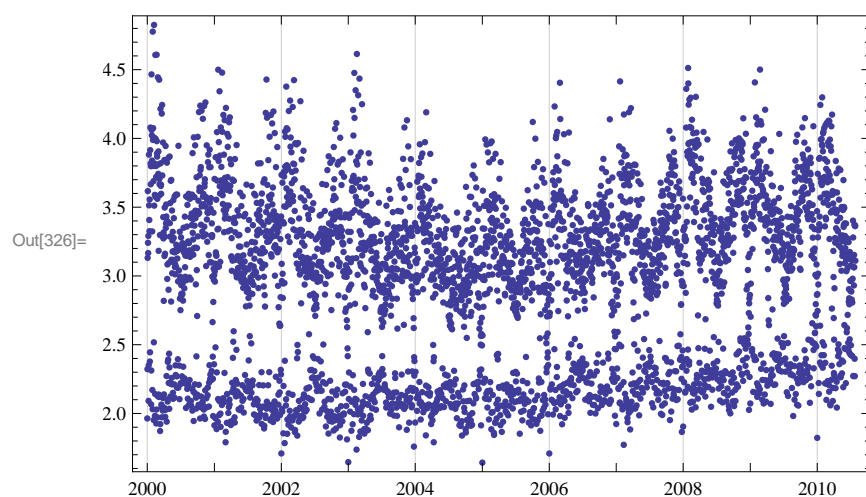
In[319]:= meansS2 = anS2[[-1, 2, 1]];
effectsS2 = meansS2[[All, 2]] - meansS2[[1, 2]];
effs2 = Transpose[{meansS2[[All, 1]], effectsS2}];
args2 = Table[ToString[effs2[[i, 1]]], {i, 1, Length[effs2]}];

For[i = 1, i ≤ Length[args2], i++,
  f[args2[[i]]] = effs2[[i, 2]]]

For[i = 1, i ≤ Length[datas2], i++,
  vars20[[i, 3]] = datas2[[i, 5]] - f["Month[" <> ToString[datas2[[i, 1]] <> "]" ] -
    f["WeekDay[" <> ToString[datas2[[i, 2]] <> "]" ] -
    f["Holiday[" <> ToString[datas2[[i, 3]] <> "]" ] -
    f["SchoolHol[" <> ToString[datas2[[i, 4]] <> "]" ]]]

vars2MinusEff = Table[{vars20[[i, 1]], vars20[[i, 3]]}, {i, 1, Length[vars20]}];
DateListPlot[vars2]
DateListPlot[vars2MinusEff]

```



```

In[328]:= anB1 = ANOVA[datab1, {Month, WeekDay, Holiday, SchoolHol},
  {Month, WeekDay, Holiday, SchoolHol}];

```

```

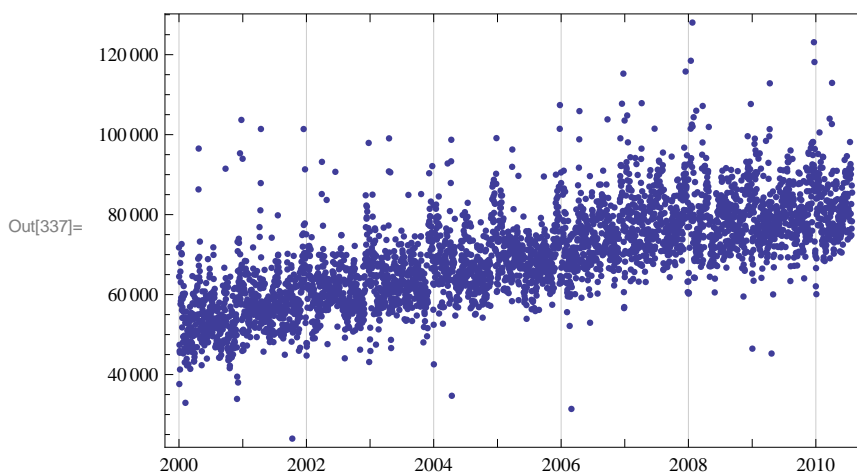
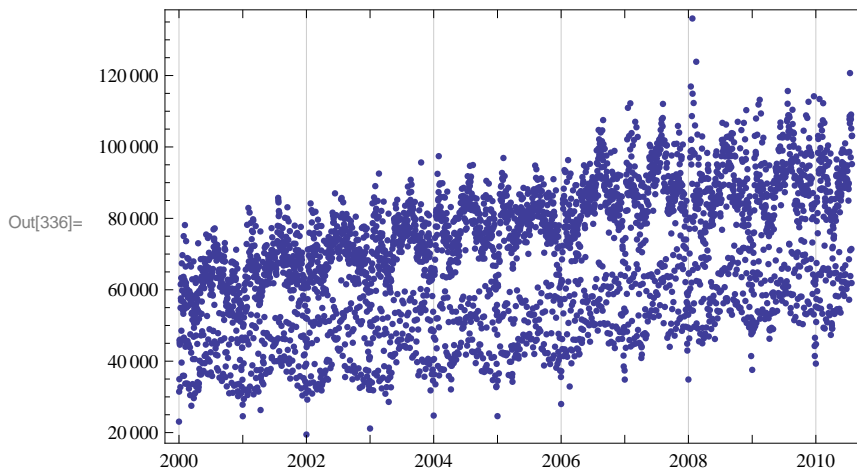
In[329]:= meansB1 = anB1[[-1, 2, 1]];
effectsB1 = meansB1[[All, 2]] - meansB1[[1, 2]];
effB1 = Transpose[{meansB1[[All, 1]], effectsB1}];
argB1 = Table[ToString[effB1[[i, 1]]], {i, 1, Length[effB1]}];

For[i = 1, i ≤ Length[argB1], i++,
  f[argB1[[i]]] = effB1[[i, 2]]]

For[i = 1, i ≤ Length[databl], i++,
  varB10[[i, 3]] = databl[[i, 5]] - f["Month[" <> ToString[databl[[i, 1]] <> "]" ] -
    f["WeekDay[" <> ToString[databl[[i, 2]] <> "]" ] -
    f["Holiday[" <> ToString[databl[[i, 3]] <> "]" ] -
    f["SchoolHol[" <> ToString[databl[[i, 4]] <> "]" ]]

varB1MinusEff = Table[{varB10[[i, 1]], varB10[[i, 3]]}, {i, 1, Length[varB10]}];
DateListPlot[varB1]
DateListPlot[varB1MinusEff]

```



```

In[338]:= anB2 = ANOVA[databl, {Month, WeekDay, Holiday, SchoolHol},
  {Month, WeekDay, Holiday, SchoolHol}];

```

```

In[339]:= meansB2 = anB2[[-1, 2, 1]];
effectsB2 = meansB2[[All, 2]] - meansB2[[1, 2]];
effB2 = Transpose[{meansB2[[All, 1]], effectsB2}];
argB2 = Table[ToString[effB2[[i, 1]]], {i, 1, Length[effB2]}];

For[i = 1, i ≤ Length[argB2], i++,
  f[argB2[[i]]] = effB2[[i, 2]]]

For[i = 1, i ≤ Length[datatab2], i++,
  varB20[[i, 3]] = datatab2[[i, 5]] - f["Month[" <> ToString[datatab2[[i, 1]] <> "]" ] -
    f["WeekDay[" <> ToString[datatab2[[i, 2]] <> "]" ] -
    f["Holiday[" <> ToString[datatab2[[i, 3]] <> "]" ] -
    f["SchoolHol[" <> ToString[datatab2[[i, 4]] <> "]" ]]

varB2MinusEff = Table[{varB20[[i, 1]], varB20[[i, 3]]}, {i, 1, Length[varB20]}];
DateListPlot[varB2]
DateListPlot[varB2MinusEff]

```

