An Urban Energy Baseline Model for Measurement & Verification of Building Energy Efficiency Retrofits in Abu-Dhabi

by

Luiz Augusto Friedrich

A Thesis Presented to the

Masdar Institute of Science and Technology

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in

Engineering Systems and Management

Abstract

Increasing awareness of climate change, pollution, reduced infrastructure investment availability and escalating fossil fuel prices have set pressure over policy-makers and energy sector planners to better utilize the electricity infrastructure by reducing or shifting peak load and to conserve energy.

Demand Side Management (DSM) measures ranging from advanced building controls to indoor-climate control and building equipment/envelope enhancement are designed to address this problem. Accurate determination of the ex-post impact of such measures is a widely recognized barrier to the wider deployment of DSM initiatives. The task is complicated by the dynamic nature of the system, the coupled interaction of multiple sub-systems and the high correlation with weather and other perturbations.

A baseline model of electricity consumption was establish for the city of Abu Dhabi, UAE, which will enable the comparison between different DSM interventions focused on curbing cooling demand and the business-as-usual (baseline) case. In the procedure proposed, also referred to as inverse load modeling, a novel hourly regression model of the aggregate load was estimated based on historical data. Different day schedules and seasonalities were accounted for, as well as the marginal effect of weather parameters such as temperature, relative humidity, wind speed and solar irradiance.

The model produced accurate results; adjusted R-squared of 0.9938 for the

training data (year 2010). The load during the first half of 2011 was used as verification data-set resulting in a RMSE of 34.19MW, equivalent to 1.99% of the growth-trend corrected annual peak load, and a MAPE of 2.75%.

It was identified that the cooling load accounts for approximately XX% of the overall electricity usage, while during the peak summer hour this value reaches YY% of the total electricity load. These values are to be minimized with policies and building codes that target overall energy efficiency, in order to curb emissions and wasteful energy consumption.

This research was supported by the Government of Abu Dhabi to help fulfill the vision of the late President Sheikh Zayed Bin Sultan Al Nayhan for sustainable development and empowerment of the UAE and humankind.

Acknowledgments

Acknowledgments

Luiz Augusto Friedrich,

Masdar City, January 1, 2013.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

1.1 Background and Motivation

The UAE has seen significant population growth in the last decade [27]. Alongside the inherent benefits of population and workforce increase, a series of problems also arose. Every human activity has a biosphere demand, that can be compared to the Earth's regenerative capacity. The Ecological Footprint, the measurement of the ecosystem's demand, is the sum of all cropland, grazing land, forest, fishing, timber and fibers, as well as the forestry area required to absorb the waste from energy production and the overall infrastructure area, regardless of where they are located on the planet. According to the World Wildlife Fund's Living Planet Report 2008 [43], the UAE had the world's worst ecological footprint per person, experiencing only minor improvements by the publication of the 2012 report [44]. This report called for proper planning and policy-making in order to rein in the country's environmental impact, in an era where international pressure towards mitigation of pollutants and green house gases *(GHG)* emissions has increased.

International agreements, such as the Kyoto Protocol which proposed an emissions cap and trade scheme applicable to only a group of countries (those called Annex I countries), has proven to be insufficient as overall world emissions have been increasing and reaching year after year record high levels [14]. Other carbon taxes and localized emissions punishment measures also generate what is called *Carbon leakage*, "defined as the increase in CO_2 emissions outside the countries taking domestic mitigation action divided by the reduction in the emissions of these countries. It has been demonstrated that an increase in local fossil fuel prices resulting, for example, from mitigation policies may lead to the re-allocation of production to regions with less stringent mitigation rules (or with no rules at all), leading to higher emissions in those regions and therefore to carbon leakage. Furthermore, a decrease in global fossil fuel demand and resulting lower fossil fuel prices may lead to increased fossil fuel consumption in non-mitigating countries and therefore to carbon leakage as well" [10].

Motivated by climate change and recent environment protection conferences as the United Nations Conference on Sustainable Development RIO+20, held in Rio de Janeiro, Brazil, on June 2012 and the Doha Climate Change Conference, held in Doha, Qatar in November 2012, or binding agreements which tends to produce uneven economic onus, the market force caused by worldwide increasing fossil fuel prices has created a new appeal for energy efficiency. Energy efficiency is undoubtedly one of the most constructive and cost-effective ways for achieving GHGs emission reductions, since the opportunity cost of the saved fuel alone is often greater than the cost of implementing corresponding energy efficiency measures. Thus, it becomes important to study urban energy consumption and to apply this knowledge for planning and policy making to improve urban environment and at the same time decrease the energy consumption, pollutants and GHGs emissions in cities.

Broadly speaking, energy efficiency measures can be divided into two main categories: Supply side and Demand side. Supply side measures focus mainly on electricity production plants and all the systems directly connected to it. Generally the electricity generation process goes through a series of energy conversions, as the options for directly converting primary energy to electricity are limited. In a thermal power plant for example, the primary energy is often converted to high pressure steam, which then is converted into mechanical energy through a steam turbine, directly connected to generators where the electricity is produced. Improving overall conversion efficiencies and processes from the primary energy source to the electricity being transmitted to the customers can therefore produce considerable improvements. On the other hand, demand side measures are those applied on the end-use side of energy. This energy is most of the time electricity consumed by urban communities, as it is the most used and the most easily convertible into other forms of energy for a broad range of applications such as indoor air conditioning and hot water production as well as for powering lighting and home appliances/devices such as refrigerators and computers. Worldwide aggregate buildings energy consumption is estimated to reach 50%, while this value reaches 70% for the Gulf region [30]. The industrial sector, even though has great contribution towards overall energy consumption, is beyond the scope of this work.

Increasing fossil fuel prices and reduced infrastructure investment availability is prompting energy sector planners to reduce and optimize energy generation requirements and enhance the utilization of the existing electricity infrastructure. Energy generation, transmission and distribution infrastructures have to be sized to supply the annual peak load, even if that peak occurs only during a short period and is far higher than the average system utilization rate. At the urban scale, better infrastructure utilization can be achieved through Demand Response: peak

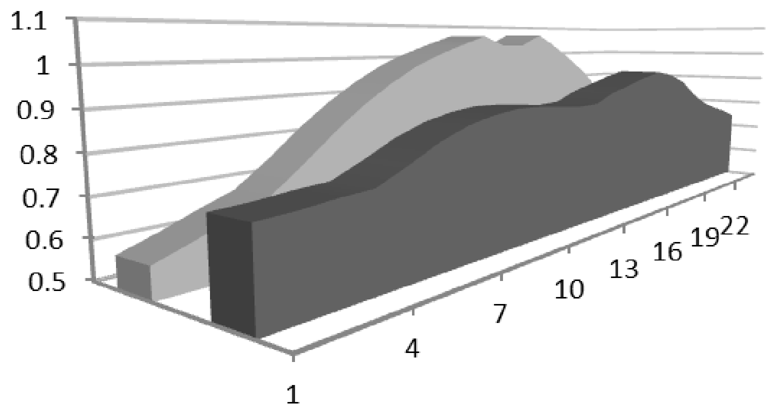

Figure 1.1: Peak Shaving/Load Shifting

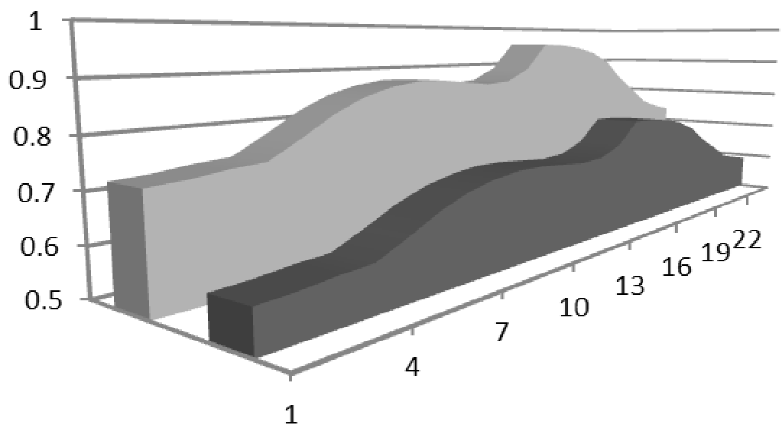

Figure 1.2: Energy Conservation

shaving–directly reducing the load during peak hours–and load shifting–shifting energy usage from high demand periods to lower demand (usually during late hours in the night until early morning)–as presented in Figure 1.1. Both actions, applied systematically, can result in reduced generation and transmission/distribution capacity requirements. Although Demand Response impacts emissions favorably [45], GHGs emission and fossil fuel use reduction is mainly achieved via energy conservation, where the overall load in the system is reduced as exemplified in Figure 1.2, also leading to a more efficient allocation of generation sources.

In some scenarios where price of fossil fuel sold to generators or retail energy prices (or both) are subsidized or when the investment in energy efficiency retrofits is paid for by a different entity than the one actually benefiting from the resulting savings (the "principal-agent" dilemma), a standard method for identifying and accounting the overall energy savings is necessary to provide a level field for the actors in the energy efficiency business. Estimating savings is a difficult task especially when monetary flux is also included, given the complexity and dynamics of the systems involved, the uncertain role of energy prices, lack of information on driving variables, unpredictability of end-user behavior and weather variability. The objective of this study is to establish a baseline model of air conditioning electricity consumption which will enable the measurement and verification of demand-side energy efficiency interventions by defining a business-as-usual as baseline. Another potential application of the baseline model is to identify the parameters having a physical interpretation and then simulate the effect of a planned DSM intervention by analyzing the sensitivity of the load to said parameters. In this way, simple ex-ante scenarios can be developed providing a quick and cost-effective approach to the a priori assessment of different DSM options.

1.2 Context, Problem Definition & Research Objectives

The UAE covers an area of 83,600 square kilometers, most of which is either sand deserts or salt flats. The country is a federation of seven "Emirates". Abu Dhabi, the largest of the seven Emirates constituting the UAE, lies in a T-shaped island in the Persian Gulf. The high temperatures and humid climate of the region during most of the year has a significant impact on the role of cooling load in the overall electricity consumption, identified to correspond to 40% of the total annual electrical load and 61% on the peak summer day [2]

Given average building insulation and extreme weather conditions, electricity load is highly correlated with weather, producing a high impact on energy consumption in the UAE and in most of the countries in the Gulf region, especially during the summer months. Demand side management of buildings' indoor air conditioning, can be approached from several angles: improvements of the cooling plants and chilled water or cool air distribution networks, building envelope insulation and solar gain reduction, day-light versus artificial light optimization and finally savings and peak load reductions via advanced controls.

To better understand the energy efficiency improvements, we proceed to divide cooling load into two types: Sensible cooling load and latent cooling load. The sensible cooling load refers to the load directly responsible for keep the building's indoor dry bulb temperature within a certain value range, while the latent cooling load strives to dehumidify the incoming fresh air so that indoor comfort conditions are not compromised (e.g., ASHRAE Standard 55).

Breaking down system-wide hourly electricity consumption in order to segregate base-load (constant and calendar-driven portions of the total load) from weather-driven latent and sensible cooling loads, can provide important information for DSM policy makers and implementers. Measurement and verification of DSM interventions plays an important role in defining savings and economic feasibility of energy efficiency projects, therefore a reliable and accurate model for this assessment has direct and tangible financial impacts.

1.3 Contributions

The main objectives of this study are to define a baseline mathematical model of the business-as-usual Abu Dhabi energy consumption which relates different components and variables affecting electricity consumption as well as daily load profiles.

The following objectives are proposed:

- To develop a low complexity linear model (some light non-linearity will be explored) using in-sample[1] data.
- To produce accurate prediction results tested against out-of-sample[2] data.
- To define a model with relatively low computational requirements for parameter estimation.
- Ensure that the estimated model coefficients accurately represent the underlying physical behavior.

The model provides a baseline energy consumption estimate for measurement and verification *M&V* of DSM projects, as well as the possibility of analyzing the contribution of each load constituent to the overall electricity consumption, in order to develop policies targeting specific consumption drivers.

1.4 Thesis Organization

The remainder of this thesis is organized as follows. Chapter 2 provides a literature review of load modeling and forecasting applications as well as different urban level energy modeling approaches and classifications. Chapter 3 discusses the data analysis, preparation and normalization procedures, the definition of an in-sample data set for model estimation and an out-of-sample data set for model validation. Chapter 4 provides a complete explanation of the model, variables being considered and reasons for such choices as well as a sensitivity analysis study. Chapter 5 describes the model validation procedure. Chapter 6 presents some applications

[1] In-sample data: portion of dataset used for the model fitting and parameters estimation.

[2] Out-of-sample data: part of the dataset not used during the parameter estimation phase, saved for testing the model's forecasting capabilities.

of the model. The final Chapter includes the conclusion and outlook for future research.

CHAPTER 2

Literature Review

Energy is a vital pre-requisite of any modern society. Analysis based on energy models are necessary for sustainable progress and development of nations. Since the introduction of electricity markets, load forecasting has been and still is today an active area of research. Operators continually run short-term forecasts (five minutes to one week ahead) to ensure system stability, plant scheduling, load dispatching, security assessment and reserve capacity. Mid-term forecasts (one week to a year ahead) are mainly used for generation optimization while long term forecasts (one year to ten years ahead) are required for investment planning. Inaccurate forecasts have high financial costs.

2.1 Model Classification

Different models are developed based on the type and amount of data available. The data used for energy demand modeling can be composed not only by the energy consumption history, but also by building blueprints, average building sizes and construction characteristics, periods of occupation, number of inhabitants, appliances, operation schedules as well regional macroeconomic and meteorological data, to name a few. There are two main classes of methods for predicting and analyzing aggregate urban energy usage [78], the bottom-up and the top-down approach. A comparison of both classes and an analysis of the possible approaches in each category is done in [8].

2.1.1 Bottom-up Model

The traditional bottom-up approach to model urban energy use is usually focused on data obtained from surveys and field measurements of the energy consumption profile and the characteristics of an individual unit or a statistically representative sample of units, identifying fixed demand per unit floor area or per household and extrapolating the results to infer global urban usage characteristics. This approach usually fails to model the total energy use in buildings and at urban level with sufficient resolution, due to a non-linear relationship between load and floor area. A more precise approach consists of a series of building models used together in a way to represent the overall building stock, individual simulation of each building focusing on energy use characteristics, are aggregated together based on the overall representativity of each building type in the studied area in order to predict urban energy use [40, 47, 24]. This approach can produce better results as compared to the traditional method, but its main weaknesses lie in the high dependency on the precision of the building prototype models, in the collected data of the buildings, in how the individual data is aggregated, and in the assumptions made with regard to changing demographic factors, hours of occupancy, indoor climate control system in use etc. [48]. The work described in [82] shows the application of an urban energy model based on individual representative buildings combined into “building clusters" for specific districts and extrapolated to determine potential CO_2 emission reduction for Osaka city, Japan, considering different energy efficiency and energy saving measures.

Bottom-up models work at a disaggregated level, and thus the detail, precision and size of the underlying building database directly influences the final result of the model [75]. Other phenomena such as the Urban Heat Island *(UHI)* effect cannot be properly accounted for and estimated, since in practice the overall urban energy consumption is different from the sum of all its constituting isolated buildings due to the urban micro climate which impacts cooling and heating energy requirements [51]. Due to the complex dynamics of the system, the non-linearity and coupling of sub-systems and the high correlation with weather and other random perturbations, the bottom up approach has many limitations when applied to urban/district energy models.

2.1.2 Top-down Model

As opposed to the bottom-up approach, the top-down approach consists in utilizing high level data, such as the approximate residential aggregate energy consumption–since on site energy production data is usually not available–macroeconomic data as gross domestic product (GDP), population growth, buildings construction/demolition rate and meteorological data to define overall urban level consumption characteristics and trends. For the top-down approach, contributions can be broadly classified in four main categories [17]:

- Time series models,
- Artificial neural networks (ANN),
- Similar day look-up,
- Regression-based approaches.

Time series models are best suited for capturing the time dependence of the load, seasonal effects (summer, winter), day-of-the-week (working day, non-working day), calendar holidays, different hours of the day and trends in general [3]. Some region specific studies using this method are relevant: energy consumption and economic growth for south-east-Asia was conducted by [5], energy production and consumption in northern Spain [20], energy demand on sectorial UK [41], X11 seasonally adjusted economic model [26], comparison of hourly load profiles for Brazil [76] and short-term day-ahead and week-ahead load forecasting model used by the Spanish system operator [18] to name a few.

Artificial Neural Networks *(ANN)* methods, often classified as "black-box"-type models, automatically interpolate among the electricity load and hour, day-of-week, day-of-year and temperature data in a training data set based on a predefined architecture; its main advantage is to learn complex and non-linear relationships that are difficult to model with conventional techniques [56, 23, 66, 54]. Due to the nature of the ANN methods, the coefficients (weights) cannot be matched with the underlying physical process, therefore the marginal impact of each input variable is not clear form the model coefficients.

The similar day look-up approach is based on searching historical data with similar characteristics (weather conditions, day-of-week, day-of-year) to the day of forecast. Linear combination of different similar days can be made to improve the forecast [9]. This model approach is purely geared toward forecasting and offers no correspondence with the underlying physical model, which does not allow the direct assessment of the sensitivity for each underlying load driver.

Regression models utilize the strong correlation of load with relevant factors such as weather, hour-of-the-day and day-of-the-week. Work using this method has proceeded on two fronts, single-equation models and multiple-equation models with different equations for different hours of the day [25, 69, 29]. The method used herein is composed of a single-equation model; therefore all data can be used for estimating one model, as opposed to breaking the dataset into smaller sets, usually done by having a separate equation for each hour. Long-term forecast based on linear and linear-log regression models was proposed in [80], mathematical modeling based on regression analysis was used to forecast load demand up to year 2000 for Yugoslavia [81].

Regression analysis was the method chosen in this study for combining and pondering the different effects, due to its relatively low computational cost and broad range of application. This approach also produces interpretable coefficients for the different contributors of the load, facilitating the analysis of each component individually, as opposed to other top-down approaches where the physical correspondence of the coefficients are lost. A comparison between six modeling techniques was done in [79] for short term urban level energy forecast, the regression model yielded the best results for up to one day ahead, being only outperformed by one of the methods (exponential smoothing filter with autoregressive component) for one to two days ahead forecast.

In [13], a linear regression model was used to estimate the elasticities of GDP, price and GDP per capita for the domestic and non-domestic Italy's energy consumption, and for long term forecasting and comparing it with available projections.

Previous studies have looked at the correlation between weather variables and electricity consumption [72, 4]. A model for forecasting electricity consumption for Spain [65] uses only temperature data while the possible influence of other weather variables were left for future work. In the model proposed herein, temperature is used together with relative humidity, wind and solar irradiance to account for different weather influences.

Some studies where the modeled areas have a heating season and a cooling season, the model can be broken down into analyzing each season individually. [32, 53] model only the load for the summer months, while [21] uses an unsupervised segmentation to break the time series in winter and summer, to be analyzed separately. One common method in practice is to transform temperature into degree-days or degree-hours based on determined set points [1, 36, 58].

The model here described does not include ARIMA components, as they are difficult to interpret from an economic/physical point of view, and can be unstable for longer term forecasts. Instead, a Fourier series is used to model daily and monthly seasonalities [67], as well as, dummy variables to differentiate working days, from Fridays and Saturdays and exogenous variables representing the weather parameters influence on the load. A similar approach was used by [28], where a stochastic trend component, together with fixed and time-varying regression effects were used to short term forecasting. For mid-term forecast, [17] uses Fourier series, as well as an elegant way to account for the non-linearity and progressive temperature threshold impact for when heating (or cooling) starts impacting the load, the model was integrated to the *Electricité de France* (EDF) load forecasting system.

2.2 Demand Side Management

The vast majority of literature in the field of electricity load forecasting are based on ARMA/ARMAX models as they yield good results for short term forecasts, a niche of the energy markets which has a high number of players and high financial cost associated with uncertainty and prediction errors. For this purpose, ARMA/ARMAX models produce a good estimate due to high autocorrelation of electricity load. For mid to long term modeling where the significance with underlying physical process is desired, linear regression is usually the most used modeling technique.

There are two categories of forecasts, *ex ante* meaning before the event, based only on information known before the forecast period, used for estimating future outcomes based only on past information. A second method called *ex post*, consists in using some information about the forecast period that is known, being used mainly for scenario study, as in the case here described as baseline definition for measurement DSM interventions and impact assessment.

Most M&V guidelines are divided into two general approaches: the one looking isolated into the affected equipment and systems without considering the whole-facility, and the one where the whole-facility is considered and individual equipment performances are aggregated into a broader bound. The main difference is how the data is collected, and the number of measuring/collecting instruments used. In the first approach, individual metering of the affected systems are necessary, while the latter aggregate facility data can be used. The metered data is used to infer savings, as the actual saved energy cannot be directly measured. For this assessment, all M&V options utilize models to predict a baseline energy used for comparison. The most used model for defining a baseline is based on regression analysis, where a series of independent and measurable parameters that affect the load are used to estimate the energy usage. Similarly to the electricity forecasting models, the regression model used for determining a baseline of electricity consumption involves independent variables as weather, calendar and occupancy to name a few. A series of standards and guidelines are available for M&V of energy savings for facilities, the most widely used being the International Performance Measurement and Verification Protocol (IPMVP) [63] and ASHRAE's Guideline 14 [7]. The assessment and estimation of urban-level energy efficiency measures and policy impacts is still an active of research due to the intrinsically of each system and region.

A study based on bottom-up methodology was conducted for Osaka City, Japan [74] to measure the potential energy savings from different DSM measures. The model included detailed information about appliances, lighting and heating/cooling devices as energy requirements, utilization and schedules. The study considered the simulation of 19 households and 20 building categories, using 5 different levels of insulation, energy efficiency standard for room air conditioners, daylight saving time and other energy conservation measures in order to reduce overall energy consumption of each building, and based on the representativity of each building category in the studied area, infer overall savings.

The non-linear relationship between electricity consumption and temperature was studied in [12] for 15 countries of the European Union, in order to identify the potential impact of global warming on electricity consumption, as well as the market penetration evolution of air-conditioning units, as the dependency of electricity consumption and temperature was identified to increase with time.

A bottom-up model is described in [16] where the urban energy consumption model is used to compare different DSM measures in order to determine the optimal share of each measure using a linear program to meet emission reduction targets in the most cost effective way.

In the work here described, a mid-term model of hourly electricity load for Abu Dhabi was developed using a technique also referred to as inverse model, to be distinguished from load forecasting, since it is performed off-line (ex post) on historical time-series data (typically one year or more), with the objective of establishing a baseline model of electricity consumption for measurement and verification of demand-side energy efficiency interventions, and enabling the model parameters to represent the sensitivity for some of the load-affecting components, in order to predict overall impact on load from some planned DSM interventions.

CHAPTER 3

Data Source and Preparation

The quality of the baseline electricity model is undoubtedly dependent on the quality of data used for calibrating and testing. This study relies on substation-level hourly electricity data measured by the SCADA system of Abu-Dhabi emirate's utility as well as robust hourly weather data (including solar irradiance) monitored by Masdar City's comprehensive weather station for the calendar year 2010 (January 1^{st} – December 31^{st}) and the first half of 2011 (January 1^{st} – June 30^{th}). A subset of substations which are all within the municipality of Abu-Dhabi was selected that in aggregate constitute a better proxy for the buildings load (residential, office/institutional and retail). Hourly electricity consumption data from 26 low-voltage substations (listed in appendix XXX) ranging from 11kV to 33kV was used. This subset includes all substations that directly supply street transformers attending final customers and that were fully operational since the beginning of 2010.

3.1 Outliers Detection and Treatment

An outlier is defined as an observation or a data point that lies an abnormal distance from neighbor values and does not match the general behavior of the data source. The identification of an outlier can be represented as the selection of samples that are considerably different or inconsistent with the neighbor data points and trend, this process is composed by two main tasks: defining what would be characterized as an exceptional data point, and second on how to identify such points in an efficient matter. Recently, proposed outlier detection algorithms can be classified as statistics-based, distance-based or density-based, to name a few [22, 39].

The statistics-based approach [46] can be considered the simplest outlier detection mechanism. The idea behind this method is based on the specific distribution probability of the underlying data, therefore data points outside that range can be considered outliers. The data points are classified based on a discordance test, given the distance from the distribution's mean. The hypothesis that a data point is part of an underlying distribution is tested with a confidence interval α, against the counter-hypothesis that the point is not part of the model's distribution. Statistics-based approach is simple and has linear calculation complexity in terms of the data size. However for complex and multi dimension datasets, the distribution of the data and its parameters requirements can be a difficulting factor for the application of this method.

Distance-based outlier detection, initially proposed by [49, 50] defines an outlier based on a given p and d, if a fraction p of the points lie at a distance greater than d from the point being analyzed, that characterizes an outlier.

The density-based outlier detection proposed by [15], is not a binary outlier/not-outlier indicator, but it identifies for each object a called local outlier factor (LOF), that is given by the distance of *k-nearest neighbors* based on an integer parameter k, the number of neighboring points to be computed.

From these three approaches, the outlier identification used in the work here described was the statistics-based method. In many cases this approach is not enough for a proper identification of the outliers, specially under multi dimensional data, which was not the case for the electricity load data. For the aggregate data, the summation of all substations in the subset mentioned, this approach identified six points as outliers, representing 0.068% of the data points, none being consecutive. Figure 3.1 presents a histogram of the aggregate electricity load data, and a clear visualization of the five outliers close to zero, and one close to 400. By going to the individual substations level, these points identified in the aggregate data were missing points in some of the substation datasets. When aggregated those substations generated a value out of the range of the other data points. Linear interpolation between neighboring points was used to correct for the outliers. Figure 3.2 presents the electricity load over time since the first hour of 2010 until the 24^{th} hour of June 30^{th}, 2011.

Dry bulb temperature *(Temp)*, relative humidity *(RH)*, wind speed *(Wind)*, global horizontal irradiance *(GHI)* and direct normal irradiance *(DNI)* data was obtained from Masdar City's weather station. The same outlier identification approach was used without identifying any points out of the series range.

3.2 Variables Cross-correlation

The Gauss-Markov theorem states that the least squares estimators have smallest variance of any unbiased estimators. This does not assure that the variance is always small in an absolute sense. If for example two variables in a model are perfectly correlated, its variance is infinite [35] (perfect multicollinearity problem). The more common case is the non-perfect multicollinearity, when two or more variables in a model are highly correlated. Under non-perfect multicollinear-

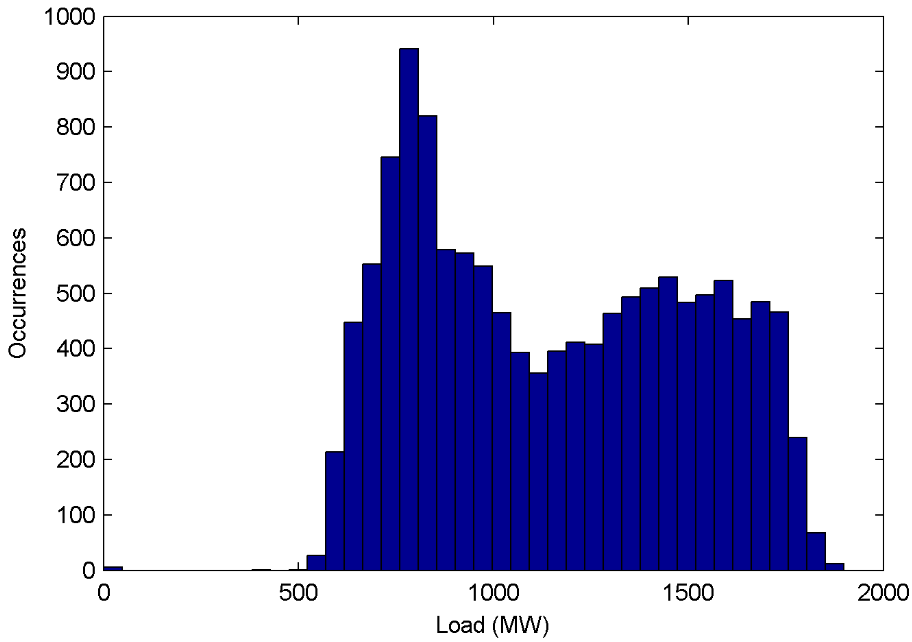

Figure 3.1: Electricity load data distribution histogram

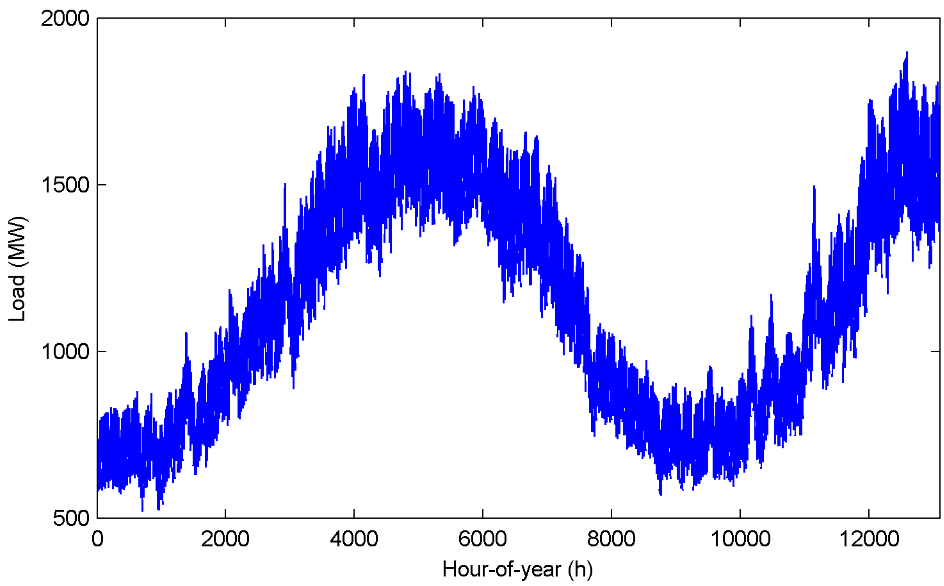

Figure 3.2: Electricity load data over time

ity problem, the least squared estimators maintain its properties but the larger the variance of the coefficients, the less efficient the estimation is and the higher the probability to get erratic estimates. The main problems that arises are:

- Adding or removing a variable from the model produces wide swings in the parameter estimates or the sign of the estimate.
- Affected coefficients tend to have large standard errors.
- Coefficients with "wrong" sign or implausible magnitudes based on the physical underlying process they are representing.

In order to avoid multicollinearity in the model, a correlation matrix between the relevant hourly weather variables (Table 3.1) was used to identify highly correlated variables. As expected, solar irradiances have positive correlation between the different measurements *(GHI, DNI, DHI, DNI_H, DNI_V)*, calculated as described in Section 4.3, in the order of 0.9. To avoid multicollinearity between the solar irradiance variables, DNI was determined by comparing a simple regression using each of the solar irradiance components, to be the most significant of all irradiance measurements, and was the only variable retained for representing solar irradiation.

Table 3.1: Correlation matrix

	GHI	DNI	DHI	DNI_H	DNI_V	Temp	RH	Wind	SH
GHI	1	0.91	0.88	0.71	0.95	0.54	-0.64	0.43	-0.13
DNI	0.91	1	0.66	0.91	0.95	0.4	-0.56	0.38	-0.17
DHI	0.88	0.66	1	0.5	0.7	0.53	-0.6	0.48	-0.09
DNI_H	0.71	0.91	0.5	1	0.74	0.23	-0.42	0.32	-0.17
DNI_V	0.95	0.95	0.7	0.74	1	0.49	-0.61	0.38	-0.15
Temp	0.54	0.4	0.53	0.23	0.49	1	-0.65	0.4	0.44
RH	-0.64	-0.56	-0.6	-0.42	-0.61	-0.65	1	-0.48	0.35
Wind	0.43	0.38	0.48	0.32	0.38	0.4	-0.48	1	-0.01
SH	-0.13	-0.17	-0.09	-0.17	-0.15	0.44	0.35	-0.01	1

Insert histogram and correlation plots

3.3 Data Normalization

In order to be able to easily identify and compare the marginal impact of each variable in the model, the data was normalized so that it ranges from zero to one (0 – 1). The weather and load data was normalized by subtracting from each measured parameter its lowest data point value, and dividing the remainder by the highest minus the lowest values for each parameter.

In areas where a heating and a cooling season is present, the dataset is usually divided for each season and two separate models are produced. Since the heating load is not apparent in the area here studied, a single model was developed with the whole dataset. The cut-off temperature, the point at which electricity starts being used for indoor space cooling, was considered using two different approaches to account for the non-linear behavior of sensible cooling requirement and load: The first was based on analyzing the root-mean-square error *(RMSE)* change-point considering the cut-off threshold temperature, and the second one is based on applying a "threshold function" to account for the non-linearity.

The RMSE change-point approach consists in varying the cut-off point of the temperature in the model until a point of minimum in identified. This method was applied by [2] to the same geographical area under study here for the year of 2008, and identified that 18.5 degrees Celsius was the optimal cut-off point. The same test was conducted for the year of 2010 (under study) and similar (18.47 °C) cut-off temperature value obtained using line search minimization, being 18.5 °C cut-off point used for this approach. Figure 3.3 represents the electricity load vs temperature, where data-points below the threshold had the normalized temperature set to zero, and a polynomial function later fit to model the non-linear behavior. Similarly specific humidity data was normalized using the cut-off value of 0.008 kg(water)/kg(dry air), the threshold beyond which dehumidification is assumed to occur.

The second approach was based on the assumption that a single cut-off value for an entire urban area cannot be determined, therefore a "threshold function" ψ similar to the one implemented in the Eventail load model of Electricité de France, described in [17] was used. This method consists in applying a transformation function to the sensible cooling proxy[1] and to the latent cooling proxy, in order to account for the non-linear relation to load. The threshold function $\psi(x, \chi_0, \sigma)$ of a random variable x is the expected value of $max(x-\chi, 0)$, where χ is normally distributed with mean χ_0 and standard deviation σ. The transformation based on ψ corresponds to a left-truncated (one-sided) normal distribution function as presented in Equation 3.1 [35], where Φ is the standard normal cumulative distribution function and φ is the standard normal probability density function. There is no closed form solution for ψ so it needs to be estimated numerically or from tables.

$$\bar{x} = \psi(x, \chi_o, \sigma) = \Phi\left[(x-\chi_0)/\sigma\right] \left\{ x - \chi_0 + \sigma \frac{\varphi\left[(x-\chi_0)/\sigma\right]}{\Phi\left[(x-\chi_0)/\sigma\right]} \right\} \tag{3.1}$$

Figure **??** presents the normalized values of electricity consumption for the year used in the model calibration. Figure represents the distribution of the weather variables and the electricity load after the normalization process.

Insert figure of this threshold function

[1] Proxy Variable: an indicator used to account for a variable in a model which simply has no fully observable/measurable counterpart [35].

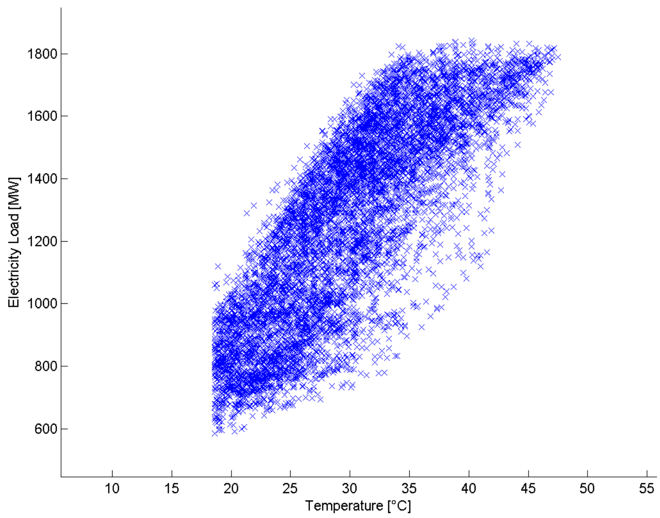

Figure 3.3: Electricity load vs Temperature

CHAPTER 4

Model Definition

In order to define a pre-DSM baseline of electricity consumption for pre/post analysis of DSM investments in energy conservation and peak load reduction, a model accounting for the different variables affecting the electricity consumption is necessary. System-wide electrical load can be divided into two main sources: the offset which is independent of climate and the cooling load assumed to be entirely climate-driven, since a heating season is not apparent in the studied area. The offset varies with hour-of-day, day-type and season, without having direct relation with instantaneous weather parameters, and cooling load which can be further divided into two sources: sensible cooling and latent cooling load. Sensible cooling refers to the load directly responsible to keep the building's dry bulb temperature within a certain value range, mainly depended on outdoor dry-bulb temperature, while the latent cooling load is related to the wet-bulb temperature, being affected by humidity. To identify the impact of each component, a set of calendar and weather derived parameters were considerer.

Abu Dhabi has experienced a growth trend in the electricity consumption which was modeled in two different ways: an additive linear and a multiplicative term. For the monthly seasonality, variations between months not related to weather parameters, four different methods were tested: dummy variables, fuzzy variable, Fourier series and moving average filter. The weather dependent fraction of the load has a significant contribution to the electricity load in Abu Dhabi, specially during the summer months were accounted. A combination solar irradiation, dry-bulb temperature and humidity indicators were tested based on literature information on the subject.

This Chapter describes each of the parameters and methods used in the model identification, being followed by Chapter 5, where the actual results are presented.

4.1 Trend

Abu Dhabi has seen significant population and energy consumption growth in the last decade. To account for this trend in electricity consumption through the year, and therefore the non-stationarity of the data series, two methods were tested: an additive linear trend and a multiplicative exponential trend.

Within literature, different approaches are used in order to produce a stationary series from the original data. Differencing is a method commonly used in ARMA type models in order to correct for non-stationarity of the data, since stationarity is a pre-requisite for the proper analysis of such series. Some authors [76] are critical of using differencing without the proper unit-root test, and due to other problems that arise with such method. When the trend is in fact deterministic, a non-invertible moving average component is introduced when differencing, which causes estimation problems in the model.

By differencing the coefficients of the model are relative to the rate of change (differenced) data, therefore an interpretation of such coefficients becomes non-trivial. Given the objectives of the model here described, to have interpretable coefficients in order to analyze the impact of each load contributing component, differencing was not tested as a method to account for the trend in the data series.

Consider $E(t)$ the total electricity consumption on hour t, $X(t)$ the vector with the exogenous variables affecting the load other than the growth on hour t, and β the coefficients of the model, for a time-series of length T.

4.1.1 Linear

The linear trend is based on the assumption that the rate of immigration, new buildings and new appliances result in a constant hour-over-hour increase in the electricity consumption throughout the year. This method considers that at every hour a portion of the load (Gl) is increased from the previous hour, independently of other factors, Gl being constant through the year. A linear growth trend is defined by Equation 4.1:

$$E(t) = Gl \times t + X(t) \times \beta \qquad \forall t \in T \tag{4.1}$$

The growth term Gl is identified during the same routine as the other parameters in the model as an additive term after the data normalization process.

4.1.2 Multiplicative

The multiplicative trend is based on the assumption that there is a rate of immigration, new buildings and new appliances which affects all parameters in the model, not only a constant hour-over-hour increase on load directly. Equation 4.2 shows how a multiplicative coefficient $\frac{Gm \times t}{8760}$ multiplies all other coefficients, creating a non-linear model in the parameters.

$$E(t) = \frac{Gm \times t}{8760} \times X(t) \times \beta \qquad \forall t \in T \tag{4.2}$$

To identify *Gm* in the multiplicative trend and to keep the model solvable as a linear regression, the average load for the first 21 days (504 hours) and the last 21 days of the year was used to determine a percentage increase on load. This trend is then removed from the electricity load data before the data normalization process and the regression analysis. The growth rate identified (13.9%) is higher than the population growth of 7.7% averaged between 2005 and 2010 [27], due to an increase in electricity use intensity, the amount of electricity used per individual.

4.2 Seasonality

Electricity demand is known to exhibits seasonal fluctuations. This fluctuations are caused by yearly seasonal patterns, as well as higher frequency, weekly and daily fluctuations.

From the cooling load and non-cooling load (offset) distinction, four methods were tested in order to determine the best result when modeling the seasonal variation of the offset. The seasonal variation are the changes in consumption not directly related to short term/high frequency changes in weather parameters, therefore the objective was to identify changes on a monthly basis or longer timespan.

4.2.1 Dummy Variables

Dummy variable, also called binary variable, indicator variable and boolean indicator are variables which take the value of either 0 or 1. In a regression model, a dummy variable with a value of 0 will cause its coefficient to disappear from the equation. Conversely, the value of 1 causes an additive coefficient to the model, which acts as a offset to the intercept term. In a linear regression modeling as described here, this type of specification is used to define subsets of observations that have different intercepts (each month) without the creation of separate models (multiple-equation models) [33]. The offset component of the total load is the summation of the main offset (first hour of first workday of the month of January) with incremental terms for each month. With this scheme, 11 dummy variables *(0/1)* were used to account for the monthly changes on load. A study using a similar approach of having step-wise constant monthly offset is described in [26]. Table 4.1 represents the assignment of dummy variables to each month and Figure 4.1 illustrates the Dummy variables distribution over the year, each color representing one variable.

Considering β_0 as the intercept term for the first hour of first workday of the month of January, and $\beta_1 \ldots \beta_{11}$ as the monthly coefficients for each dummy variable, Equation 4.3 gives the overall contribution of the monthly seasonality $(\hat{y}_s)$ and an error term ε.

$$\hat{y}_s(t) = \beta_0 + D1 \times \beta_1 + \ldots + D11 \times \beta_{11} + \varepsilon \tag{4.3}$$

Table 4.1: Month dummy variables

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11
January	0	0	0	0	0	0	0	0	0	0	0
February	1	0	0	0	0	0	0	0	0	0	0
March	0	1	0	0	0	0	0	0	0	0	0
April	0	0	1	0	0	0	0	0	0	0	0
May	0	0	0	1	0	0	0	0	0	0	0
June	0	0	0	0	1	0	0	0	0	0	0
July	0	0	0	0	0	1	0	0	0	0	0
August	0	0	0	0	0	0	1	0	0	0	0
September	0	0	0	0	0	0	0	1	0	0	0
October	0	0	0	0	0	0	0	0	1	0	0
November	0	0	0	0	0	0	0	0	0	1	0
December	0	0	0	0	0	0	0	0	0	0	1

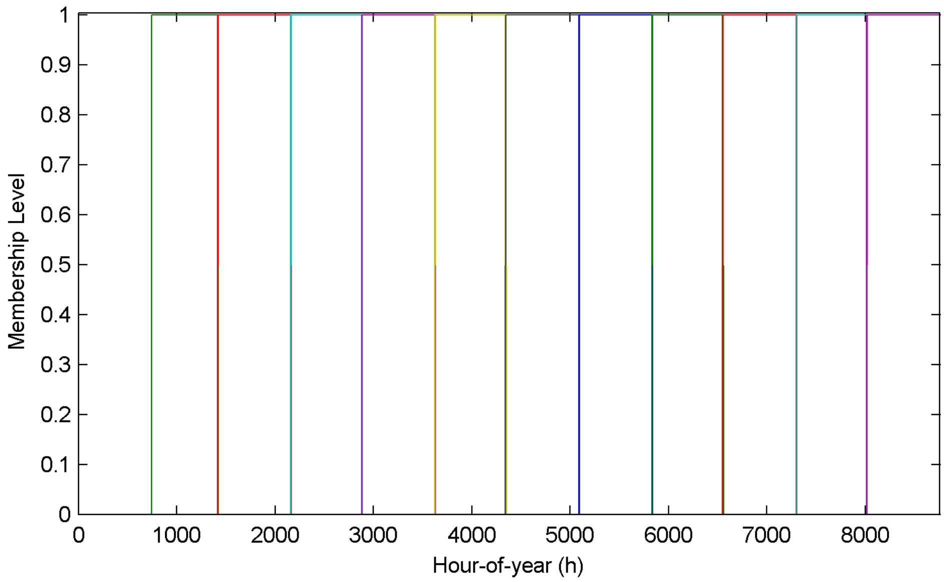

Figure 4.1: Dummy variables membership function.

4.2.2 Fuzzy Variables

Using the previous method of Dummy variables, a clear borderline between electricity consumption in one month and a subsequent month is delimited. Since in practice behavior and therefore electricity consumption does not abruptly change when the calendar month changes, a proposed improvement using fuzzy sets was tested.

The concept of fuzzy variables is that the membership of an element is not strictly false or true as in the Dummy variable method, but a gradual set to range continuously through the interval [0,1] as presented in Figure 4.2 for the example of a "tall men" using a sigmoid membership function. Fuzzy logic is used to deal with the imprecision of many real-world problems and vague concepts for example hot, tall, young, round, wealthy, etc.

In order to eliminate the abrupt changes in the seasonality parameters between consecutive months, a fuzzy set combination [83] was used to produce a smooth transition between them. Figure 4.3 illustrates the Dummy variables from Table 4.1

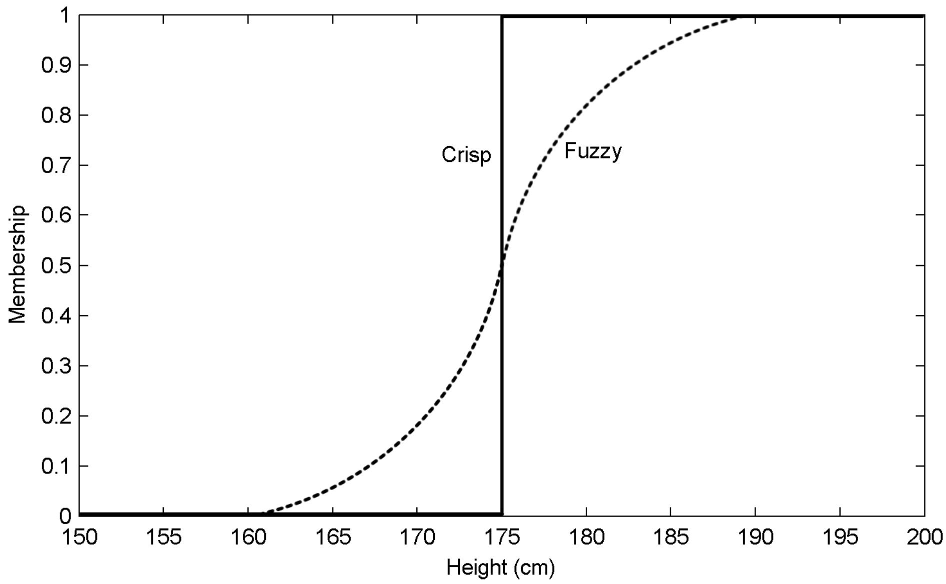

Figure 4.2: Two definitions of the set of "tall men", a crisp and a fuzzy set. [73]

after the fuzzyfication process, each color representing one variable. The marginal impact of the monthly seasonal parameter is combined using a triangular membership function to fuzzify each input, which represents a smooth linear behavior change between two consecutive months. The fifteenth day of each month's seasonal parameter is entirely and only dependent on that month's parameter, while other days seasonal value is determined by a linear interpolation of the parameter for the two neighboring months. A similar approach was used by [19], where a layer of nodes (fuzzy input neurons) was used in a Artificial Neural Network model for monthly electricity demand forecasting. Similarly to the Equation 4.3, Equation 4.4 represents the contribution of the monthly seasonality ($\hat{y}_s$), where $F1$ to $F11$ represents the fuzzy variables.

$$\hat{y}_s(t) = \beta_0 + F1 \times \beta_1 + ... + F11 \times \beta_{11} + \varepsilon \tag{4.4}$$

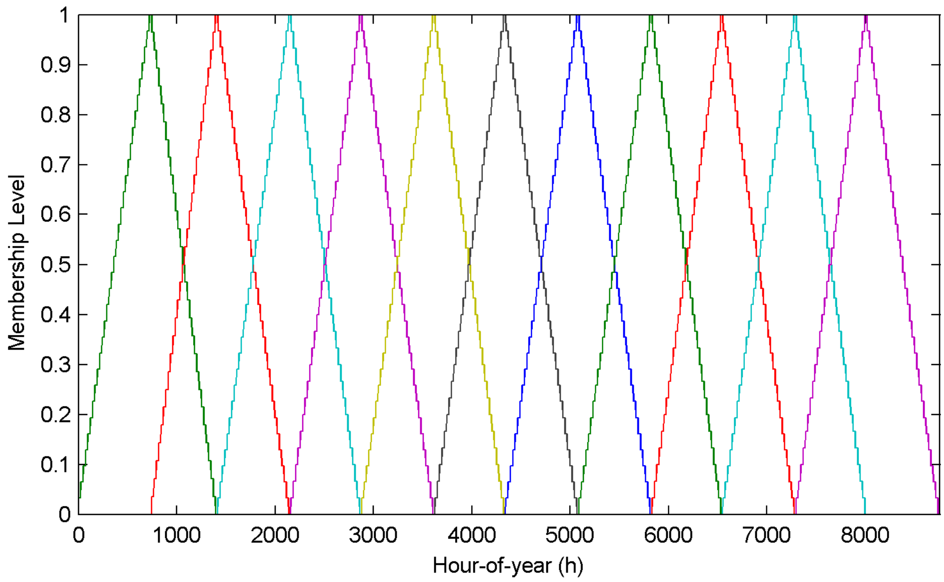

Figure 4.3: Fuzzy variables membership function.

4.2.3 Fourier Series

To account for the periodic calendar-driven variation on the load not explainable by the exogenous weather parameters, Fourier series decomposition was used to model seasonality in a deterministic way. Fourier series decomposition is based on the fact that a periodic function can be expressed by the sum of sine and cosine functions. If the function under analysis is *absolutely integrable*[1], that is sufficient for the convergence of the Fourier integral, this condition being sufficient, but not necessary [34].

Considering a function $f(t)$, a continuous Fourier series can be expressed as Equation 4.5, being a summation of possibly infinite harmonics of sine and cosine waves. $2\pi/\tau$ is the main frequency and k represents the harmonics multiplier.

$$f(t) = a_0 + \sum_{k=1}^{\infty} [a_k \sin(2k\pi t/\tau) + b_k \cos(2k\pi t/\tau)] \tag{4.5}$$

[1] An integrable function f on $I \to \mathscr{R}$ is said to be absolutely integrable on I if $|f|$ is also integrable on I [11].

The variation on load through the year was observed before by [68] to follow a *sinusoidal* behavior, therefore advocating the usage of Fourier decomposition using sines and cosines mainly concerned with the measurement of the calendar-driven component as yearly, weekly and daily seasonal effects.

As can be seen and Figure XXX, due to the relative high frequency of the data here studied (hourly), can be observed in Figure XXX a yearly *sinusoidal* variation, as well in Figure XXX an intra-week and intra-day *sinusoidal* variation.

In order to model the seasonal effect in the electricity load that is not captured by the exogenous weather variables, the following Fourier terms are considered: Equation 4.6 represents the intra-day variation, where τ_1 represents the daily period, i.e., 24 *h* and *ord* represents the order of the series. Since this function represents the average workday, a series of dummy variables was included in order to model the variation between a workday hour, and a Friday, Saturday, Holiday or Ramadan hour (Section 4.6).

$$\sum_{i=1}^{ord}[b_{2i-1}\sin(2i\pi t/\tau_1)+b_{2i}\cos(2i\pi t/\tau_1)] \tag{4.6}$$

Similarly Equation 4.7 represents the weekly variation, where τ_2 represents the weekly period, i.e., 168 *h* and *ord* represents the order of the series.

$$\sum_{i=1}^{ord}[c_{2i-1}\sin(2i\pi t/\tau_2)+c_{2i}\cos(2i\pi t/\tau_2)] \tag{4.7}$$

Equation 4.8 represents the yearly variation while τ_3 represents the length of the standard year, i.e., 8760 *h* corresponding to 365 days which both 2010 and 2011 happen to be. (alternatively, one could use the true length of the solar year, i.e., 8766 *h* corresponding to 365.25 days [28]).

$$\sum_{i=1}^{ord}[d_{2i-1}\sin(2i\pi t/\tau_3)+d_{2i}\cos(2i\pi t/\tau_3)] \tag{4.8}$$

Different combinations of Fourier series orders and frequencies were tested in order to achieve the best seasonality representation.

4.2.4 Moving Average Filter

Another approach for accounting the seasonality and variations not related to measurable exogenous variables is the moving average technique. This approach is used to smooth high frequencies oscillation generally caused by weather, to isolate the seasonality effects and to dampen the noise. By defining the moving average filtered data as $\hat{m}_t$, for the year seasonality, a moving average filter of order 168 (1 week) was used Equation 4.9, while for the weekly seasonality a filter with order 24 (1 day) was used Equation 4.10.

$$\hat{m}_t^y = \frac{1}{168}(0.5x_{t-64} + x_{t-63} + ... + x_{t+63} + 0.5x_{t+64}) + \varepsilon \tag{4.9}$$

$$\hat{m}_t^m = \frac{1}{24}(0.5x_{t-12} + x_{t-11} + ... + x_{t+11} + 0.5x_{t+12}) + \varepsilon \tag{4.10}$$

4.3 Solar Irradiation

The sun, as any other object with absolute temperature above absolute zero emits radiation. Radiation is transmitted, absorbed or scattered in different amounts, depending on the wavelength, 97% of this emitted radiation being in the range of 290 nm to 3000 nm for the sun. The interaction of the solar radiation with the Earth's atmosphere result in three solar components, being *(DNI)* further decomposed into two: [77]

- Direct normal irradiance *(DNI)* – Defined as the amount of solar radiation from the solar disk (beam) per unit of area on a surface always normal to the solar rays coming in a straight line from the sun.

- Diffuse horizontal irradiance *(DHI)* – Defined as the amount of solar radiation per unit of area that does not arrive in direct path from the sun, being scattered by the atmosphere and aerosols.
- Global horizontal irradiance *(GHI)* – Defined as the total hemispheric irradiance per unit of area, being the geometric sum of the DNI and DHI.
- Direct normal irradiance vertical *(DNI_V)* – Component of DNI that is incident on vertical surfaces, usually used to account for the irradiance on walls.
- Direct normal irradiance horizontal *(DNI_H)* – Component of DNI that is incident on horizontal surfaces, used to model the irradiance on roofs.

When specific measurement data for each component is not available, the solar azimuth angle *(SZA)* is an important and necessary information to calculate the geometric summation of GHI from DNI and DHI and to decompose DNI into its components. The relation between GHI, DNI, DHI and SZA is given by the Equation 4.11. The decomposition of DNI is given by the Equation 4.12 for the vertical component and Equation 4.13 for the horizontal component.

$$GHI = DNI \times \cos(SZA) + DHI \tag{4.11}$$

$$DNI\_V \times \cos(SZA) \tag{4.12}$$

$$DNI\_H \times \sin(SZA) \tag{4.13}$$

Direct normal irradiance measurements were decomposed into vertical *(DNI_V)* and horizontal *(DNI_H)* planes irradiance calculated as described in [71], based the solar position for Abu-Dhabi for every hour-of-day and day-of-year calculated using a function by Vincent Roy [2], following the algorithm described in [71].

4.3.1 Sol-Air Temperature

Solar irradiation absorbed by external building surfaces often provide the majority of the envelope thermal gain [52]. The notion of *sol-air* temperature is an improvement over the simplified practice of accounting only for the difference between outdoor envelope temperature and ambient dry-bulb temperature [62, 31], since it accounts for solar irradiation absorption. The *sol-air* radiation can be defined as "the equivalent outdoor temperature which will cause the same rate of heat ow at the surface and the same temperature distribution throughout the material as results from the outdoor air temperature and the net radiation exchange between the surface and its environment" [70] as represented in the equation 4.14.

$$T_{sol-air} = T_o + \frac{\alpha \times I - \Delta Q_{ir}}{h_o} \tag{4.14}$$

Where:

T_o: Outdoor temperature.

α: Average solar radiation absorptivity of the surfaces.

I: Global solar irradiance incident on the surfaces.

ΔQ_{ir}: Extra infrared radiation due to difference between external air temperature and apparent sky temperature.

h_o: Heat transfer coefficient for radiation and convection.

In addition, this formulation enables us to explicitly consider the impact of wind speed via its influence on the convection transfer parameter h_o. Different combinations of solar components were tried in the model and combination with wind were tested.

[2]Function available at: http://www.mathworks.com/matlabcentral/fileexchange/4605-sunposition-m/content/sun_position.m

4.4 Air Dry-bulb Temperature

Ambient air dry-bulb temperature and electricity load is known to have a non-linear relationship as described in 3.3. After ambient air dry-bulb temperature being normalized using a cut-off point (18.5 °C) impact on load was modeled as a polynomial function to account for the saturation effect at very high temperature (reduced COP of chillers) as well as the gradual take-off of cooling near the cut-off temperature [12]. A similar non-linear relationship with temperature is presented by Hagan [37] for load forecasting.

Heating load was studied in [38], and a similar behavior between ambient air dry-bulb temperature and electricity load was identified for cooling load in Abu-Dhabi, where at high outdoor temperature the ability to achieve the desired cooling can be limited by the overall capacity of the cooling system in place. Also if the cooling system cannot supply the necessary cooling load, a saturation effect is expected to be identified, where an increase in temperature has a lower marginal impact on electricity consumption. To account for the non-linearity between temperature and load, [69] simply uses the square of temperature, while here a sixth-order polynomials, retaining only even exponents was used for a better representation of the gradual increase of temperature impact close to the cut-off and the saturation effect.

To account for delays caused by heat transfer between building walls and roofs, [6] used a set of lagged temperature variables, and the same approach being tested in the model here defined. Seasonality of the lagged temperature parameter was tested and discarded.

A second approach to account for the heat transfer inertia through the building was based in the previously cited Eventail load model of Electricité de France [17], where the concept of a smoothed temperature is used as a proxy of the temperature felt inside the premises, which can be inferred to be directly proportional to the sensible cooling load required. Assuming hourly sampled values, the smoothed temperature T_s is defined by Equation 4.15 as an exponentially weighted moving average filter. The value used for κ has significant sensibility to the forecasting accuracy of the model, in this case the same value suggested in [17] (0.98), equivalent to a filter time constant of $\tau = 50\ h$ ($\kappa = e^{-1/\tau}$) which produced was a good estimate for the Abu Dhabi case. In order to have a similar proxy as the Sol-Air temperature 4.3.1, a horizontal DNI component was included to define the sensible cooling load proxy. As for the optimal value of ω, the multiplier for the horizontal DNI component in the smoothed temperature, a value of 13 °C.m^2/kW was determined through line search minimization of the forecasting error.

$$T_s(t) = \kappa T_s(t-1) + (1-\kappa)(T(t) + \omega I) \tag{4.15}$$

The smoothed temperature represents the heat transfer inertia through the building, and in order to account for the direct impact of temperature, for example affecting the chiller's COP or heat entering the building directly through air infiltration, a composite temperature θ is defined by Equation 4.16. In the same way as ω in the smoothed temperature, α was identified by line search minimization of the forecasting error as 0.19, thereby giving considerable weight (more than 80%) to the smoothed temperature.

$$\theta = (1-\alpha) T_s + \alpha T \tag{4.16}$$

The composite temperature θ approach produces a cleaner final model, as a single coefficient is identified for the combined effect of all weather variables affecting the overall Abu Dhabi's electricity sensible cooling load.

4.5 Air Humidity

Since most cooling units in the UAE are air cooled and do not use wet cooling towers, chiller's coefficient of performance *(COP)*[3] is generally not affected by the ambient air humidity level. However, through infiltration, ventilation and fresh air intake, outdoor air humidity has an impact on latent cooling load.

To account for the latent cooling fraction of the load, different variables derived from relative humidity and temperature were used following thermodynamics equations as:

- specific humidity
- enthalpy
- Humidex index

Specific humidity *(SH)* values were converted using thermodynamics equations as the vapor pressure calculated from Equation 4.17 and Equation 4.18, where Tdc is the dew-point temperature, Tc the ambient temperature in Celsius, and RH relative humidity. $Pressure$ assumed to be 985 mBar.

$$e = \frac{(6.112 \times \exp^{\frac{17.67 \times Tdc}{243.5 + Tdc}} \times RH)}{100} \tag{4.17}$$

$$SH = \frac{0.622 \times e}{Pressure - (0.378 \times e)} \tag{4.18}$$

In order to model this constituent of the electrical load, SH, was used as one of the input variables of the model. The impact of SH is modeled as a polynomial due to the gradual start-up and the saturation effect, similar to temperature. Both

[3] The coefficient of performance *COP* is defined as the ratio between the cooling provided, divided by the total energy consumption (here defined as electricity) to provide that cooling.

polynomials are sixth-order, retaining only even exponents; in addition, each polynomial is formulated in such a way that the polynomial annuls at 0 (ie, at min Temp or min SH) when is cooling load is not affecting electricity load. Specific humidity was also used in the sensitivity analysis of electricity consumption in business districts of Tokyo, Japan [42].

Enthalpy *(ENT)* was calculated using Equation 4.19, combination of Specific Humidity and Temperature. ENT is given in kJ/kg, SH is the Specific humidity in kg/kg and $Temp$ the dry-bulb temperature. Enthalpy was considered as a linear component in the model.

$$ENT = Temp \times (1.01 + 0.00189 \times (SH \times 1000)) + 2.5 \times (SH \times 1000) \quad (4.19)$$

Thirdly the Humidex index *(HDX)*, a measurement index first used in 1965 in Canada to describe how hot and humid weather feels to the average person, using an easy to understand index. This index also relates the dry-bulb temperature with Specific humidity using the calculated vapor pressure e. The index calculation is straight forward given by Equation 4.20.

$$HDX = Temp + (\frac{5}{9} \times (e - 10)) \quad (4.20)$$

Air infiltration can be related to wind speed [57, 55]. To account for this phenomenon, a set of variables relating wind with humidity, temperature and GHI were also included in the model.

4.6 Day-of-week, Holiday, Ramadan and Time-of-day

Since 1^{st} of September 2006, the regular work week for most UAE office workers ranges from Sunday to Thursday, while Friday and Saturday constitute the weekend. In addition, certain major holidays affect the aggregate load profile and present similarities with the Friday profile, therefore public holidays are accounted as being Fridays. The month of Ramadan, a period of 29 or 30 days also requires changes on schedules and activities being specifically accounted for in the model. A total of 4 day-types are to be distinguished: Weekday, Friday, Saturday and Ramadan.

To properly account for the hourly and daily variability of electricity consumption, the proposed model includes dummy variables representing different hours of day for regular workdays, as well as Fridays, Saturdays and Ramadan workday. A similar approach was adopted in a forecasting model for Greece [59].

The offset component of the total load is the summation of the main off-set (first hour of first workday of the month of January) with *incremental* terms obtained by multiplying time-of-day/day-of-week parameters with corresponding binary *(0/1)* dummy variables [64]. More precisely, the dummy variables represent:

- $\delta_i^w(t)$ Hourly parameter for hours (i) 2 through 24 of workdays: account for the offset variation between different hours of a regular (non-Ramadan) workday.
- $\delta_i^f(t)$ Hourly parameter for hours 1 through 24 of Fridays: Fridays and Saturdays are not regular workdays in the studied region, therefore the offset for these days have different profile from regular workdays. Holidays have similar schedules as Fridays; therefore holidays happening on what would be a regular workday are treated as Fridays.
- $\delta_i^s(t)$ Hourly parameter for hours 1 through 24 of Saturdays: Similarly to

Fridays, Saturday are not regular workdays, but some offices and retail facilities operate to some extent.

- $\delta_i^r(t)$ Hourly parameter for hours 1 through 24 of Ramadan workdays: Main population behavior and schedules change during Ramadan workdays. A dummy variable and parameter was used to account for these changes. Ramadan Fridays and Saturdays are not treated differently since Ramadan's main electrical impact is considered to be due to the change in working hours.

These four items, together with the seasonality parameter described in the Subsection 4.2 were tested in the model, accounting for long term offset variations following Equation 4.21.

$$\sum_{i=2}^{24}[w_i \delta_i^w(t)] + \sum_{i=1}^{24}[f_i \delta_i^f(t) + s_i \delta_i^s(t) + r_i \delta_i^r(t)] \tag{4.21}$$

For the hourly variation when using the Fourier Series 4.2.3, the work-hour variation is modeled by a series of sines and cosines, while the increment or decrement of Friday, Saturday or Ramadan hour is done using the dummy variables as explained in this section.

4.7 Price Elasticity

Electricity price elasticity is usually accounted for in similar studies, since price fluctuations often produce significant changes in electricity consumption [13]. Abu Dhabi has seen stable electricity prices in the past years; therefore the impact of price elasticity cannot be inferred from the data available and the study period here utilized. Furthermore, for reasons beyond the scope of the present study, the pricing level is unlikely to be used in a DSM context within the emirate of Abu-

Dhabi in the near future. For the reasons mentioned, price elasticity was therefore not considered in the present study, but could be a important addition to future work as electricity price variation happens.

CHAPTER 5

Model Results & Verification

The previous chapters described the source of data, the pre-processing and treatment to generate the load constituents proxies in order to identify the best representative model of electricity consumption for Abu-Dhabi.

Since the initial models tested, autocorrelation of the residuals was identified, a phenomenon common when using time series data being interpreted as independent samples, fact that due to the nature of the problem here stated could not be directly solved. Under the Gauss-Markov assumption the residuals are assumed to be independent distributed and non-correlated to each other: $Cov(\varepsilon(i), \varepsilon(j) = E(\varepsilon(i)\varepsilon(j)) = 0, i \neq j$. The Ordinary Least Squares *(OLS)* framework is known to produce under estimation of the errors when this assumption is violated, therefore affecting the t-statistics and hypothesis testing for the parameters coefficients identified. In order to overcome this problem, an identification method robust to autocorrelation and heteroskedasticity in the error terms was used. The Newey-West estimator, devised by Whitney K. Newey and Kenneth D. West in 1987 [60], provides an estimate of the covariance matrix of the parameters to improve the OLS regression.

A series of different models were tested with a combination of the variables and methods previously discussed, in order to identify the best model representation of the overall electricity load behavior. The models were compared on the coefficient identification phase using the Adjusted R^2 given by the parameter estimation tool. The best identified models were further tested to measure the accuracy of the prediction $\hat{y}(t)$. The Root Mean Square Error *RMSE* defined by Equation 5.1 and the Mean Absolute Percentage Error *(MAPE)* Equation 5.2, which has been traditionally used to measure accuracy in load forecasting, were used to test the models' forecast capability against the measured data from the first half of 2011.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}(y(t) - \hat{y}(t))^2}{T}} \tag{5.1}$$

$$MAPE = \frac{100}{T}\sum_{t=1}^{T}\left|\frac{(y(t) - \hat{y}(t))}{y(t)}\right| \tag{5.2}$$

5.1 Model Representation

As discussed in Chapter 4, electricity load $y(t)$ can be decomposed into two of components. The baseload $y_b(t)$ that varies only with calendar components, not dependent on weather parameter, and the cooling load $y_c(t)$ which is assumed to be completely dependent on weather. Together with these two components, there is the growth trend described on Section 4.1, a additive linear trend applied to the model on Equation 5.3 and the multiplicative trend that applies to both the baseload and cooling load are presented on Equation 5.4. Both methods were tested in combination with the other parameters, as the best trends representation was not clear from initial tests.

$$y(t) = Gl \times t + y_b(t) + y_c(t) \tag{5.3}$$

$$y(t) = \frac{Gm \times t}{8760} \times (y_b(t) + y_c(t)) \tag{5.4}$$

The base model considered the linear trend, with Fourier series of order 4 to account for seasonalities, composite temperature for the sensible cooling component and Enthalpy accounting for the latent cooling load.

The table below will summarize the results: To be developed...

Table 5.1: Add caption

Prediction	Base	Fourier 6	Fourier 8	Growth Mult
RMSE	2.0996 (2.04%)	2.0996 (2.04%)	2.0493 (1.99%)	2.1982 (2.19%)
MAPE	2.77%	2.76%	2.75%	3.01%
AdjR2	0.9867	0.9866	0.9872	0.9835

CHAPTER 6

Applications

The accessibility of reliable electricity is responsible for a great share of the development achieved in recent decades, despite the fact that many regions of the world still lag the availability of advanced energy sources and reliable supply. Electricity demand models are used in this scenario to forecast capacity requirements, area of service and overall infrastructure investment. This application will usually demand a long term forecast model.

The competitiveness in the electricity market has increase in recent years due to the implementation of energy markets, where different players are involved in buying and selling electricity, usually on a day-ahead scheme, having to rely on demand forecasts to estimate load and subsequently prices. Due to the variability and dynamics of the system, an accurate forecast of the load is not always possible, incurring in great financial impact of forecast errors. A large number of publications has been made targeting short term forecast, specifically targeting this sector.

Mid-term forecast models, as proposed in this work, when applied strictly into the future, have applications mainly on system optimization. The specific application here described is the assessment of Demand Side measures by establishing a baseline for comparison. The objective is not to predict the future, but to identify what would have happened if a set of changes to the system were not made. By comparing the actual data after a energy efficiency measure or policy to the predicted consumption of the system without changes. A secondary objective of the model established was to maintain significance with the underlying physical process. Having coefficients representing physical parameters that can be improved with energy efficiency measures, enables an estimation of the impacts on load of such measures when using a sensitivity analysis.

6.1 Renewable Energy - Solar

One of the well debated demand side solution for reducing CO_2 emissions and overall transmission and distribution costs is the adoption of distributed generation using rooftop photo-voltaic *(PV)* panels. The assessment of such measure, the inclusion of a electricity generation equipment on the demand side of the electricity system, can be made in a relative straight forward way by adding another term to the model, with negative coefficient to represent electricity generation. The Standard Test Conditions *STC* based on the standard IEC 61215 for Crystalline Silicon Modules and IEC 61646 for Thin Film Modules state a test condition of $1000W/m^2$ of DNI, and 25 °C. Based on the rated capacity, a sensitivity analysis can be conducted to analyze the impact of different PV penetration rate in the studied area.

6.2 Policy Analysis

As an example of the model application in M&V, after a city level program for energy efficiency was conducted, it is desired to calculate the overall impact of such measure. Due to the high dynamics of the electricity demand, and the impact of exogenous parameters as weather, the consumption of one day, cannot be directly compared to a different day.

For such assessment the model here proposed can be used to generate a baseline energy consumption, based on the conditions before the energy efficiency program, and compare to the actual measured data after the program. Figure XXX below presents a fictitious data from an energy efficiency program where the main target was to promote the increase of thermostats set points, reducing air conditioning load. An actual trial of such measure was conducted by [61], where the participating house-holds were knows, and the load could then be compared to a test group. In the case where the test group cannot be delimited, using a baseline energy model is the most common approach.

CHAPTER 7

Conclusion

The aim of this Master's thesis was to develop a model of electricity consumption for the city of Abu Dhabi, UAE, in order to overcome a known barrier of accurate determining the *ex-post* impact of the deployment of Demand Side Management *(DSM)* initiatives. DSM measures ranging from advanced building controls to indoor-climate control and building equipment/envelope enhancement are designed to address the increasing awareness of climate change, pollution, reduced infrastructure investment availability and escalating fossil fuel prices. In order to achieve this objective, a low complexity linear model (linear in the parameters as the relationship between load and weather is known to be non-linear) was developed.

A series of parameters and methods were tested in order to identify the best representation of the drivers of electricity load, maintaining the physical correspondence with the underlying process. Hour-of-day, day-type and seasonalities were considered to model the fraction of the load not related to weather. For the weather related fraction, cooling load as a heating season was not observed in the studied area, a combination of parameters such as temperature, specific humidity, wind speed and solar irradiance, as well as the time delays involved between the outside condition and the actual impact on the electricity load were tested.

The hourly electricity load for 2010 was used in the parameter estimation phase, with adjusted R-squared reaching 0.9938. The first half of 2011 data was used for testing the model prediction capabilities, resulting in a RMSE of 34.19MW, equivalent to 1.99% of the growth-trend corrected annual peak load, and a MAPE of 2.75%. It was identified that the cooling load account for approximately XX% of the overall electricity usage, while during the peak summer hour this value reaches YY% of the total electricity load.

This model proved to achieved accurate results in developing a baseline electricity consumption profile, to identify what would have happened if a set of changes to the system (DSM) were not made. By comparing the baseline generated from the model, with the data actually being measured, it is possible to determine the actual savings and impacts generated from the energy efficiency intervention/policy.

Was observed in the analysis of the model residuals that autocorrelation is still present, this indicates that some information was not fully accounted for in the model. The same problem was found in similar studies, explained by the limitation of a linear regression model to fully represent the dynamic behavior of load. For future research a model using a more dynamic modeling approach (state-space, artificial neural networks) might be able to fully utilize this information present in the data, which could produce better results with the cost of added complexity, or lost physical significance.

Bibliography

[1] J. Al-Zayer and A. A. Al-Ibrahim. Modelling the impact of temperature on electricity consumption in the eastern province of saudi arabia. *Journal of Forecasting*, 15(2):97–106, 1996.

[2] M. T. Ali, M. Mokhtar, M. Chiesa, and P. Armstrong. A cooling change-point model of community-aggregate electrical load. *Energy and Buildings*, 43(1): 28–37, 1 2011.

[3] N. Amjady. Short-term hourly load forecasting using time-series modeling with peak load estimation capability. *Power Systems, IEEE Transactions on*, 16(3):498–505, 2001. ID: 1.

[4] F. Apadula, A. Bassini, A. Elli, and S. Scapin. Relationships between meteorological variables and monthly electricity demand. *Applied Energy*, 98: 346–356, 10 2012.

[5] J. Asafu-Adjaye. The relationship between energy consumption, energy prices and economic growth: time series evidence from asian developing countries. *Energy Economics*, 22(6):615–625, 2000.

[6] H. Asan and Y. S. Sancaktar. Effects of wall's thermophysical properties on time lag and decrement factor. *Energy and Buildings*, 28(2):159–166, 10 1998.

[7] ASHRAE. Guideline 14-2002, measurement of energy and demand savings, 2002.

[8] M. Aydinalp, V. I. Ugursal, and A. S. Fung. Modelling of residential energy consumption at the national level. *International Journal of Energy Research*, 27(4):441–453, 2003.

[9] E. Badar-Ul-Islam and S. A. Qureshi. Comparison of conventional and modern load forecasting techniques based on artificial intelligence and expert systems. *International Journal of Computer Science*, 8(5):504, 2011.

[10] Bashmakov-I. Bernstein L. Bogner J. Bosch P. Dave R. ... & Zhou D. Barker, T. *Mitigation from a cross-sectoral perspective.*, chapter 11.7. 2. Summary for Policymakers IPCC Fourth Assessment Report, Working Group III. 2007.

[11] Robert Gardner Bartle. *A modern theory of integration*. American Mathematical Society, 2001.

[12] M. Bessec and J. Fouquau. The non-linear link between electricity consumption and temperature in europe: A threshold panel approach. *Energy Economics*, 30(5):2705–2721, 9 2008.

[13] V. Bianco, O. Manca, and S. Nardini. Electricity consumption forecasting in italy using linear regression models. *Energy*, 34(9):1413–1421, 9 2009.

[14] T. A. Boden, G. Marland, and R. J. Andres. Global, regional, and national fossil-fuel co2 emissions. Technical Report 10, Carbon Dioxide Information Analysis Center Environmental Sciences Division Oak Ridge National Laboratory, 2009.

[15] M. M. Breunig, H. P. Kriegel, R. T. Ng, and J. Sander. Lof: identifying density-based local outliers. In *ACM Sigmod Record*, volume 29, pages 93–104. ACM, 2000.

[16] R. A. Brownsword, P. D. Fleming, J. C. Powell, and N. Pearsall. Sustainable citiesmodelling urban energy supply and demand. *Applied Energy*, 82(2): 167–180, 2005.

[17] A. Bruhns, G. Deurveilher, and J. S. Roy. A non-linear regression model for mid-term load forecasting and improvements in seasonnality. In *Proceedings of the 15th Power Systems Computation Conference*, pages 22–26, 2005.

[18] J. R. Cancelo, A. Espasa, and R. Grafe. Forecasting the electricity load from one day to one week ahead for the spanish system operator. *International Journal of Forecasting*, 24(4):588–602, 2008.

[19] P. Chang, C. Fan, and J. Lin. Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach. *International Journal of Electrical Power & Energy Systems*, 33(1):17–27, 1 2011.

[20] S. G. Chavez, J. X. Bernat, and H. L. Coalla. Forecasting of energy production and consumption in asturias (northern spain). *Energy*, 24(3):183–198, 1999.

[21] B. J. Chen and M. W. Chang. Load forecasting using support vector machines: A study on eunite competition 2001. *Power Systems, IEEE Transactions on*, 19(4):1821–1830, 2004.

[22] S. Chen, W. Wang, and H. van Zuylen. A comparison of outlier detection algorithms for its data. *Expert Systems with Applications*, 37(2):1169–1178, 3 2010.

[23] T. W. S. Chow and C. T. Leung. Neural network based short-term load forecasting using weather compensation. *Power Systems, IEEE Transactions on*, 11(4):1736–1742, 1996. ID: 1.

[24] J. A. Clarke, C. M. Johnstone, M. Lever, L. B. McElroy, F. McKenzie, G. Peart, L. Prazeres, and P. A. Strachan. Simulation support for the formulation of domestic sector upgrading strategies. In *Proceedings of 8th International IBPSA conference*, pages 219–226, 2003.

[25] R. Cottet and M. Smith. Bayesian modeling and forecasting of intraday electricity load. *Journal of the American Statistical Association*, 98(464):839–849, 2003.

[26] E. B. Dagum. Modelling, forecasting and seasonally adjusting economic time series with the x-11 arima method. *Journal of the Royal Statistical Society.Series D (The Statistician)*, 27(3):203–216, 09/01 1978.

[27] Statistics Center Abu Dhabi. Statistical yearbook of abu dhabi 2011. Technical Report 2011, 2011.

[28] V. Dordonnat, S. J. Koopman, M. Ooms, A. Dessertaine, and J. Collet. An hourly periodic state space model for modelling french national electricity load. *International Journal of Forecasting*, 24(4):566–587, 2008.

[29] V. Dordonnat, S. J. Koopman, and M. Ooms. Dynamic factors in periodic time-varying regressions with an application to hourly electricity load modelling. *Computational Statistics & Data Analysis*, 2011.

[30] Esam Elsarrag and Yousef Alhorr. Using dynamic facade for indoor air quality, thermal comfort and energy efficient air conditioning. *Khartoum University Engineering Journal*, 1(2), 2011.

[31] D. G. Erbs, S. A. Klein, and W. A. Beckman. Sol-air heating and cooling degree-days. *Solar Energy*, 33(6):605–612, 1984.

[32] E. A. Feinberg and D. Genethliou. Load forecasting. *Applied mathematics for restructured electric power systems*, pages 269–285, 2005.

[33] S. Garavaglia and A. Sharma. A smart guide to dummy variables: four applications and a macro. In *Proceedings of the Northeast SAS Users Group Conference*, 1998.

[34] Bernd Girod, Rudolf Rabenstein, and Alexander Stenger. *Signals and systems*. Wiley, 2001.

[35] William H. Greene. *ECONOMETRIC ANALYSIS*. 2003.

[36] P. C. Gupta and K. Yamada. Adaptive short-term forecasting of hourly loads using weather information. *Power Apparatus and Systems, IEEE Transactions on*, (5):2085–2094, 1972.

[37] M. T. Hagan and S. M. Behr. The time series approach to short term load forecasting. *Power Systems, IEEE Transactions on*, 2(3):785–791, 1987.

[38] A. Henley and J. Peirson. Nonlinearities in electricity demand and temperature: Parametric versus nonparametric methods. *Oxford Bulletin of Economics and Statistics*, 59(1):149–162, 1997.

[39] H. Huang, J. Lin, C. Chen, and M. Fan. Review of outlier detection. *Application Research of Computers*, 8:002, 2006.

[40] J. Huang, H. Akbari, L. Rainer, and R. Ritschard. *481 prototypical commercial buildings for 20 urban market areas*. US Department of Commerce, National Technical Information Service, 1991.

[41] L. C. Hunt, G. Judge, and Y. Ninomiya. Underlying trends and seasonality in uk energy demand: a sectoral analysis. *Energy Economics*, 25(1):93–118, 2003.

[42] T. Ihara, Y. Genchi, T. Sato, K. Yamaguchi, and Y. Endo. City-block-scale sensitivity of electricity consumption to air temperature and air humidity in business districts of tokyo, japan. *Energy*, 33(11):1634–1645, 2008.

[43] WWF International. Living planet report 2008. Technical Report 2012, 2008.

[44] WWF International. Living planet report 2012. Technical Report 2012, 2012.

[45] Demand Side Management Programme International Energy Agency. Task 18 - demand side management and climate change. Technical report, 2009.

[46] H. Jiawei and M. Kamber. Data mining: concepts and techniques. *San Francisco, CA, itd: Morgan Kaufmann*, 5, 2001.

[47] PJ Jones, S. Lannon, and J. Williams. Modelling building energy use at urban scale. In *7th International IBSPA conference, Rio de Janeiro, Brazil, August*, 2001.

[48] M. Kavgic, A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, and M. Djurovic-Petrovic. A review of bottom-up building stock models for energy consumption in the residential sector. *Building and Environment*, 45 (7):1683–1697, 7 2010.

[49] E. Knorr. Algorithms for mining distance-based outliers in large datasets. In *Proc. 24th Int. Conf. Very Large Data Bases*, pages 392–403, 1998.

[50] E. M. Knorr, R. T. Ng, and V. Tucakov. Distance-based outliers: algorithms and applications. *The VLDB Journal*, 8(3):237–253, 2000.

[51] M. Kolokotroni, I. Giannitsaris, and R. Watkins. The effect of the london urban heat island on building summer cooling demand and night ventilation strategies. *Solar Energy*, 80(4):383–392, 2006.

[52] T. H. Kuehn and J. W. Ramsey. Thermal environmental engineering, 1998.

[53] S. S. K. Kwok, R. K. K. Yuen, and E. W. M. Lee. An intelligent approach to assessing the effect of building occupancy on building cooling load prediction. *Building and Environment*, 46(8):1681–1690, 2011.

[54] C. N Lu, H. T Wu, and S. Vemuri. Neural network based short term load forecasting. *Power Systems, IEEE Transactions on*, 8(1):336–342, 1993. ID: 1.

[55] N. Malik. Field studies of dependence of air infiltration on outside temperature and wind. *Energy and Buildings*, 1(3):281–292, 4 1978.

[56] P. Mandal, T. Senjyu, and T. Funabashi. Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market. *Energy Conversion and Management*, 47(1516):2128–2142, 9 2006.

[57] G. E. Mattingly and E. F. Peters. Wind and trees: Air infiltration effects on energy in housing. *Journal of Wind Engineering and Industrial Aerodynamics*, 2(1):1–19, 1 1977.

[58] W. Mendenhall and T. Sincich. *A second course in statistics*. Prentice Hall, 1996.

[59] S. Mirasgedis, Y. Sarafidis, E. Georgopoulou, D. P. Lalas, M. Moschovits, F. Karagiannis, and D. Papakonstantinou. Models for mid-term electricity demand forecasting incorporating weather influences. *Energy*, 31(23):208–227, 0 2006.

[60] W. K. Newey and K. D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica: Journal of the Econometric Society*, pages 703–708, 1987.

[61] G. R. Newsham, B. J. Birt, and I. H. Rowlands. A comparison of four methods to evaluate the effect of a utility residential air-conditioner load control program on peak electricity use. *Energy Policy*, 39(10):6376–6389, 10 2011.

[62] P. W. O'Callaghan and S. D. Probert. Sol-air temperature. *Applied Energy*, 3 (4):307–311, 10 1977.

[63] Efficiency Valuation Organization. International performancemeasurement andverification protocol.concepts and options for determining energy and water savings -volume 1. Technical report, 2012.

[64] A. D. Papalexopoulos and T. C. Hesterberg. A regression-based approach to short-term system load forecasting. *Power Systems, IEEE Transactions on*, 5 (4):1535–1547, 1990.

[65] Angel Pardo, Vicente Meneu, and Enric Valor. Temperature and seasonality influences on spanish electricity load. *Energy Economics*, 24(1):55–70, 1 2002.

[66] D. C. Park, M. A. El-Sharkawi, R. J. Marks II, L. E. Atlas, and M. J. Damborg. Electric load forecasting using an artificial neural network. *Power Systems, IEEE Transactions on*, 6(2):442–449, 1991. ID: 1.

[67] E. Pelikn. Middle-term electric load forecasting by time series decomposition. *Electricity Load Forecast using Intelligent Technologies*, pages 167–176, 2002.

[68] D. Pilipovic. *Energy risk: Valuing and managing energy derivatives*. McGraw-Hill New York, 2nd edition, 2007.

[69] R. Ramanathan, R. Engle, C. W. J. Granger, F. Vahid-Araghi, and C. Brace. Short-run forecasts of electricity loads and peaks. *International Journal of Forecasting*, 13(2):161–174, 1997.

[70] K. R. Rao. *Some investigations on the sol-air temperature concept*. Commonwealth Scientific and Industrial Research Organization, Australia, Melbourne :, 1970. M1: Accessed from http://nla.gov.au/nla.cat-vn1101424.

[71] Ibrahim Reda and Afshin Andreas. Solar position algorithm for solar radiation applications. *Solar Energy*, 76(5):577–589, 2004.

[72] P. J. Robinson. Modeling utility load and temperature relationships for use with long-lead forecasts. *Journal of Applied Meteorology*, 36(5):591, 1997.

[73] J. C. Rodrigues. Theoretical foundations of fuzzy logic, 2008.

[74] Y. Shimoda, T. Asahi, A. Taniguchi, and M. Mizuno. Evaluation of city-scale impact of residential energy conservation measures using the detailed end-use simulation model. *Energy*, 32(9):1617–1633, 2007.

[75] LD Shorrock and JE Dunster. The physically-based model brehomes and its use in deriving scenarios for the energy use and carbon dioxide emissions of the uk housing stock. *Energy Policy*, 25(12):1027–1037, 1997.

[76] L. J. Soares and M. C. Medeiros. Modeling and forecasting short-term electricity load: A comparison of methods with an application to brazilian data. *International Journal of Forecasting*, 24(4):630–644, 2008.

[77] T. Stoffel, D. Renn, D. Myers, S. Wilcox, M. Sengupta, R. George, and C. Turchi. *Concentrating Solar Power: Best Practices Handbook for the Collection and Use of Solar Resource Data (CSP)*. National Renewable Energy Laboratory (NREL), Golden, CO., 2010.

[78] L. Suganthi and Anand A. Samuel. Energy models for demand forecasting a review. *Renewable and Sustainable Energy Reviews*, 16(2):1223–1240, 2 2012.

[79] J. W. Taylor, L. M. De Menezes, and P. E. McSharry. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, 22(1):1–16, 2006.

[80] SC Tripathy. Demand forecasting in a power system. *Energy conversion and management*, 38(14):1475–1481, 1997.

[81] V. M. Vlahovi and I. M. Vujoevi. Long-term forecasting: a critical review of direct-trend extrapolation methods. *International Journal of Electrical Power & Energy Systems*, 9(1):2–8, 1987.

[82] Y. Yamaguchi, Y. Shimoda, and M. Mizuno. Proposal of a modeling approach considering urban form for evaluation of city level energy management. *Energy and Buildings*, 39(5):580–592, 2007.

[83] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, /6 1965.