MIKKO NURMIRANTA
DATA DRIVEN LOAD MODELING AND CUSTOMER BEHAVIOR CHANGE DETECTION

Master of Science thesis

Examiner: Prof. Hannu Koivisto
Examiner and topic approved by the
Faculty Council of the Faculty of
Engineering Sciences
on 4th January 2017
ABSTRACT

MIKKO NURMIRANTA: Data driven load modeling and customer behavior change detection
Tampere University of Technology
Master of Science thesis, 54 pages
May 2017
Master’s Degree Programme in Automation Technology
Major: Intelligent Systems
Examiner: Prof. Hannu Koivisto
Keywords: change detection, Pearson divergence, smart grids, load forecasting, customer behavior

The strive for more efficient delivery of electricity and management of production resources in order to control costs and emissions has lead to innovations in power engineering and combined with input from other fields smart grids were developed. More variety and flexibility in production, better customer engagement and improvements in delivery reliability are just some perks offered by smart grids. In particular the vast amount of timely data collected from the customers enables massively improved network state analysis which is necessary for cost-effective distribution. The accuracy needs to be maintained and the detection of changes in consuming habits is useful for this purpose.

In this thesis a method for detecting changes in consumption patterns is proposed. By building a forecasting model for the consumer and comparing the load forecasts with load measurements changes not caused by external factors such as outdoor temperature should be possible to be detected. Two models, a periodic autoregressive model and an artificial neural network are used, with the former coming out as more accurate and simpler to use. The relationship between model forecasts and load measurements is monitored with cumulative sum and Pearson divergence and based on test cases the changes in load shapes can be detected and especially the method based on Pearson divergence achieves promising results. However forecasting the load of individual customers appears to challenging and better models are needed, even if the actual goal of change detection is accomplished.
TIIVISTELMÄ

MIKKO NURMIRANTA: Dataperusteinen sähkökulutuksen mallinus ja kulutuskäyttävymisen muutoksen havainnointi
Tampereen teknillinen yliopisto
Diplomityö, 54 sivua
Toukokuu 2017
Automaaatiotekniikan koulutusohjelma
Pääaine: Oppivat järjestelmät
Tarkastajat: Prof. Hannu Koivisto
Avainsanat: muutoksen havainnointi, Pearson divergenssi, älyverkot, sähkökulutuksen mallinnus, asiakaskäyttäväminen

Pyrkimyks teokkaampaan sähköjakeleun ja tuotantoarvioinnin hallintaan kulujen ja päästöjen hillitsemiseksi on johtanut innovaatioihin sähkövoimatekniikan alalla ja usean tieteenalan yhteistyön tuloksena on saatu älyverkot. Älyverkot tarjoavat moninaisia etuja perinteiseen sähköverkkoon verrattuna, esimerkiksi enemmän vaihtoehtoja ja joustavuutta sähkötuotannossa, paremmat asiakkaiden osallistumismahdollisuudet sekä sähköverkon lisääntymyt luotettavuus. Erityisesti asiakkailta kerätty sähkökulutuksen tuntimittausdata mahdollistaa tarkemman analyysin verkon tilasta, joka on taloudellisen sähkötuotannon kannalta keskeistä. Analyysin tarkkuutta on jatkuvasti ylläpidettävä ja mahdollisten muutosten havainnointi on tämän kannalta hyödyllistä.

PREFACE

This thesis was done for the project RESPONSE with funding from the Academy of Finland. I would like to thank all members of the project for letting me to be a part of it and also for providing new learning experiences. The electricity distribution is going through serious reformations and for me being able to use my knowledge and skill set in such a meaningful project was a priceless opportunity. Most notably thanks to Professor Hannu Koivisto for providing guidance along the way. Despite a few setbacks which I assume are common in research work this thesis was able to be completed. Also thanks to M.Sc. Antti Mutanen. It may not have always seemed so but I thoroughly appreciate all the comments and discussion we shared during this project. Last but definitely not least I would like to thank all relevant teaching staff at Tampere University of Technology for inspiring me to pursue this field at the time I was unsure about the future.

Mikko Nurmiranta
24th May 2017
CONTENTS

1. Introduction ......................................................... 1

2. Finnish electricity industry background ................................ 4
   2.1 Statistics ...................................................... 4
   2.2 Sources of change ........................................... 6
   2.3 Load research .................................................. 8

3. Change detection ..................................................... 10
   3.1 Cumulative sum ............................................... 11
   3.2 Divergence method ........................................... 12

4. Load forecasting ....................................................... 15
   4.1 Periodic autoregressive model ................................ 17
   4.2 Feedforward neural network model .......................... 18
   4.3 Forecasting performance ..................................... 20

5. Method validation ..................................................... 29
   5.1 Workshift change ............................................. 30
   5.2 Customer abnormality ......................................... 32
   5.3 Electric sauna stove installation ............................. 35
   5.4 No change ...................................................... 39
   5.5 Unknown change ............................................... 42

6. Conclusion ........................................................... 47

Bibliography ............................................................ 48
LIST OF FIGURES

2.1 Electricity consumption in Finland in 2015, 82.5 TWh in total. . . . 4
2.2 Electricity used by Finnish households in 2015, 20.9 TWh in total. . 5

4.1 Normalized hourly load dependency measured on outdoor tempera-
ture for a certain customer with electric heating. . . . . . . . . . . . . . 16
4.2 A feedforward network with one hidden layer with $D$ input units, $H$
hidden units and $C$ output units. [38] . . . . . . . . . . . . . . . . . . . . . 19
4.3 Illustration of the recursive next-day forecasting procedure. Only
the forecasts marked with red dots inside the ellipses are taken into
account when measuring method forecasting accuracy, the forecasts
on marked with blue dots are only used in the recursion process. . . . 21
4.4 Average forecasting RMSE for January 2010. . . . . . . . . . . . . . . . 24
4.5 Average hourly loads for each group of customers in January 2010. . . 24
4.6 Average forecasting RMSE for July 2010. . . . . . . . . . . . . . . . . . 25
4.7 Average hourly loads for each group of customers in July 2010. . . . 25
4.8 Forecasting performance on a direct heating customer. . . . . . . . . 26
4.9 Histogram and fitted normal distribution on forecasting errors. . . . 27
4.10 Forecasting performance on a customer without electric heating or
electric sauna stove. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
4.11 Histogram and fitted normal distribution on forecasting errors. . . . 28

5.1 Example weekly load profiles before and after the change. . . . . . . . . 30
5.2 Illustration of forecasting performance before the inserted change. . . 31
5.3 Illustration of forecasting performance after the inserted change. . . 31
5.4 PE score for the year, vertical line marks the day of inserted change. . 31
5.5 CUSUM score for the year, vertical line marks the day of inserted change. .......................... 32
5.6 Hourly measured loads from years 2009 and 2010. ............................................. 33
5.7 Forecasts and measurements around the moment of change. ............................. 33
5.8 Relative PE score using multivariate sample presentation and a 7-day time window. Vertical lines mark the beginning and the end of abnormal customer behavior. ............................... 34
5.9 Relative PE score using univariate sample presentation and a 7-day time window. Vertical lines mark the beginning and the end of abnormal customer behavior. ............................... 34
5.10 CUSUM test with a 7-day time window. Vertical lines mark the beginning and the end of abnormal customer behavior. .................................................. 35
5.11 Daily profiles for the average row house customers with and without electric sauna stove. ........................................... 36
5.12 Example of forecasting performance before the electric sauna stove installation without the additional noise. .................................................. 37
5.13 Example of forecasting performance after the electric sauna stove installation without the additional noise. .................................................. 37
5.14 Change detection with Pearson divergence. Vertical line marks the day of change. .................................................. 38
5.15 Change detection with CUSUM algorithm. Vertical line marks the day of change. .................................................. 39
5.16 Average daily loads for a certain customer from years 2009 and 2010. Notice the timing on level shifts. .................................................. 40
5.17 Average daily temperatures from years 2009 and 2010. ............................................. 40
5.18 Illustration of forecasting performance, from Thursday to Wednesday. High Saturday peak is captured although lacking in magnitude. .................................................. 41
5.19 Relative Pearson divergence with multivariate sample presentation, vertical lines mark the days of level shifts. No change can be determined. 41
5.20 Relative Pearson divergence with univariate sample presentation, vertical lines mark the days of level shifts. No change can be determined. 42

5.21 CUSUM test, vertical lines mark the days of level shifts. No change can be determined. 42

5.22 Comparison of average days of a certain direct electric heating customer from years 2009 and 2010. 43

5.23 Pearson divergence with multivariate sample presentation. 44

5.24 Pearson divergence with univariate sample presentation. 44

5.25 Cumulative sum test. The assumed changepoint is at much later date than with the Pearson test. 45
LIST OF TABLES

4.1 1-hour-ahead forecasting performance on training data. . . . . . . . . 23
4.2 1-hour-ahead forecasting performance on test data. . . . . . . . . . 23
4.3 Next day (up to 39-hour-ahead) forecasting performance on test data. 23
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMR</td>
<td>Automatic meter reading</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial neural network</td>
</tr>
<tr>
<td>ASHP</td>
<td>Air source heat pump</td>
</tr>
<tr>
<td>CUSUM</td>
<td>Cumulative sum</td>
</tr>
<tr>
<td>DG</td>
<td>Distributed generation</td>
</tr>
<tr>
<td>DR</td>
<td>Demand response</td>
</tr>
<tr>
<td>GSHP</td>
<td>Ground source heat pump</td>
</tr>
<tr>
<td>kWp</td>
<td>Kilowatt peak</td>
</tr>
<tr>
<td>MAE</td>
<td>Mean Absolute Error</td>
</tr>
<tr>
<td>PAR model</td>
<td>Periodic autoregressive model</td>
</tr>
<tr>
<td>PE divergence</td>
<td>Pearson divergence</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root-mean-square error</td>
</tr>
<tr>
<td>RuLSIF</td>
<td>Relative unconstrained least-squares information fitting</td>
</tr>
<tr>
<td>SI</td>
<td>Subspace identification</td>
</tr>
<tr>
<td>SSA</td>
<td>Singular-spectrum analysis</td>
</tr>
<tr>
<td>SOM</td>
<td>Self-organizing maps</td>
</tr>
<tr>
<td>STLF</td>
<td>Short-term load forecasting</td>
</tr>
<tr>
<td>SVM</td>
<td>Support vector machine</td>
</tr>
</tbody>
</table>
Climate change and rising production costs among other factors have driven governments, industries and consumers to reconsider their energy consumption and production policies and patterns, leading to the evolution of smart grids. Smart grid can be described as follows: "a smart grid is the use of sensors, communications, computational ability and control in some form to enhance the overall functionality of the electric power delivery system" [16]. Traditional power grids have no other function besides the one-way delivery of electricity from the power plants to end-users. Smart grids make two-way communication between the supplier and consumer possible, enabling stronger customer engagement on the electricity network through mechanisms such as distributed generation (DG) and demand response (DR). Customers in the possession of small wind turbines or photovoltaic panels previously meant for personal use only may feed excess electricity to the network, sell it at market price and gain financial benefits, resulting in potentially multitude of renewable electricity producers scattered around. This would not be possible in a traditional network incapable of dealing with reversed electricity flows.

Demand response has gathered increasing amount of attention in recent years. Demand response means that the demand side, that is the end-use customers, change their consuming patterns in response to some incentives from the supplying side, such as the price of electricity. For the supplying side demand response actions aim to keep the generation balanced and to reduce network congestion by shifting the demand from a time of high electricity usage to a time frame where the load levels are lower, also lowering costs in the process. The consumer side benefits too as the price of electricity tends to be cheaper at the time of little consumption. A simple example of demand response is the heating of domestic water. One kind of electricity tariff offers different rates for day and night use, the response to cheaper nighttime electricity is to schedule the heating of water to those hours. Nowadays tariffs that offer different rates on hourly basis are also available and whether users change their consumption behavior according to hourly changing price is an intriguing topic. A gamification approach to demand response was proposed as a part of the CITYOPT project [69]. A game-like tablet application notifies users of next day load peaks and asks for users to shift or reduce their load at the time. If users choose to comply
they will be awarded with points. The app was tested in 140 households in Nice, France and over 80% of these households participated in load reduction at peak times, despite minimal monetary gains. In large scale such participation rate would remarkably reduce the load at peak hours.

In this thesis the customer behavior refers to customer load with no further segmentation but the load can be analyzed more in-depth. In [72] customer load is stated to be formed of two components, customer behavioral determinants and physical determinants. Physical determinants show heavy correlation with climate variation and building design but low correlation with human behavior. Heating, cooling and lighting make up the most load in this category. Behavioral determinants consist of consuming habits by the use of household appliances and show little correlation with seasonal variation. Working hours, time of absence, the number of occupants for example influence this type of consumption. Load profiling of Irish customers was performed in [32]. It was found that the number of bedrooms in a dwelling and the number of occupants have a significant effect on which customer group the dwelling belongs to and it is clear these two factors are also related to the two consumption types stated above, number of rooms affects the electricity used for heating and lighting (physical determinant) and number of occupants affect the use of household appliances (behavioral determinants) for example. Regarding demand response, it is unlikely that the use of appliances that involve constant human presence such as entertainment or cooking appliances would be shifted to the period of low network usage, on the other hand appliances that are based on storing energy or maintaining a set temperature, such as freezers or an electric heating system for example may very well be subject to demand response actions. In particular heating of both living spaces and domestic water are of interest as they make up a major part of the residential electricity consumption in Finland as seen in the next chapter.

In Finland load prediction is commonly based on customer classes [41] and therefore detecting changes in customer behavior is essential, regardless of the cause. The classification done by using old measurement data no longer matches the post-change behavior, leading to errors in load prediction and network state estimation. Residential customers tend to have a lot more variation in their behavior as opposed to industrial customers, which makes change detection a challenging task as there is no certainty if the change is caused random by variation or a deliberate, possibly permanent change in human behavior.

The remainder of this thesis is organized as follows. Information on Finnish consumption statistics, future outlook and previously done load research is provided in Chapter 2. The means for customer behavior change detection are introduced
in Chapter 3. The models used for load forecasting are described in Chapter 4. The change detection methods in combination with the load models are tested in Chapter 5. Chapter 6 concludes the thesis.
2. FINNISH ELECTRICITY INDUSTRY BACKGROUND

This chapter provides background information on the current state of the Finnish electricity industry, consumption statistics and load research. Some possible causes of change are also discussed.

2.1 Statistics

In 2015 the total electricity consumption in Finland was 82.5 TWh, with industry using almost half of total as seen in Figure 2.1. In particular the pulp and paper industry is a major contributor to electricity consumption, consuming 48% of total electricity used by industry and 23% of all electricity used in 2015. [50]

![Electricity consumption in Finland in 2015, 82.5 TWh in total.](image)

Regarding residential consumption most electricity used goes towards heating spaces, although only slightly more than is used by household appliances, as shown in Figure 2.2. About heating, electricity is only the third largest solution measured by total energy consumed as district heating and wood are the largest although the
three solutions were separated by only 3 TWh in 2015. Energy consumed towards heating depends largely on building type. Detached houses make up the most of electricity consumed for heating. Apartment buildings are heated almost solely on district heat, on terraced houses district heating has roughly a 66% share and the remaining share is mostly electricity. Outdoor temperature obviously influences the energy consumption spent on heating, all energy forms included the energy used for space heating in 2015 was 8 TWh less than in 2010, corresponding to 16%. The year 2015 was also the warmest year recorded at the time. However the energy spent on electric heating has remained relatively stable, 10.1 TWh in 2010 versus 9.5 TWh in 2015, a 6.7% decrease. The regulation issued by the Council of European Union to phase out incandescent light bulbs [9] appears to be productive as the electricity spent on lighting has dropped by 30 percent from 2010 to 2015 and is also largely responsible for the drop in energy consumption of household appliances as the energy spent on cooking and other devices besides lighting has remained level over the years. [51]

![Pie chart showing electricity usage in Finnish households in 2015]

**Figure 2.2 Electricity used by Finnish households in 2015, 20.9 TWh in total.**

In November 2016 the Finnish government published their Energy and Climate Strategy [66]. The plan is, among other things, to completely phase out coal based energy by 2030 and increase the use of renewable energy sources to cover 50% of energy use. In 2015 hard coal accounted for 12.1% of heat and electricity production [52]. Potential for distributed solar power generation is mentioned but despite advances in efficiency and battery technology it will not be able to completely replace conventional power plants since the need for electricity is highest during cold winters but due to Finland’s northern location the amount of sunlight is minimal at those times.
2.2 Sources of change

As mentioned earlier in Finland solar power during winter is out of the question however the conditions in summertime are comparable to those of Central Europe where residential solar power generation is a lot more widespread. The production potential in southern Finland from March to October was estimated to be around 850–900 kWh/kWp [43], where kWp is kilowatt peak and stands for the peak power of a solar module under optimal conditions. For example residential customers of a certain network operator consumed approximately 3.4 kWh of electricity on average during March, 2010–October, 2010. While the reality is not as straightforward a simple calculation shows that for this period of the year a 2 kWp system could provide half of the electricity consumed by the average customer of this operator, some could even be completely self-sufficient. However as it stands for most people it is not financially feasible to install photovoltaic panels. Despite the prices of panels going down dramatically in recent years the payback time is still well over 10 years, depending on various factors such as the price of electricity and the size of the panel system [43]. Unlike some other countries the Finnish government currently does not offer any subsidies for installing solar panels.

Assuming the aforementioned government strategy actualizes radical changes are expected on the consumption side as the plan is to have 250 000 electric vehicles (including plug-in hybrids and hydrogen vehicles) in road use by the year 2030, which is a massive increase from 2015 when the total of registered electric road vehicles, including plug-in hybrids, was at 1952 vehicles [54]. Such an increase would without doubt show in the household electricity consumption especially at evenings as that is when the vehicles would most likely be charged. Norway is the leading nation regarding electric vehicles per capita with over 10 000 vehicles at the end of 2012 (including plug-in hybrids) and a study on the network impact from the charging of electric vehicles carried out by Seljeseth et. al [58] indicates that weaker grids may face problems caused by the charging. Smart charging is one example of demand response actions that could weaken the impact of charging on the network by shifting the charge timing to hours with less load e.g. the early morning.

In Finland thousands of households change their heating solutions every year [68]. Some of them choose electricity as the basis of their new heating system which is bound to increase their total electricity consumption and change their daily load curves. In particular oil as a heating solution has been losing users at an increasing rate but actually the net change for electricity was also negative during the years 2006-2009. For customers moving away from electric heating the load curves will change too of course. Note that the cited report only considers existing detached
2.2. Sources of change

houses. According to the Finnish Heat Pump Association [13] in 2015 around 50% of the newly built detached houses have ground source heat pumps (GSHPs) as the main heating system. Air source heat pumps (ASHPs) too have grown remarkably in popularity with tens of thousands of units sold every year since 2004. While GSHPs may only affect the consumption during colder time of the year ASHPs can also be used for cooling and so have an effect on summer time consumption as well. The effect of installing ASHP in detached house with direct electric heating was investigated in a small study consisting of 166 houses from around Finland [49]. The installation of an ASHP reduced the annual electricity consumption by 1 000–5 000 kWh depending on the consumption before the installation. Most of the ASHPs in the study were manufactured during the years 2005–2008 and it is likely that newer models have improvements in efficiency and therefore bigger savings could be achieved.

As stated in Chapter 1 the number of occupants seems to be a factor in customer loads. As children grow up and eventually leave to live on their own the use of everyday appliances reduced. Less laundry may increase the time between washing machine usage. Personal computers and entertainment devices are no longer used at evenings at the same intensity. Time and energy spent on cooking will go down. Less warm water is used hence the electricity spent on heating of water will decrease and so on. In general large number of occupants means bigger houses. Once they are no longer actively used the heating in the empty rooms can be turned down by a few degrees if possible which would lead to some savings on houses with electric heating.

Home energy management systems (HEMS) are a growing market triggered by the call for energy conservation. By automatically controlling the temperature in each room individually while taking into account the occupants’ life style customers with electric heating can avoid unnecessary heating and experience great savings in electricity bills. As practically every house in Finland is equipped with a smart meter the measurement data from a house with a HEMS should provide even more accurate information on the customer lifestyle, leading to better load profiling. Smart meters are also required to have support for DR thus the control of the heating system can be performed in a way that is the cheapest without giving up comfort. In Finland almost half of all detached houses are heated by electricity [53] so significant savings on a national level could be achieved with a large scale implementation of HEMS.
2.3 Load research

In Finland customer load has been subject to extensive research since 1983 when a collaboration between 40 electricity utilities began [59]. Eventually in 1992 46 customer classes based mainly on load measurements, customer type, building type and heating solution were released. The measurement data at the time was expensive to collect and therefore data from only 667 customers was used to produce the classes, however they were seen appropriate enough for network planning and load prediction and even today some of these profiles are in use. In 20 years consumption patterns have changed radically and are likely to keep changing in the future therefore these profiles lose their accuracy. The advent of smart grids has opened up new possibilities for refining customer load profiles and other techniques for more accurate network state estimation.

Smart meters that allow Automatic Meter Readings (AMR) have already been common in Finland for years and since 2014 Distribution Network Operators (DNOs) have been required by law to equip at least 80% of their customers with such meters. DNOs are also required to share meter readings to their customers, save the data for at least six years and have support for demand response actions [65]. The availability of measurement data has been taken advantage of in various load studies. Customer load profiling using ISODATA algorithm was done in [41] while Self-Organizing Maps (SOM) combined with K-means and hierarchical clustering algorithms to create customer profiles was used in [48]. SOM and K-means was also used in [56], also taking into account socioeconomical factors and building types. Each of the mentioned profiling methods show improvement over the profiles used by the respective DNOs. The effect of load profile updating was studied in [40]. New load profiles were made in two ways, by keeping the original customer classification intact and using AMR measurements to update the class specific load profiles, and by doing a complete re-classification by clustering. In both cases the forecasts based on new load profiles were better than the originals. Several methods for short-term load forecasting (STLF) were compared in [29]. A neural network model came out as the most accurate however a model based on load profiles obtained by clustering and a model combining Kalman-filter predictor and physically based components were not far behind with each method having their own strengths. A model selection scheme for neural networks based on genetic algorithm was studied in [42]. The tedious task of choosing appropriate input variables for a neural network architecture in short-term load forecasting application was alleviated by using the proposed method. Change detection has been studied as well. Clustering methods to detect between-years change in customer behavior was used in [5]. Customers were first clustered into groups and weekly profiles were then constructed for each cluster. For
each customer Fuzzy C-Means is used to obtain a degree of membership to each cluster using weekly load measurements. The degree of membership to each cluster is then used as a basis for change: if the sum of absolute differences between weekly memberships for 5 or more consecutive weeks exceed a certain threshold, a change in behavior is claimed.

CLIC Innovation Ltd. is a cluster of companies, research institutions and universities founded with the purpose to find solutions in bioeconomy, cleantech and energy [6]. Among their many research projects they have programs that heavily related to Smart Grids and load research, such as the already concluded Smart Grids and Energy Markets (SGEM) [8] and the ongoing project RESPONSE - Improved Modelling of Electric Loads for Enabling Demand Response by Applying Physical and Data-Driven Models. The Academy of Finland provides funding for the project RESPONSE which has the following objectives [7]:

- Developing improved models for short term load forecasting and optimization of electric load dynamic responses to load control actions and weather variations.

- Analyzing and developing criteria for comparing the performance of the short term load forecasting.

- Enhancing utilization of smart metering data and other spatial information in updating and verification of load response models.

- Proving that the developed load modelling methods improve the state estimation and network analysis accuracy.

Accurate models of demand responses help managing the electricity balance in the network while also reducing costs and emissions. For models to remain accurate changes must be detected and appropriately dealt with. This thesis is a part of project RESPONSE.
3. CHANGE DETECTION

Change detection is a problem of noticing the instance of time when the properties of the observed signal before and after the point differ. The methods for solving the problem can be classified in two categories, real-time and retrospective. Real-time or online methods are useful in applications that require immediate response, such as robotics and biometrics [1], as the observations are processed right as they are received. Retrospective or offline methods make up for the longer detection time by being generally more robust and accurate [31]. Change point detection has been subject to exhaustive research for decades due to wide range of applicable fields, such as fault detection, outlier detection, speech signal segmentation, intrusion detection, econometrics, climate change detection and medical diagnostics just to mention a few. As a result several approaches have been developed. An online change detection method for customer loads using Bayesian framework was proposed in [62]. Customer load is modeled with multivariate Gaussian distribution and the load profile is updated after each day with fresh load measurements. The probability of change happening at time instant $t$ is then evaluated and if change is detected the load profile is reset and rebuilt. The method was tested on 32 customers in the UK for 180 days. The results showed several kinds of behavior, some customers showed changes every few days, some did not change even once during the trial period. However no reference points were used for the load profiles. The load profiles evolved over time starting from a standard multivariate Gaussian distribution so whether the customer behavior changed from the year before for example was not studied.

A method to detect change points in time series based on subspace identification (SI) was proposed in [27]. The columns of an extended observability matrix constructed from the matrices of a linear state-space model span a certain subspace. A change point can be detected by calculating the distance between the two subspaces constructed from two subsequent time intervals. Both online and offline versions are presented, along with a form that takes input sequences into account thus making it method worth considering for load measurements as well. The method compared favorably to another subspace method based on singular-spectrum analysis (SSA), in which the distance between a subspace spanned by the eigenvectors of a so-called lag-covariance matrix and measurement vectors serves as the basis for change de-
3.1 Cumulative sum

Researchers at Binghamton University proposed a method based on SSA for detecting disturbances in power grids [71]. Using simulation data the SSA based method performed better than another disturbance detection algorithm [30] in terms of detection speed and accuracy.

To detect changes from load measurements the nature of change must be defined first. Changes within the year caused by external factors such as temperature or specific type of day are of no interest here, instead we are looking for changes that result in a different state in the long term, possibly permanently, and are caused by unpredictable human interaction, for example a change in heating solution or in work shifts. To reduce the effect of external variables a model which takes these into account is fit for the customer, the future load is predicted using said model and the forecasts compared with actual measurements. This approach allows us to tie together load forecasting and behavior monitoring. If the relationship between the measurements and forecasts begins to change it is therefore reasonable to assume a change in customer behavior has happened. To monitor that relationship we use two methods, cumulative sum and relative Pearson divergence.

3.1 Cumulative sum

Cumulative sum (CUSUM) chart method is a technique of statistical quality control introduced by E.S. Page [44]. Page was interested in the change in quality of a production output of a continuous manufacturing process. By sampling products at regular intervals and measuring certain attributes to determine their relation to expected values it should be possible to detect if the process is out of control and the quality of products will get worse if no action is taken. Page himself presented several rules and over time several variations of the algorithm have been developed (see e.g. [2,36,47,61]), in this thesis a rule introduced by Page is followed. The rule can be written as

$$S_0 = 0$$
$$S_t = \max(0, S_{t-1} + x_t),$$

(3.1)

where $x_t$ is a score assigned to each observation taken at the time $t$. The score is in this thesis is calculated on daily basis and the score function used resembles one already proposed for load forecasting environment. A medium voltage distribution network state estimation tool including a nonlinear autoregressive exogenous (NARX) model for load forecasting and a CUSUM algorithm is proposed in [21]. In the paper the score function is based on error percentages, however in this thesis the mean absolute errors (MAE) are used and so the score function for day $t$ is
\[ x_i = MAE_t - k \]

\[ MAE_t = \frac{1}{24w} \sum_{i=t-w}^{t} \sum_{j=1}^{24} |y_{ij} - \hat{y}_{ij}|, \quad (3.2) \]

where \( y_{ij} \) and \( \hat{y}_{ij} \) correspond to the load measurement and forecast at hour \( j \) on day \( i \), \( w \) is the window length in days and \( k \) is a user defined allowance term. In words the score function is the difference between mean absolute hourly error within the time window of \( w \) days and the allowance term \( k \). As can be seen from the formula the CUSUM score increases when the MAE is larger than \( k \) and decreases when the MAE is smaller than \( k \), with zero being the lower bound.

Especially in an online setting attention needs to be paid on the time window from which the MAE is calculated. The window shouldn’t be too short in order to avoid false alarms caused by anomalies but on the other hand it shouldn’t be too long either as the change should be noticed as early as possible.

### 3.2 Divergence method

One way to measure the similarity of sets of data is to compare the underlying probability distributions. However to calculate the dissimilarity one must know or estimate the probability densities. Density estimation is a hard problem (see [67]), poor estimation of individual densities will lead to poor performance in dissimilarity calculation as well, therefore it is convenient to estimate the density ratio directly without the need for individual densities. This also frees us from making any assumptions about the distribution of the samples, that is the load measurements and forecasts in this thesis. Let \( P \) and \( P' \) be two probability distributions of samples \( \mathcal{X} = \{x_i\}_{i=1}^{n} \) and \( \mathcal{X}' = \{x'_{ij}\}_{j=1}^{n'} \) with \( p(x) \) and \( p'(x) \) their respective density functions. Let us model the density ratio \( \frac{p(x)}{p'(x)} \) with the kernel model

\[ g(x; \theta) = \sum_{i=1}^{n} \theta_i K(x, x_i) \quad (3.3) \]

where \( \theta \) is a parameter vector estimated from data and \( K(x, x') = \exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) \) a Gaussian kernel function. The kernel width \( \sigma \) is a user defined parameter.

The parameters of model 3.3 are learned by a method called relative unconstrained least-squares information fitting (RuLSIF). The method outperformed both SI and SSA and some other methods in change detection [31] which is why this method was chosen for this thesis. For more on the theory behind the density ratio estimation
3.2. Divergence method

the reader is referred to [26,63,64,70]. A paper [74] using this method for detecting photovoltaic panel installations may also be of interest as it is in fact also related to detecting changes from customer load measurements.

The dissimilarity measure used is the relative Pearson divergence, defined as

$$P_{\alpha}(P||P') = \int p'_\alpha(x)\left(\frac{p(x)}{p'_\alpha(x)} - 1\right)^2dx$$  \hspace{1cm} (3.4)

where $p'_\alpha(x) = \alpha p(x) + (1 - \alpha)p'(x)$ and $0 \leq \alpha < 1$. Using Pearson divergence leads to a squared loss function which has an analytical solution and is therefore more efficient and robust than estimating e.g. Kullback-Leibler divergence that leads to a nonlinear log-loss function which needs to be solved iteratively [64]. Let us define $r_\alpha(x) = \frac{p(x)}{\alpha p(x) + (1 - \alpha)p'(x)}$. Using $\alpha$ as a smoothing parameter leads to better accuracy in the case the density ratio $p(x)/p'(x)$ is unbounded [70]. To estimate $r_\alpha(x)$ we use the model 3.3 with the parameter vector $\theta$ learned by minimizing the following squared-loss function

$$J(x) = \frac{1}{2} \int p'_\alpha(x)\left(r_\alpha - g(x; \theta)\right)^2dx$$

$$= \frac{1}{2} \int p'_\alpha(x)r_\alpha^2(x)dx - \int p(x)r_\alpha(x)g(x; \theta)dx$$

$$+ \frac{\alpha}{2} \int p(x)g(x; \theta)^2dx + \frac{1 - \alpha}{2} \int p'(x)g(x; \theta)^2dx. \hspace{1cm} (3.5)$$

Ignoring the constant and approximating integrals with sample averages the problem is reduced to

$$\min_{\theta} \frac{1}{2} \theta^T H \theta - h^T \theta + \frac{\lambda}{2} \theta^T \theta \hspace{1cm} (3.6)$$

where $\lambda$ is a user defined regularization parameter, $h$ and $H$ $n$-dimensional vector and $n \times n$-matrix respectively defined as

$$h_i = \frac{1}{n} \sum_{i=1}^{n} K(x_i, x_i)$$

$$H_{i,j} = \frac{\alpha}{n} \sum_{i=1}^{n} K(x_i, x_i)K(x_i, x_j) + \frac{1 - \alpha}{n} \sum_{j=1}^{n} K(x'_j, x_i)K(x'_j, x_j). \hspace{1cm} (3.7)$$
The problem has an analytical solution

$$\theta = (H + \lambda I_n)^{-1}h. \quad (3.8)$$

With the parameters for density ratio estimator acquired the relative Pearson divergence between $P$ and $P'$ can be estimated from samples $\mathcal{X}$ and $\mathcal{X}'$ with

$$\hat{PE}_\alpha(P||P') = -\frac{\alpha}{2n} \sum_{i=1}^{n} g(x_i)^2 - \frac{1-\alpha}{2n} \sum_{j=1}^{n} g(x'_j)^2 + \frac{1}{n} \sum_{i=1}^{n} g(x_i) - \frac{1}{2}. \quad (3.9)$$

Note that relative Pearson divergence is not symmetric (therefore it is not a metric), i.e. $PE_\alpha(P||P') \neq PE_\alpha(P'||P)$. To gain a symmetrized PE score one can add the asymmetric measures together

$$PE_{\alpha, sym} = PE_\alpha(P||P') + PE_\alpha(P'||P). \quad (3.10)$$

In this thesis the symmetric version of relative PE score is used in all occasions involving the Pearson divergence.

A word on sample presentation. If load during the day is described with a multivariate random variable samples are presented as vectors i.e. $x_i \in \mathbb{R}^d$, where $d$ is the resolution. The data used in this thesis is sampled at one hour intervals which means $d = 24$. Another approach is to consider the hourly load as a univariate random variable in which case information about the time of day is lost but in return effectively more samples are obtained. For example consider a 7-day time window. In multivariate case there 7 samples for a 24-variate distribution, while for the univariate case there are $7 \times 24 = 168$ samples. The multivariate approach is preferred in this thesis unless explicitly stated otherwise as the intra-day behavior is of interest. However in the case of inconclusive results the univariate approach is applied as well.
4. LOAD FORECASTING

Short-term load forecasting (STLF) refers to forecasting electric loads from one hour up to one-week ahead and is an extensively studied subject with several different approaches, such as time series analysis [10, 46], support vector machines (SVMs) [3, 12, 23, 55], artificial neural networks (ANNs) [39, 42, 57], fuzzy models [45, 73] and physically based models [28]. So far most STLF studies have focused on forecasting aggregated loads, and for a good reason. Substation-level forecasts are accurate enough for proper network state estimation. Residential customers have too much variation in their daily activities in order to form truly reliable forecasts. The effect of aggregation level to forecasting performance was explored in [60], including residential and small business customers. Predicting human behavior is hard enough as it is but the sampling interval of the data poses another challenge. Forecasting residential loads with a Kalman filter based method and various sampling rates was studied in [17]. Halving the sampling rate from 1 hour to 30 minutes also roughly halved the error rate for 1-hour ahead forecasts. Further reduction to 15 minutes did not bring about such dramatic increase in accuracy. The data used in this thesis is sampled at 1-hour intervals.

Generally the load is modeled by using two types of variables, time related variables and weather variables. Customer loads tend to be of seasonal nature since customers have their patterns that show up repeatedly in their load profiles if not daily then at least on a weekly basis. For example office buildings tend start up in the morning, hold relatively steady consumption through the afternoon and shutdown in the evening, resulting in fairly smooth behavior. For residential customers things are not as straightforward as households can be wildly different but stereotypically there would be morning peak before leaving for work or school, downtime in the afternoon and another, bigger peak in the evening. As such the time of day needs to be taken account for accurate forecasting. Consumption habits depend on the type of day as well. Customers have different consumption behavior on weekends and holidays compared to normal working days. Within-year seasonal variation can be taken into account by including a month variable or the length of day for example. As time in daily life is perceived as a categorical attribute in the form of calendars or clocks, these aforementioned variables can be modeled as such too through binary coding.
To include the cyclic nature of these time variables a sine-cosine transformation can be used as well. Due to the seasonality the measurement time series tend to show reasonable autocorrelation at lags that are multiples of 24 which is why past load past measurements up-to one week may be used as an input variable.

Depending on the heating solution the outdoor temperature can have a significant effect on electricity consumption. Figure 4.1 shows the hourly load dependency on the outdoor temperature in 2009 for a certain customer with direct electric heating. As can be seen for this customer the electricity load is higher when temperatures go below 0°C, indicating that the heating is turned on. Customers without electric heating may also exhibit similar behavior to a smaller extent. Cold weather and the infamous darkness of Finnish winters drive people to spend more time inside, increasing the use of lighting and household appliances. On the other hand commercial customers in an office environment may show similar behavior on the warmer side of the temperatures due to increased cooling and ventilation. The temperature affects with a delay which is why past temperature measurements are worth considering as input variables in addition to the temperature at the forecasted hour. The relationship between the customer load and the outdoor temperature may also be non-linear which is why machine learning techniques such as ANNs and SVMs are popular choices for load forecasting since they are both universal approximators meaning they are capable of approximating any continuous function with an arbitrary level of accuracy [19, 24].

![Normalized hourly load dependency measured on outdoor temperature for a certain customer with electric heating.](image_url)
In this thesis a periodic autoregressive (PAR) model, following the presentation given in [12], and feedforward neural network model are used. Two different models are used because based on testing the slightly more accurate PAR model also appears to be a lot more rigid with regards to changing conditions than the neural network model, therefore to bring out the differences between CUSUM and Pearson divergence change detection methods another model is used depending on the case.

4.1 Periodic autoregressive model

The strong seasonality induced by the customer rhythm needs to be taken into account when modeling the customer load. In time series analysis several methods for dealing with seasonality has been developed (see [15] and the references therein). Because the daily cycles are very much relevant to load forecasting deseasonalization by preprocessing the data, that is the removal of seasonal variation, is out of the question. Another way is to include the seasonal variation in the time series model. Perhaps the best known and most used model for seasonal time series is the seasonal autoregressive integrated moving average (seasonal ARIMA, SARIMA [4]) model, used for STLF e.g. in [46]. As stated the time series model used in this thesis is a PAR model [15]. A PAR model of order \( p \) can be written as

\[
y_t = C_s + \phi_{1,s}y_{t-1} + \ldots + \phi_{p,s}y_{t-p} + \epsilon_t,
\]

where \( s \) is the number of seasons, \( C_s \) is a seasonally varying constant, \( \phi_{i,s} \) are AR model parameters and \( \epsilon_t \) is assumed to be white noise. To put it simply the seasonality is taken into account by fitting an autoregressive model separately for each season.

Using PAR model with exogenous variables enables us to capture the daily cycle in the load. The number of "seasons" corresponds to the number of hours in a day, \( s = 24 \), so in fact the load model is a collection of 24 different models. For the input the past 24 hourly load measurements are used, along with 168 hour lagged load. Calendar variables day of week and month are modeled as binary dummy variables \( w_d \) and \( m_d \) with eves and public holidays modeled as Saturdays and Sundays respectively. The hourly temperature \( T_{h,d} \) is taken into account with \( \mathbf{T}_{h,d} = [CR_{h,d}, HR_{h,d}, XHR_{h,d}]^T \), where \( CR_{h,d} = \max(0, T_{h,d} - 20{^\circ\text{C}}) \), \( HR_{h,d} = \max(0, 17{^\circ\text{C}} - T_{h,d}) \) and \( XHR_{h,d} = \max(0, -T_{h,d}) \). The hourly load \( y \) for day \( d \) is
then modeled as follows:

\[
\begin{align*}
y_{1,d} &= C_1 + \phi_{1,1} y_{24,d-1} + \phi_{1,2} y_{23,d-1} + \cdots + \phi_{1,24} y_{1,d-1} + \phi_{1,25} y_{1,d-7} \cdots + \alpha_{1}^T \mathbf{w}_d + \beta_{1}^T \mathbf{m}_d + \gamma_{1,d}^T \mathbf{T}_{1,d} + \epsilon_{1,d} \\
y_{2,d} &= C_2 + \phi_{2,1} y_{1,d} + \phi_{2,2} y_{24,d-1} + \cdots + \phi_{2,24} y_{2,d-1} + \phi_{2,25} y_{2,d-7} \cdots + \alpha_{2}^T \mathbf{w}_d + \beta_{2}^T \mathbf{m}_d + \gamma_{2,d}^T \mathbf{T}_{2,d} + \epsilon_{2,d} \\
\vdots \\
y_{24,d} &= C_{24} + \phi_{24,1} y_{23,d} + \phi_{24,2} y_{22,d} + \cdots + \phi_{24,24} y_{24,d-1} + \phi_{25,24} y_{24,d-7} \cdots + \alpha_{24}^T \mathbf{w}_d + \beta_{24}^T \mathbf{m}_d + \gamma_{24,d}^T \mathbf{T}_{24,d} + \epsilon_{24,d},
\end{align*}
\] (4.2)

Worth noting is that to avoid exact collinearity between variables one of the components in each set of binary vectors must be dropped [25]. This means that day of the week is modeled as \( \mathbf{w}_d \in \{0, 1\}^8 \) and month as \( \mathbf{m}_d \in \{0, 1\}^{11} \).

Because the model for hourly loads is linear the parameters \( C_i, \phi_i, j, \alpha_i, \beta_i \) and \( \gamma_{i,j} \) can be solved using ordinary least-squares techniques, also allowing for regularization techniques such as lasso or ridge regularization. However, these are not used in thesis.

### 4.2 Feedforward neural network model

Artificial neural networks mimic the human brain with interconnected group of neurons used to learn relationships between inputs and outputs from experience. The neurons in a feedforward network (also called multilayer perceptron) are organized in input, hidden and output layers and are connected in such a way that no cycles are formed. The number of hidden layers can vary, a feedforward network with one hidden layer is illustrated in Figure 4.2.
Computations within the network happen in the neurons. Consider the feedforward network pictured in Figure 4.2. For example the output of the hidden neuron \( H \) with an activation function (also called a transfer function) \( f_H \) is \( z_{nH} = f_H(\sum_{j=1}^{D} x_{nj}v_{jH}) \). One can see that a nested structure is formed from input to output, for the network in Figure 4.2 the output \( y_{nk} \) with the activation function \( g_k \) is \( y_{nk} = g_k(\sum_{j=1}^{H} w_{jk}f_j(\sum_{i=1}^{D} x_{ni}v_{ij})) \). Common choices for the activation function in the hidden layer are sigmoids, for example log-sigmoid \( \text{logsig}(t) = \frac{2}{1+e^{-t}} \) and hyperbolic tangent \( \text{tanh}(t) = \frac{e^t - e^{-t}}{1+e^{-t}} \). Depending on application the output layer can have sigmoidal activation function, especially used in classification tasks, or a linear function as is typical for regression cases.

The training of a feedforward network is based on minimizing a cost function, for example the sum of squared errors (SSE) \( J(v) = \frac{1}{2} \sum_{i=1}^{d} (t_i - y_i)^2 = \frac{1}{2} ||t - y|| \), where \( t \) and \( y \) are the target and network output vectors of length \( d \) and \( v \) is a vector of all weights in the network. The network weights are updated with a steepest descent backpropagation algorithm, the most basic form of it is summarized in Algorithm 1 for a network with one hidden layer [34]. In short first an input signal is sent forward to calculate network output values (forward propagation), then the error signal is sent back through the network to generate delta values used to update the network weights. The minimum of the cost function is found when the derivative is zero but due to the highly nonlinear structure an analytical solution is not possible and one must settle for gradient descent techniques, in Algorithm 1 the learning rate parameter \( \eta \) is used to control the distance taken to the direction of the negative gradient at each iteration step, short distance leads to slower learning while too
long distance does not converge at all. One drawback of gradient descent methods is that in the case of nonconvex error surfaces the solution is likely be to a local optimum. More accurate and robust alternatives for the standard steepest descent backpropagation have been developed over the years, such as the scaled conjugate gradient method [35], the Levenberg-Marquardt algorithm which is often considered the standard method, [18] and the Bayesian regularization algorithm [14, 33]. For more in-depth discussion on the properties and training of neural networks the reader is referred to e.g. [11, 20, 22, 34, 38].

**Data:** Training samples \( \{(x_i, t_i)\}_{i=1}^N \)

**Result:** Network weights \( v \) that minimize the cost function

Initialize network weights \( v \);

while termination condition is not met do

for each training sample \( x_i \) do

    calculate the output \( o_j \) of every unit \( j \) in the network;

    for every output unit \( k \), calculate its error term \( \delta_k = o_k(1 - o_k)(t_{ik} - o_k) \);

    for every hidden unit \( h \), calculate its error term

\[
\delta_h = o_h(1 - o_h) \sum_{k \in \text{outputs}} v_{kh}\delta_k;
\]

    update each network weight \( v_{ji} \leftarrow v_{ji} + \Delta v_{ji} \), where \( \Delta v_{ji} = \eta \delta_j x_{ji} \)

end

end

**Algorithm 1:** Backpropagation algorithm

The feedforward network used in this thesis consists of one hidden layer with 15 neurons and one output, the load at next hour \( y_{h,d} \). The inputs for the network are the past 24 load measurements, outdoor temperature at the target hour as well as 8, 16, 24, and 48 hour lagged temperatures. In addition the hour of day, day of the week, month and day of the month are used as inputs but instead of binary coding a sine-cosine transformation is used. For example the hour of day using this transformation is \( d_h = [\sin((h-1)\frac{2\pi}{24}), \cos((h-1)\frac{2\pi}{24})] \), \( h \in \{1, ..., 24\} \). The activation functions for the hidden neurons and output unit are hyperbolic tangents and linear, respectively. The network is constructed with Matlab Neural Network Toolbox using Levenberg-Marquardt backpropagation as the learning algorithm and early stopping as the method to control overfitting.

### 4.3 Forecasting performance

The forecasting performance of the specified models is examined in this section. First the 1-hour-ahead forecasting performance is examined on training and test data. The long horizon forecasting is referred to as next-day forecasting and it
is performed in the following way. To keep matters consistent within the project RESPONSE the forecasts covering day $d$ are made at 9:00 on day $d - 1$ using the measurement data available at that time, including the load measurement at 9:00. This means that the forecast horizon is 39 hours. The recursive forecasting procedure is illustrated in Figure 4.3. The models are trained to predict the load at the next hour so to cover the forecast horizon the forecasts are made recursively with past model outputs replacing the past measurements over time. This and the long horizon mean that towards the end of the forecasting horizon the models have only one load measurement to rely on from 168 hours before, therefore model accuracy is critical.

![Figure 4.3 Illustration of the recursive next-day forecasting procedure. Only the forecasts marked with red dots inside the ellipses are taken into account when measuring method forecasting accuracy, the forecasts on marked with blue dots are only used in the recursion process.](image)

The AMR data used comes from a Finnish distribution operator Koillis-Satakunnan Sähkö Oy, from years 2009 and 2010. In addition to load measurements hourly temperature measurements from the same area are used, in 2009 the temperatures ranged from $-24.7^\circ C$ to $27.6^\circ C$ and in 2010 from $-29.5^\circ C$ to $32.8^\circ C$, which means the models have to deal with temperatures outside the identification range. The 2009 data is used for training and 2010 data for testing. The load measurement data is standardized to zero mean and unit variance, as is the temperature data when using the neural network model. This evaluation scheme is not ideal since the temperature data are actual measurements, in a real world situation weather forecasts would have to be used instead.
4.3. Forecasting performance

The forecasts are performed on customers from three groups based on the classification information received from the operator: "housing + direct electric heating", "row house/apartment, no electric heating, no electric sauna stove" and "row house/apartment, no electric heating, electric sauna stove", referred to as group 1, 2 and 3 respectively. The customers in these groups are expected to behave differently with respect to weather conditions and daily activity. The forecasts are made for 50 individual customers selected from each class. In addition to individual customers the forecasts are also made for aggregated loads, first for a group consisting of the 50 customers selected before, then for a group consisting of another 150 individuals in addition to the 50 customers selected before, so in total a group of 200 customers. The individual loads are visually inspected for obvious outliers, missing values and such. The performance measure used is root-mean-square error (RMSE), defined

\[
RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2},
\]

where \(y_t\) and \(\hat{y}_t\) are hourly measurements and forecasts respectively at hour \(t\) and \(n\) is the number of hours in the test set of data. The forecasts are made in Matlab 2016a on a PC with a 3.20GHz Intel Core i7 960 CPU and 24 GB of RAM, the results are compiled in Tables 4.1, 4.2 and 4.3. Note that the RMSE are calculated by using the load values that are normalized by using the procedure explained earlier in this section. The PAR model is generally more accurate except on the 1-hour-ahead forecasts on test data. Since the neural network training is susceptible to getting stuck at local optimum the starting values are very important. The initial values are random so the networks were trained again in the case of suspiciously large RMSE values in comparison to the PAR model. This leads us to a key difference between the methods: to make sure the neural network is performing as well as possible several training sessions are needed. Unless one is interested in forecasting the load only a few hours ahead there is little reason to use the neural network over the PAR model. Of course this remark only applies to the specific network structure and input variables chosen in this thesis.
4.3. Forecasting performance

\textbf{Table 4.1} 1-hour-ahead forecasting performance on training data.

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th></th>
<th>Group 2</th>
<th></th>
<th>Group 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PAR</td>
<td>NN</td>
<td>PAR</td>
<td>NN</td>
<td>PAR</td>
<td>NN</td>
</tr>
<tr>
<td>Individuals: Best</td>
<td>0.2017</td>
<td>0.2142</td>
<td>0.4357</td>
<td>0.2991</td>
<td>0.5255</td>
<td>0.3601</td>
</tr>
<tr>
<td>Individuals: Average</td>
<td>0.4999</td>
<td>0.5288</td>
<td>0.6950</td>
<td>0.7300</td>
<td>0.6808</td>
<td>0.7287</td>
</tr>
<tr>
<td>Individuals: Worst</td>
<td>0.6891</td>
<td>0.7493</td>
<td>0.8347</td>
<td>0.9017</td>
<td>0.8727</td>
<td>0.9151</td>
</tr>
<tr>
<td>Aggregate of 50</td>
<td>0.1117</td>
<td>0.1241</td>
<td>0.3091</td>
<td>0.3255</td>
<td>0.2857</td>
<td>0.3024</td>
</tr>
<tr>
<td>Aggregate of 200</td>
<td>0.0663</td>
<td>0.0725</td>
<td>0.1787</td>
<td>0.1922</td>
<td>0.1721</td>
<td>0.1838</td>
</tr>
</tbody>
</table>

\textbf{Table 4.2} 1-hour-ahead forecasting performance on test data.

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th></th>
<th>Group 2</th>
<th></th>
<th>Group 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PAR</td>
<td>NN</td>
<td>PAR</td>
<td>NN</td>
<td>PAR</td>
<td>NN</td>
</tr>
<tr>
<td>Individuals: Best</td>
<td>0.2730</td>
<td>0.2675</td>
<td>0.5330</td>
<td>0.4043</td>
<td>0.6427</td>
<td>0.4468</td>
</tr>
<tr>
<td>Individuals: Average</td>
<td>0.6205</td>
<td>0.6008</td>
<td>0.8917</td>
<td>0.8041</td>
<td>0.8321</td>
<td>0.7813</td>
</tr>
<tr>
<td>Individuals: Worst</td>
<td>0.9665</td>
<td>0.9273</td>
<td>2.1275</td>
<td>2.1000</td>
<td>1.1165</td>
<td>1.0413</td>
</tr>
<tr>
<td>Aggregate of 50</td>
<td>0.1393</td>
<td>0.1499</td>
<td>0.3535</td>
<td>0.3509</td>
<td>0.3268</td>
<td>0.3191</td>
</tr>
<tr>
<td>Aggregate of 200</td>
<td>0.0813</td>
<td>0.0882</td>
<td>0.2099</td>
<td>0.2067</td>
<td>0.2035</td>
<td>0.2003</td>
</tr>
</tbody>
</table>

The neural network shows its generalization strength on 1-hour-ahead forecasts on test data with lower RMSE values across the board. On the long horizon forecasts, however, the PAR model comes out on as the more accurate. The RMSE values for 1-hour-ahead PAR model forecasts are 9.7\%, 4.2\% and 3.8\% lower on average respectively for each group in comparison to the longer horizon forecasts.

\textbf{Table 4.3} Next day (up to 39-hour-ahead) forecasting performance on test data.

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th></th>
<th>Group 2</th>
<th></th>
<th>Group 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PAR</td>
<td>NN</td>
<td>PAR</td>
<td>NN</td>
<td>PAR</td>
<td>NN</td>
</tr>
<tr>
<td>Individuals: Best</td>
<td>0.3190</td>
<td>0.3950</td>
<td>0.6575</td>
<td>0.6611</td>
<td>0.6837</td>
<td>0.6747</td>
</tr>
<tr>
<td>Individuals: Average</td>
<td>0.6850</td>
<td>0.7042</td>
<td>0.9066</td>
<td>0.9438</td>
<td>0.8636</td>
<td>0.8658</td>
</tr>
<tr>
<td>Individuals: Worst</td>
<td>1.0572</td>
<td>1.0006</td>
<td>1.3332</td>
<td>1.3436</td>
<td>1.1201</td>
<td>1.0955</td>
</tr>
<tr>
<td>Aggregate of 50</td>
<td>0.1727</td>
<td>0.1876</td>
<td>0.3992</td>
<td>0.4404</td>
<td>0.3528</td>
<td>0.3568</td>
</tr>
<tr>
<td>Aggregate of 200</td>
<td>0.1279</td>
<td>0.1411</td>
<td>0.2726</td>
<td>0.2807</td>
<td>0.2526</td>
<td>0.2526</td>
</tr>
</tbody>
</table>

Now let us look at error on different hours of the day. We will limit the discussion on the coldest and warmest months of the year 2010 which were January and July respectively. The average forecasting RMSE of the PAR model on each hour of
the day when considering the individual customers in each group are illustrated in Figures 4.4 and 4.6.

![Figure 4.4 Average forecasting RMSE for January 2010.](image)

Group 1 consists of customers with direct electric heating and as such the consumption is heavily influenced by the variation on the colder side of temperatures. It can be seen that forecasting the customers in this group suffers the most from the longer forecasting horizon as the gap between 1-hour-ahead and next-day forecasts is wider than for the other groups on all hours of the day. However for group 1 the RMSE in general is lower than the other groups except for the late night/early morning hours. The RMSE on those hours are rather low in general, perhaps due to the predictably low activity. Customers in group 2 have electric sauna stoves. Predicting their bathing behavior appears to be challenging as the RMSE has notable peak during the typical hours even for 1-hour-ahead forecasts. Group 3 shows a very differently shaped curve but the reason for that is that the customers in fact seem to behave differently as can be seen from the average hourly load curves in Figure 4.5. Customers in group 3 as a whole have a lot more electricity usage around noon. Note that the average loads are calculated from the normalized loads. It can be seen that the RMSE curves share some similarities with the average load curves.

![Figure 4.5 Average hourly loads for each group of customers in January 2010.](image)

(a) Group 1. (b) Group 2. (c) Group 3.
4.3. Forecasting performance

![Average forecasting RMSE for July 2010.](image)

**Figure 4.6** Average forecasting RMSE for July 2010.

In July heating is no longer a factor which can be seen in the RMSE values for group 1. The RMSE values are slightly lower in general and the gap between 1-hour-ahead and next-day forecasts is slimmer compared to January forecasts. The late evening peaks in both RMSE and hourly loads for group 2 are also diminished, perhaps there is a change in the pattern people use their saunas. Group 3 again distinguishes itself from the others with a different curve shape altogether. Interestingly for group 3 the hours around noon seem to harder to forecast than the evening hours. The average hourly loads for July can be observed in Figure 4.7. The shape difference between groups 2 and 3 is not as clear as in January, the afternoon load valley for group 3 is somewhat deeper than for group 2.

![Average hourly loads for each group of customers in July 2010.](image)

**Figure 4.7** Average hourly loads for each group of customers in July 2010.

While the customers in groups 1 and 2 generally tend to have their peak loads during the evening the peak hours are distributed throughout the day in a noticeable double-peak fashion for customers in group 3. What is also interesting is that one might expect the next-day RMSE to increase monotonically all the way until midnight due to the longer forecasting horizon and the cumulation of errors caused by the reliance on past forecasts instead of measurements but this is does not seem to be the case,
in both January and July the average RMSE drops several hours ahead of midnight. What can be seen however is the slight widening in the gap between 1-hour-ahead and next-day forecasting RMSE towards the late night hours.

Forecasting individual residential customers was known to be hard and these results reflect that. The best PAR model next-day forecast (RMSE 0.3190) from Group 1 (direct electric heating) is illustrated in Figure 4.8. The forecasts are able to follow the general behavior and the fluctuations caused by the dependency on outdoor temperature are captured as well but the random peaks presumably caused by human behavior are too challenging to be forecasted.

![Figure 4.8 Forecasting performance on a direct heating customer.](image)

The error in the PAR model (Equation 4.2) was assumed to be white noise, that is normally distributed with zero mean. The following figure shows the fitted distribution of the 1-hour-ahead forecasting error for two time periods where the aforementioned direct electric heating customer regularly exhibited load spikes. It can be seen that the normal distribution is not the best fit. The violation of the normality assumption may imply that there are unmodeled dynamics.
4.3. Forecasting performance

![Diagram](image)

(a) Time period 10:00–11:00.

(b) Time period 18:00–19:00.

**Figure 4.9** Histogram and fitted normal distribution on forecasting errors.

The forecasts for a customer in Group 3 (no electric heating, no electric sauna) are illustrated in Figure 4.10, the RMSE value for next-day forecasts is 0.6837. This customer does not have electric heating so there is little dependency on outdoor temperature and therefore the load levels are similar throughout the whole test set. Again the general behavior is captured fairly well but random spikes prove to problematic.

![Diagram](image)

**Figure 4.10** Forecasting performance on a customer without electric heating or electric sauna stove.

Looking at the 1-hour-ahead forecasting error histogram at frequent load spike times
for this customer it can be seen that the normal distribution fits quite a bit better in comparison to the other customer with direct electric heating.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figures/error_distribution.png}
\caption{Histogram and fitted normal distribution on forecasting errors.}
\end{figure}

As stated the models are able to follow the general behavior trends with decent accuracy but random spikes are difficult to predict as was expected. Aggregating the data smooths these spikes out and therefore increases forecasting accuracy. Regardless of whether the next-day forecasts on single customers are accurate enough for real-life purposes they will be used for the change detection evaluation in the next chapter.
5. METHOD VALIDATION

In this chapter CUSUM and relative Pearson divergence methods are tested. The test cases consist of hand picked customers that either show real change happening within the year 2010 or possess a regular, smooth behavior that is easily artificially modified to represent changed behavior. The artificial change cases involve swapping the measurement data from one customer to another at some point in time while making sure the difference in load levels is minimal and only the shape of load curves change.

The tests are based on load forecasts. Measurements and forecasts in the same time window are compared in a sliding fashion, that is with a window of length \( w \) days measurements \( y_{d-w:d} \) and forecasts \( \hat{y}_{d-w:d} \) are fed into the algorithms, and the day \( d \) is incremented. Note that since the forecasting horizon for next day forecasts is 39 hours only the last 24 hours are taken into account so we have daily midnight to midnight forecasts.

The allowance term \( k \) in CUSUM method in Equation (3.1) is fixed to the standard deviation of residuals obtained from performing the forecasting procedure above on training data. The smoothing parameter in relative Pearson divergence is fixed to \( \alpha = 0.9 \) and the kernel width parameter in Equation (3.3) is chosen via cross validation from the set \( \sigma_{cv} = med \times \{0.2, 0.5, 0.8, 1, 1.2, 1.5\} \), where \( med \) is the median distance between samples in training data, likewise the regularization parameter \( \lambda \) in Equation (3.6) is chosen via cross validation from the set \( \lambda_{cv} = \{10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1, 10^2\} \). To give some idea of the magnitude of the PE scores the scores were also calculated on the training data and a dashed line representing the 95-percentile (obtained with the Matlab function \texttt{prctile}) of the scores is drawn on all PE test figures. One could then determine the significance of the scores by e.g. counting the number of (consecutive) days the score is above the percentile line.
5.1 Workshift change

The first test mimics an industrial customer changing work shifts. The data from one industrial customer is swapped with another industrial customer which introduces the change in work shifts. The change in load curves is illustrated in Figure 5.1. Testing the forecasting accuracy on unaltered data the PAR model offers reasonably accurate forecasts with a RMSE value of 0.2764, shown in Figure 5.3. The data is then modified so the change happens on day 213. From that point on model forecasts no longer provide accurate information on daily load profile, as shown in Figure 5.3. Change detection results using Pearson divergence and CUSUM are shown in Figures 5.4 and 5.5.

![Figure 5.1 Example weekly load profiles before and after the change.](image-url)

Figure 5.1 Example weekly load profiles before and after the change.
5.1. Workshift change

![Figure 5.2 Illustration of forecasting performance before the inserted change.](image)

![Figure 5.3 Illustration of forecasting performance after the inserted change.](image)

Great anomalous events happened early in the year and that is reflected on both the PE score and CUSUM. PE score then settles to around the 0.06 level. More smaller scale, shape-wise anomalous events happen right before the artificially inserted change, bringing the PE score up to the 0.1 level. However once the change has happened the PE score never goes down to pre-change levels. Shorter time window causes a lot more variation to the score but the relative change is still clear.
5.1. Workshift change

Figure 5.4 PE score for the year, vertical line marks the day of inserted change.

Figure 5.5 CUSUM score for the year, vertical line marks the day of inserted change.

CUSUM score remains stable at zero outside the bigger events although some activity can be noticed with the shorter time window. The events right before the inserted change go unnoticed when using the longer, 7-day time window since they are small in magnitude. The inserted change shows as ever increasing CUSUM value after the change point.
5.2 Customer abnormality

The next test shows irregular behavior from a household customer belonging to a class "housing + storage electric heating". The customer is on a tariff which offers cheaper electricity at nights, this in combination with the storage electric heating leads to peak loads at late nights when the time of cheaper electricity begins. Figure 5.6 shows hourly measured loads from years 2009 and 2010, in late summer 2010 the measured loads change significantly compared to 2009 and then return to seemingly normal level at the end of the year.

![Graph showing hourly measured loads from years 2009 and 2010.](image_url)

**Figure 5.6** Hourly measured loads from years 2009 and 2010.

The main model for this test is the neural network. Figure 5.7 shows forecasts made with the PAR model and neural network. As can be seen the PAR model continues to forecast the late night peak loads even when there are none while the NN model does not, which leads to smaller errors during the abnormal period. This enables us to see some differences in the performance of our change detection schemes as shown in the results below. Apart from the abnormal period both models perform with similar accuracy, thanks to the somewhat regular behavior by the customer. The RMSE value for NN forecasts is 0.4911.

This test also makes use of the univariate presentation of samples. Figure 5.8 shows the results with multivariate sample presentation. As can be seen the PE score stays below the reference line for a quite long time before staying consistently above, therefore one could make false judgments about the length of the irregular
period. One possible cause for the inconclusive PE test lies within the density ratio estimation itself: while the distributions change the density ratio may stay the same. For that reason the PE score stays near the upper bound regardless of the change in load shape. With the univariate presentation however, the PE score is clearly elevated during the abnormal period compared to the rest of the year and furthermore the score goes above and also below the reference line immediately at the points of change.

\[ \text{Figure 5.7 Forecasts and measurements around the moment of change.} \]

\[ \text{Figure 5.8 Relative PE score using multivariate sample presentation and a 7-day time window. Vertical lines mark the beginning and the end of abnormal customer behavior.} \]
5.2. Customer abnormality

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.9}
\caption{5.9 Relative PE score using univariate sample presentation and a 7-day time window. Vertical lines mark the beginning and the end of abnormal customer behavior.}
\end{figure}

The CUSUM test shows a difference between the two methods. CUSUM barely reacts to the abnormal period at all, this is due to the fact that while the forecasts are erroneous the mean absolute error within the time window is below the limit specified by the allowance term.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.10}
\caption{5.10 CUSUM test with a 7-day time window. Vertical lines mark the beginning and the end of abnormal customer behavior.}
\end{figure}

Now the question whether this kind of behavior is in fact abnormality and should not be taken into account for load profile purposes is worth considering. As stated it seems like the load at the end of the year 2010 is similar to the end of 2009 which suggests no permanent change has happened. However the period of different behavior is quite long, almost 4 months, so to explain the behavior with a vacation or a faulty hardware e.g. a water boiler sounds dubious.
5.3 Electric sauna stove installation

The next test is for detecting the installation of an electric sauna stove. For this test the data is averaged from several customers because finding suitable single customers for the testing is time consuming. "Customer 1" is an average customer from the class "row house, no electric heating, no sauna stove" (564 customers) and "customer 2" is an average customer from the class "row house, no electric heating, electric sauna stove" (380 customers). Only district heating and oil heating are present in these groups, 82%/18% split for the group without electric sauna and 56%/44% for the group with an electric sauna stove. The average daily profiles by the type of day for these customers are shown in Figure 5.11. Saturday is traditionally the most prominent sauna day in Finland and the visible load difference between the two customer classes in Saturday evening is presumably caused by the electric sauna stove. The test is done by swapping the measurement data between customers at day 174 in 2010. Note that the data is standardized individually to zero mean unit variance before the data swap. The average customer with an electric sauna stove has a higher mean consumption but since the data are standardized before the swap both customers have zero mean, meaning the difference in mean in raw data will not appear therefore making the change detection more challenging. The forecasts are made with the PAR model with multivariate sample presentation for the Pearson divergence estimation.

![Graph showing daily profiles for average row house customers with and without electric sauna stove.](image)

**Figure 5.11** Daily profiles for the average row house customers with and without electric sauna stove.

The model offers reasonable forecasting performance before and even after the change which truthfully is expected since the data is gathered and processed from several sources, smoothing out the noise. As household customers generally do not show such smooth behavior some white noise is added to each time series to see
the effect on method performance. The standard deviation of the added noise is set to 2% and 5% of the average difference between daily minimum and maximum values in training data. Model performance without the added noise is illustrated in Figures 5.12 and 5.13, the RMSE value on unaltered data is 0.2008, with noise added the RMSE values are 0.2078 for the 2% case and 0.2517 for the 5% case. It can be seen that the Saturday evening peak which is assumed to be caused by the electric sauna stove is not correctly forecasted.

**Figure 5.12** Example of forecasting performance before the electric sauna stove installation without the additional noise.
5.3. Electric sauna stove installation

Figure 5.13 Example of forecasting performance after the electric sauna stove installation without the additional noise.

The detection algorithms are applied as before using a 7-day window. Pearson divergence picks up the change immediately with the score staying permanently at higher level afterwards, see Figure 5.14. A small amount noise added has no notable effect on the PE score but 5% added noise makes it practically impossible to notice the change by visual inspection.

Figure 5.14 Change detection with Pearson divergence. Vertical line marks the day of change.
Results from the CUSUM test are in Figure 5.15. CUSUM basically does not notice the change at all until remarkable amount of time has passed. The continuous rise in the score starts only after day 230, around two months later than the change occurred. There is some activity in the noiseless case but not a significant amount. The additional noise not only weakens the initial reaction but also slows down the eventual increase which is to be expected. The allowance term in the CUSUM algorithm is based on the standard deviation of training residuals, more noise means higher errors and therefore higher standard deviation. Another factor is that the chosen CUSUM algorithm operates on mean errors. The forecast accuracy on "normal” days is generally good, thus the errors from the few evening peaks during the week are diminished by the averaging process enough to keep the mean error below the allowance term.

![CUSUM Algorithm Graph](image)

**Figure 5.15** Change detection with CUSUM algorithm. Vertical line marks the day of change.

### 5.4 No change

This test further illustrates the meaning of change in this thesis. Figure 5.16 shows average daily loads from the years 2009 and 2010 for an individual customer belonging to a class "row house/apartment, no electric heating, electric sauna stove". The customer seems to have a tendency for significant load level shifts in spring and autumn. However the timings of the level shifts are a lot different in 2010 in comparison to 2009. Heating should not be the deciding factor since the customer is stated have no electric heating but even if the customer information was incorrect the temperatures between the years are not significantly different (see Figure 5.17).
5.4. No change

around the shift timings to explain the disparity. Therefore one could argue this is a case of changed behavior.

**Figure 5.16** Average daily loads for a certain customer from years 2009 and 2010. Notice the timing on level shifts.

**Figure 5.17** Average daily temperatures from years 2009 and 2010.

Forecasts are made with the PAR model with resulting RMSE value of 0.7797. The forecasting performance from Thursday to Wednesday around the latter level shift is displayed in Figure 5.18. The level shift occurs on Friday (around the hour 6560 in Figure 5.18) and the discrepancy between the measurements and the model caused by the shift is gone on Sunday. Residential customers are notoriously difficult to model due to high variation but depending on the customer some elements can be forecasted accurately. This customer in particular shows high peaks on every Saturday mostly at 16 o’clock and our model captures this behavior. The customer also somewhat regularly shows smaller peaks on weekdays around 11 o’clock, however those our model struggles with.
5.4. No change

![Figure 5.18](image)

**Figure 5.18** Illustration of forecasting performance, from Thursday to Wednesday. High Saturday peak is captured although lacking in magnitude.

The change detection methods are applied as before, with 7-day windows. No change can be determined from the PE score as nothing out of the ordinary happens around the marked days, as seen in Figure 5.19. The sample presentation makes no difference. CUSUM shows no reaction either with the score staying at zero all the time, Figure 5.21.

![Figure 5.19](image)

**Figure 5.19** Relative Pearson divergence with multivariate sample presentation, vertical lines mark the days of level shifts. No change can be determined.
Figure 5.20 Relative Pearson divergence with univariate sample presentation, vertical lines mark the days of level shifts. No change can be determined.

Figure 5.21 CUSUM test, vertical lines mark the days of level shifts. No change can be determined.

As mentioned before the change in this thesis is with respect to model forecasts. Even when the customer behavior seems to change between years if the model is able capture it no change is determined. This test reflects that sentiment.

5.5 Unknown change

Figure 5.22 shows average days for a certain customer with direct electric heating. Besides the slightly higher consumption which is explained with colder temperatures (see Figure 5.17), the daily profiles are very similar in shape with minor differences in the first half of the year. In the latter half of the year the shapes of the load
curves are noticeably different in 2010, in particular the morning peak at weekdays and the higher consumption on Saturday afternoons. Some kind of change must have happened.

(a) Average days in January-June.

(b) Average days in July-December.

Figure 5.22 Comparison of average days of a certain direct electric heating customer from years 2009 and 2010.

The change detection methods are applied as before with 7-day windows. The forecasts are made with the PAR model, resulting in RMSE value of 0.9658. Figure 5.23 shows the results Pearson divergence test with multivariate sample presentation. Based on this figure the change has happened around day 230 which is in late August.
Besides a couple of instances the score never goes below the reference line afterwards.

![Figure 5.23 Pearson divergence with multivariate sample presentation.](image)

The results from the univariate test seem to agree, see Figure 5.24. The score however dips below the reference line several times so the test is not as conclusive as it is with multivariate sample presentation.

![Figure 5.24 Pearson divergence with univariate sample presentation.](image)
5.5. *Unknown change*

Results from the CUSUM test (Figure 5.25) again differ from the Pearson test. If the CUSUM test was to believed the change would happen around a month later than what the Pearson test indicates. Since this is natural, unaltered data there is of course no information about when and why the change has happened but going by the previous results obtained in this thesis the Pearson test seems more trustworthy.

![Cumulative sum test](image)

**Figure 5.25** Cumulative sum test. The assumed changepoint is at much later date than with the Pearson test.

Based on the results presented in this chapter it seems that a way to detect behavioral changes has been found. While the CUSUM method does not seem to be able to detect changes of relatively small magnitude these changes could be detected on time by using the relative Pearson divergence, not considering the delay caused by the time window. The sample presentation plays a part as it seems that a change undetected by using one presentation scheme can be observed with the other, because of this one could argue this is not a universal solution. Compared to previous change detection work done at the Tampere University of Technology, namely the work by Tao Chen [5], the proposed method of monitoring the relationship between model forecasts and measurements with Pearson divergence provides more accurate time information about the change and also allows us to skip temperature normalization and weekday alignment preprocessing steps by incorporating the outdoor temperature and daytype variables directly to our models.

The accuracy of long horizon forecasts obtained for residential customers with the
models used in this thesis may not be appropriate enough for real world use but nevertheless the change detection methods seem to work. Perhaps one could then consider for example a scenario where the actual load forecasting is performed on aggregate loads based on customer classes and the forecasts for individual customers are only used to monitor if the customer classification is fitting by using the relative Pearson divergence based change detection scheme. If a change is detected the customer may need to be reclassified.
6. CONCLUSION

Change detection is an intricate problem with several solutions proposed but not many have been applied to electric load measurements. Detecting changes that are not caused by yearly variation in temperature or other external factors is particularly important as it is then related to changes in customer behavior, which in turn leads to changes in customer classification on the operator side. This thesis proposes a method that ties short-term load forecasting and change detection by comparing measurements with forecasts either by using forecasting error or probability densities, the latter approach appears to be more accurate in terms of detecting the change in the first place and detection delay in the case the change is in fact detected. Of the forecasting models the PAR model is found more suitable due to simplicity of use and forecasting accuracy, however as forecasting individual residential customer is shown to be challenging because of high randomness the need for better models is real.

Going forward better load modeling seems like the highest priority. The change detection schemes seem to work even in the case the true nature of human behavior in the form of random load spikes are not captured by the models but some fine tuning can be done within the change detection methods as well. The parameters were chosen somewhat arbitrarily based on few personal tests and heuristics presented in the literature but surely better performance could be achieved if proper research and theoretical foundation were the basis for the choices. The reference which determines whether the change is significant or not should be studied as well, a simple 95-percentile value acquired from the training data works if one looks at the results retrospectively but if the method was to be used in an online setting something better is needed. Nonetheless the way of combining change detection with load forecasting shows great promise.
BIBLIOGRAPHY


