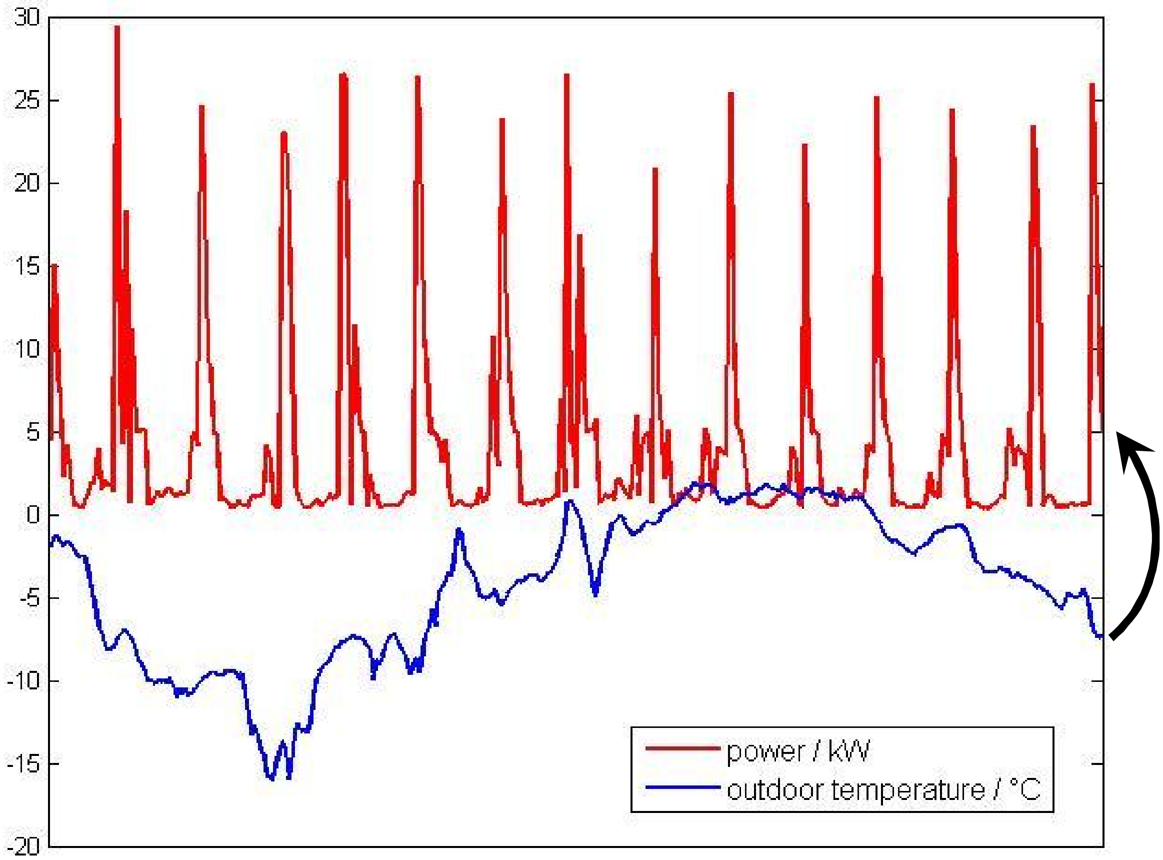
RESEARCH REPORT

VTT-R-02882-12

Prediction of energy consumption from outdoor temperature for houses electrically heated via heat storage

Authors: Henri Riihimäki, Pekka Koponen

Confidentiality: Public



Report's title	
Prediction of energy consumption from outdoor temperature for houses electrically heated via heat storage	
Customer, contact person, address TEKES Jussi Mäkelä	Order reference 1031/09 3. jakso
Project name Smart Grids and Energy markets	Project number/Short name 78560/SGEM III
Authors Henri Riihimäki, Pekka Koponen	Pages 20/-
Keywords Load prediction, measurement based approach	Report identification code VTT-R-02882-12

Summary

Load prediction is needed by different actors in electricity networks and markets. Accurate load prediction enables network operators to manage loading and electricity retailers to reduce their market and imbalance costs. It also helps in attacking the problem of spatio-temporal mismatch thus increasing the efficiency of the power system. Being able to estimate future loads further helps to predict when and where consumption peaks may occur and how they can be managed by load control actions.
This report investigates the prediction of energy consumption of houses that are electrically heated via hot water tanks that act as heat storage. Such a system is called full storage heating. The prediction is based on the measurement of outdoor temperatures averaged over periods that are 24 hours long. Results confirm two points. First, predictability is the better the newer the temperature data until the time shift between the temperature and power is about 21 hours and temperature data overlaps the heating period by 3 hours. Then the prediction performance starts to deteriorate with newer data. Secondly, predictability is better for a larger group of houses as the stochastic variations in the behaviour of single houses smooth out, which enables detection of differently behaving houses and their removal from the group, which further improves prediction accuracy.
For the analysis software tools were developed for detection of houses that behave differently from the others due to various reasons such as 1) other additional heating, 2) bigger proportion of loads that do not depend on outdoor temperature, or 3) inadequate dimensioning of heating, heat storage and insulation.

Confidentiality	Public

Espoo 27.6.2012 Written by	Reviewed by	Accepted by
H. Riihimäki, P. Koponen Research trainee, Senior scientist	Göran Koreneff Senior scientist	Seppo Hänninen Technology manager

VTT's contact address
VTT, Tekniikantie 2, P.O. Box 1000, 02044 VTT, Espoo, Finland

Distribution (customer and VTT)
Project SGEM

Preface

This work was carried out in the Smart Grids and Energy Markets (SGEM) research program coordinated by CLEEN Ltd. with funding from the Finnish Funding Agency for Technology and Innovation, Tekes.

This report studies the possibility of predicting energy consumption of an electrically heated house or a group of houses based on the measurement of outdoor temperatures. The study is part of a research on smart metering based dynamic load control that is reported in [5] and [6].

Thanks go to Pekka Koponen of VTT for his invaluable advices and help in writing this report, Jussi Matilainen of Fingrid, Joel Seppälä of Helen Electricity Network, Ossi Porri of Helsinki Energy and Anders Wickström of Fortum Electricity Markets.

Espoo 27.6.2012

Henri Riihimäki

Contents

1 Introduction

Load prediction has been one of the main themes in electric power systems and continues to be in the future Smart Grid concept. Load prediction is needed by network operators and by competitive electricity market actors. Accurate load prediction enables 1) network operators to manage voltage quality and network loading and 2) electricity retailers to reduce their imbalance costs and market risks. Knowing the energy consumed in place and time also helps to attack the problem of spatio-temporal mismatch, thus increasing the efficiency of the power system. Being able to estimate future loads also helps to predict when and where consumption peaks may occur. This in turn makes it easier for the supply side to prepare for the peaks making the energy delivery more reliable [1] and for the demand side to know when and how to apply demand response.

Static time of use control is being upgraded to dynamic load control [2,3] that controls the loads based on the variations in market prices and network state instead of a clock. Accurate predictions of daily energy demand form a necessary basis for scheduling the dynamic load control actions.

Loads are usually affected by many dynamic and stochastic variables, making the load behaviour a function of these variables. Load prediction becomes difficult when the behaviour in time of the dynamic variables is not well known. In this report we focus on one load-dynamic variable pair, the namely outdoor temperature and the electric power consumed by an electrically heated house. Outdoor temperature is an example of a highly stochastic dynamic variable. As is long known in meteorology, forecasting the weather exactly is impossible and this applies to outdoor temperature as well. Temperature can even exhibit strong fluctuations within a day and, for example in Finland, the outdoor temperature can change 20 degrees during a day in certain seasons. This makes electric heating a highly variable load [4].

There are two ways to approach the problem of load prediction: model-based approach and measurement-based approach. In model-based approach we construct a physical model of the load system from first principles. This usually means a system of differential equations comprising the known physical laws governing the behaviour of load components. Such systems have been constructed, see for example [5]. Load prediction is then based on obtaining a solution for the system of differential equations. But physical modelling can be complicated. In the case of modelling the indoor temperature behaviour of a house, the difficulty arises due to myriad of different heat capacities and pathways for heat conduction, especially in the case of direct or partially storing heating. In the case of fully storing heating, the required heating is retrieved from the storage tank. Thus the heat storage acts as a buffer decoupling the outdoor and indoor temperatures and simplifying the impact of the heat dynamics of the house [5] to the power consumption.

In measurement-based approach the modelling goes to the opposite direction. We start with measurement data and try to identify the best model to replicate the data most accurately. That usually means a curve fitting problem. Measurement-based approach is suitable when the studied system is too complex or impossible to model with physical laws. Or we may only be interested in studying if a relationship exists between some dynamic variables as is the case in this report. Often measurement-based approach is used in connection with model-based approach in which case measurements are used to tune the model parameters to match the actual system [6]. This approach of combining measurements with physical model was used in [5]. In a full storage heated house the heat storage tank and temperature control system decouple the detailed dynamics of the house and measurement-based approach is adequate for the prediction of the daily energy consumption. Thus the analysis here is measurement-based and not based on physical models.

2 The aim and method of study

In this report we present the results of a research task that focused on predicting the energy consumption of electrically heated fully storing houses based on outdoor temperature data. One motivation of the research was to study the effect of using newer temperature data when predicting the energy consumption. Today, the distribution network operator uses the average temperature of the previous day to predict the energy demand of the next day. More accurately there is a time shift of 45 hours between the input data and the output data of the prediction. Therefore there is a gap between the data collection interval and the interval for which we want to predict the energy consumption. During this gap temperature can change considerably and the energy use of electrical heating is not estimated as accurately as possible. This is more of a problem when the time reserved for heating is based on too small an estimated energy and as a result not enough heating energy is supplied to maintain the desired temperatures of indoor air and domestic hot water. On the other hand, too large an energy consumption estimation reduces the efficiency of the power system.

First we used available measurement data from two houses. The data consisted of outdoor temperature and power delivered to the houses. Figure 1 shows the time series for one house to give an idea what the data sets look like and Figure 2 shows two first weeks of the data to give more accurate picture of the variations of outdoor temperature and power.

Time span of the measurements was 1.1.2011-31.5.2011 and time interval between consecutive measurements was one hour. We also had a small data set for the other house from 15.11.2010 to 8.12.2010. This data was used to compare prediction models for the house made from fall and spring measurements. Later during the research we also received measurement data for a large group of houses. After analysis acceptable data was found for 185 houses. This data was used to study the prediction of total energy consumption of a large number of houses. Presumably, in a large group of houses the stochastic variations of energy consumptions in single houses even out and better prediction capability was expected.

The aim was to determine how we should observe the outdoor temperatures of some previous time interval in order to predict the energy consumption of the houses for the following 24 hour time interval between 21 o’clock-21 o’clock. This time interval was chosen to take with minimum delay into account the full night period, when electricity price is typically lower and the heat is stored, and consequently most of the energy consumption of a house with fully storing electric heating takes place. These consumption peaks are clearly shown in Figure 2.

The research was made using MATLAB software. Energy consumptions of time intervals 21 o’clock-21 o’clock were simply integrated from power measurements. Basically the idea was to find the best way to calculate some characteristic of the outdoor temperature of some previous time interval with the help of which we could achieve best correlation to energy consumption. In other words we were looking for a function

$$f(T) = E_{21\text{-}21}, \tag{1}$$

where T is the characteristic of the outdoor temperature for some time interval and $E_{21\text{-}21}$ is the energy consumption of the following 24 hour time interval 21 o’clock-21 o’clock. When calculating the characteristic from outdoor temperature, we focused on to investigate the 24 hour time intervals prior to the 24 hour time interval 21 o’clock-21 o’clock during which the energy consumption was predicted. Different ways to calculate some characteristics from temperature data were tried, for example standard deviation, variance and some other ways. However, the best way turned out to be the mean of the temperatures and this was used in obtaining the results presented in this report. Mean temperature and energy consumption calculated from measurement data provided us with data points. Prediction model was obtained by fitting a curve to these data points in a least-squares sense. In this report temperature interval means the interval from which the mean temperature is calculated. Likewise prediction interval means the interval 21-21 for which energy consumption is predicted.

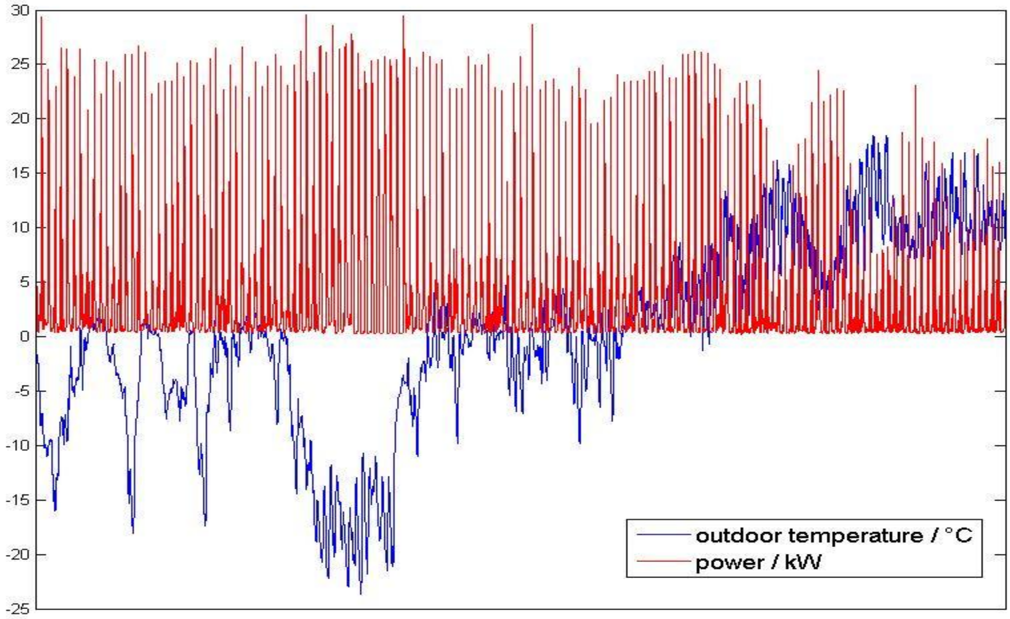


Figure 1: Time series measurement data used in the analysis of house 1

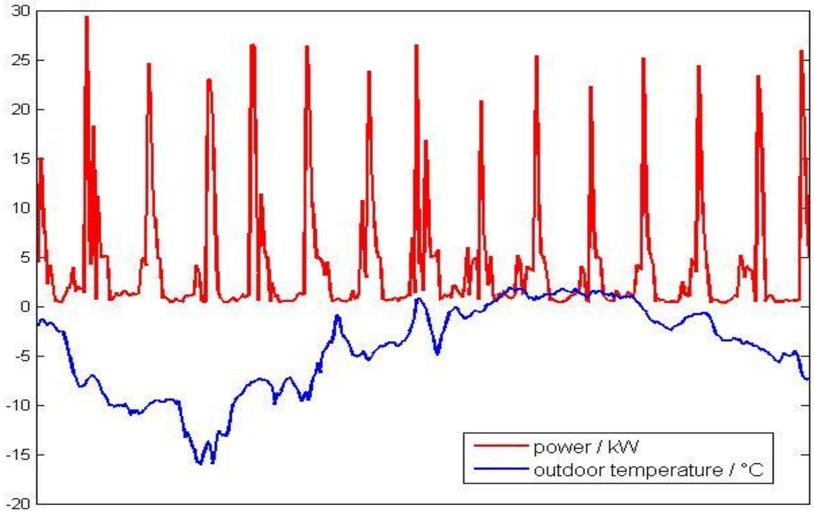


Figure 2: First two weeks of the time series measurement data of house 1.

3 Results

3.1 House 1

Mean temperatures of several 24 hour time intervals were calculated and 24 hour energy consumption of the following time interval 21-21 was plotted against the mean temperatures.

Prediction model was obtained for the house with linear curve fitted to the data points. Figure 3 shows this for two representative time intervals. The middle line is the fitted model. The lines above and below the model are 95% confidence bounds for the model. This means that 95% of predictions made by the fitted model fall between these confidence bounds. This is mainly of practical interest in the prediction of energy consumption. Because exact prediction of consumption is impossible due to the random behaviour of consumers, retailers are interested to know with what certainty the predicted consumption matches the actual use.
We also investigated the possibility of using 48 hour time intervals to calculate the mean temperature to see the effect of longer history on the dynamics. The results are shown in Figure 4 for two time intervals. Comparing the pictures shows little variation between different methods used to calculate the mean temperature. Actual comparison with some numerical key figures is made in Tables 1 and 2 later in this section.

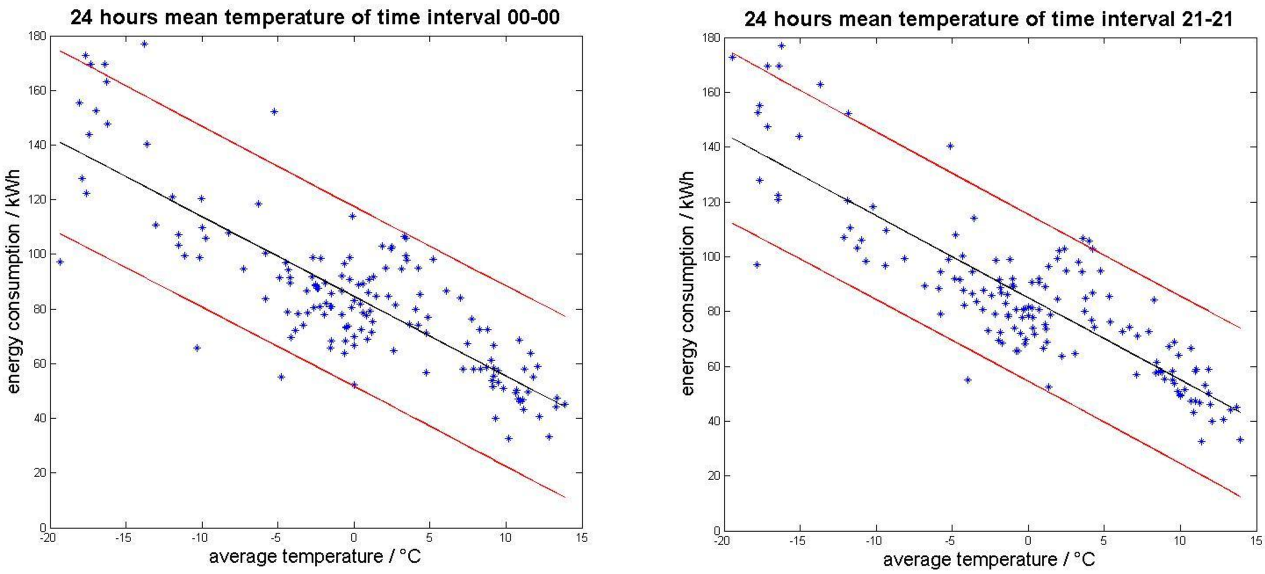


Figure 3: Energy consumption in kWh for House 1 plotted against average temperatures of 24 hour time intervals 00-00 with 45 hour time shift and 21-21 with 24 hour time shift.

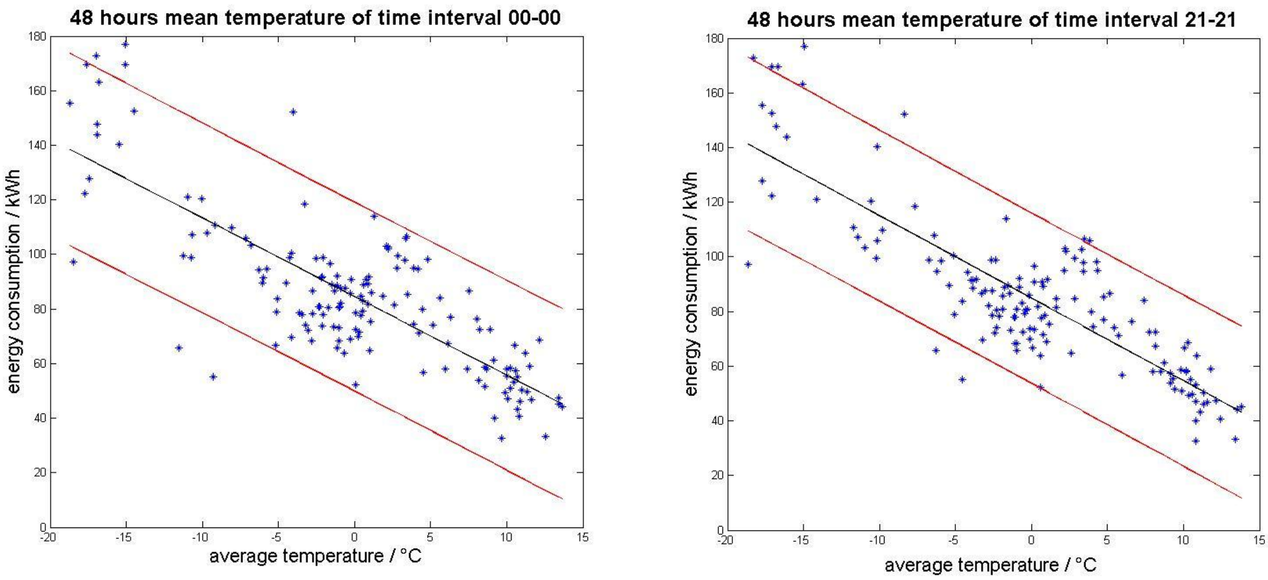


Figure 4: Energy consumption in kWh for House 1 plotted against average temperatures of 48 hour time intervals 00-00 with 69 hour time shift and 21-21 with 48 hour time shift.

The effect of using 2 days weighted mean temperature was also investigated for some temperature intervals. Based on previous results, linear correlation was assumed. Thus following function was used with Matlab-function that tries to find the minimizing parameters x_n using least-squares method:

(2) $$\overline{x_1(x_2\bar{T}_{time\ interval\ 1} + x_3\bar{T}_{time\ interval\ 2})} + x_4 - E_{24h}.$$

Equation of a line can be recognized in the sense of equation (1) from which energy consumption data is subtracted. Used Matlab-function generated the least squares sum to find the parameters x_n. Overbars represent mean values and $\bar{T}_{time\ interval\ 1}$ and $\bar{T}_{time\ interval\ 2}$ are the mean temperatures of intervals 1 intervals 2, respectively. We were interested in finding the weight coefficients x_2 and x_3. Instead of using the parameters of line, x_1 and x_4, we only used the weight coefficients to calculate the 2 days weighted mean temperature and then used the same curve fitting tool as with other results to achieve better comparability. Results showed very little variation among each other.
This method was rather unstable and suffered from convergence failures. The used Matlab-function was not able to find a solution with all initial guesses of the minimizing parameters and results were very sensitive to initial guess. Furthermore, using weighted mean brings no improvement to the correlation. These aspects can be seen in Table 1. Thus, the more simple methods using only mean temperatures seem preferable.
Table 1 summarizes the residuals, or the minimum sums of squares, at the best fit for different temperature intervals. Time shift is the time in hours from beginning of the temperature interval to the beginning of the prediction interval. Overlap means that the last hours of the temperature interval overlap the first hours of the prediction interval. Table shows that better correlation between mean temperature and energy consumption is achieved by reducing the time shift, or in other words, using temperature data closer to the prediction interval. Table 1 also shows that the best method to calculate the mean temperatures is to use only 24 hour temperature intervals.

temperature interval	time shift in hours	1 day mean temperature	time shift in hours	2 days mean temperature	2 days weighted mean temperature
00-00	45	$4{,}04\cdot10^4$	69	$4{,}44\cdot10^4$	
13-13	32	$3{,}66\cdot10^4$	56	$3{,}90\cdot10^4$	$4{,}07\cdot10^4$
17-17	28	$3{,}55\cdot10^4$	52	$3{,}75\cdot10^4$	$4{,}13\cdot10^4$
20-20	25	$3{,}48\cdot10^4$	49	$3{,}63\cdot10^4$	
21-21	24	$3{,}46\cdot10^4$	48	$3{,}60\cdot10^4$	$3{,}53\cdot10^4$
22-22 (overlap)	23	$3{,}44\cdot10^4$	47	$3{,}57\cdot10^4$	
00-00 (overlap)	21	$3{,}43\cdot10^4$	45	$3{,}53\cdot10^4$	

Table 1: Minimums of sums of squares at the best fit for different methods

Table 2 shows dispersions, or the differences between upper and lower confidence bounds, in prediction of energy consumption for different temperature intervals. Confidence bound is symmetric with respect to the model. For example for temperature interval 21-21, consider that the model predicts energy consumption of 120 kWh for mean temperature of -10 degrees Celsius. Then with 95% certainty the observed energy consumption lies somewhere between (120-30,52)kWh and (120+30,52)kWh. Table 2 confirms the results of Table 1 that the better the prediction the newer the temperature data. There are two reasons why shorter time shifts than 21 hours are not considered. Firstly reducing time shift further from 21 hours starts to worsen the prediction for houses that have enough heat storage. This is due to the dynamics of the heat storage and associated temperature control. Secondly the decision for heating periods is made several hours before the start of the heating period to allow data communication latencies when sending the control signals. Thus not only for the overlap time but also for some more hours before it the temperature is based on short term weather forecasts instead of measurements. Tables 1 and 2 assume perfect temperature forecasts.

temperature interval	time shift in hours	1 day mean temperature	time shift in hours	2 days mean temperature
00-00	45	65,95	69	69,43
13-13	32	62,78	56	64,96
17-17	28	61,83	52	63,74
20-20	25	61,18	49	62,78
21-21	24	61,03	48	62,50
22-22 (overlap)	23	60,89	47	62,22
00-00 (overlap)	21	60,55	45	61,62

Table 2: Dispersions in prediction of energy consumption (in kWh) for different methods.

3.2 Comparison House 2

We also had measurement data from another house. This data had a few days period when the power measurement had been offline. Nevertheless, the data was used for comparison. This section gathers the same results as in the previous section. Figures 5 and 6 present the data points together with the fitted model and confidence bounds.

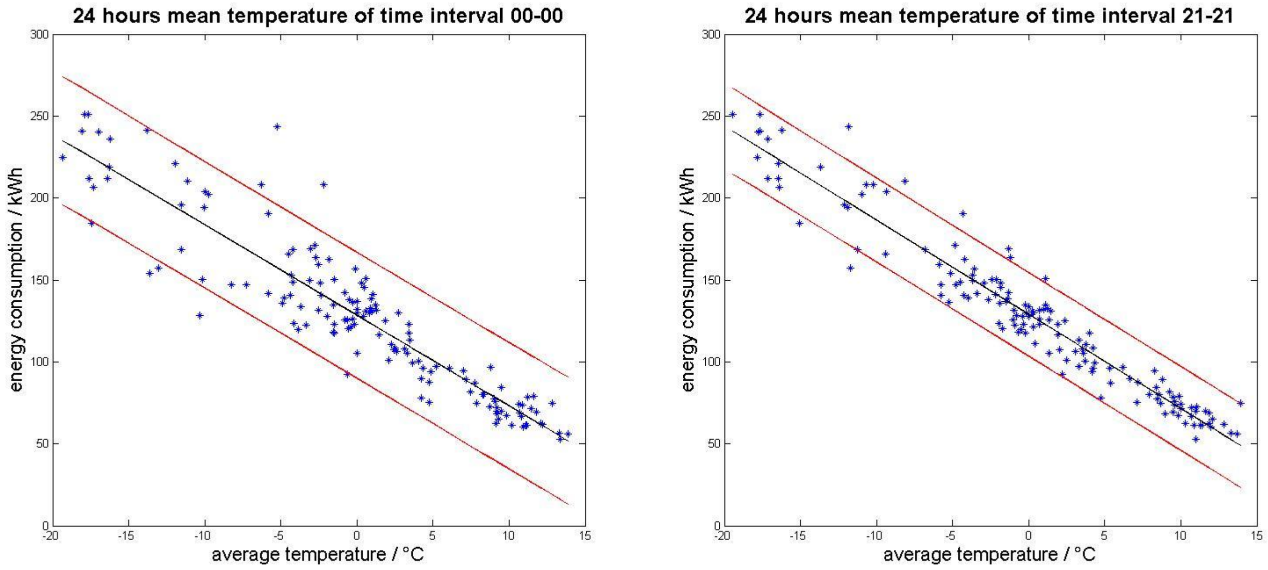


Figure 5: Energy consumption in kWh for House 2 plotted against average temperatures of 24 hour time intervals 00-00 with 45 hour time shift and 21-21 with 24 hour time shift.

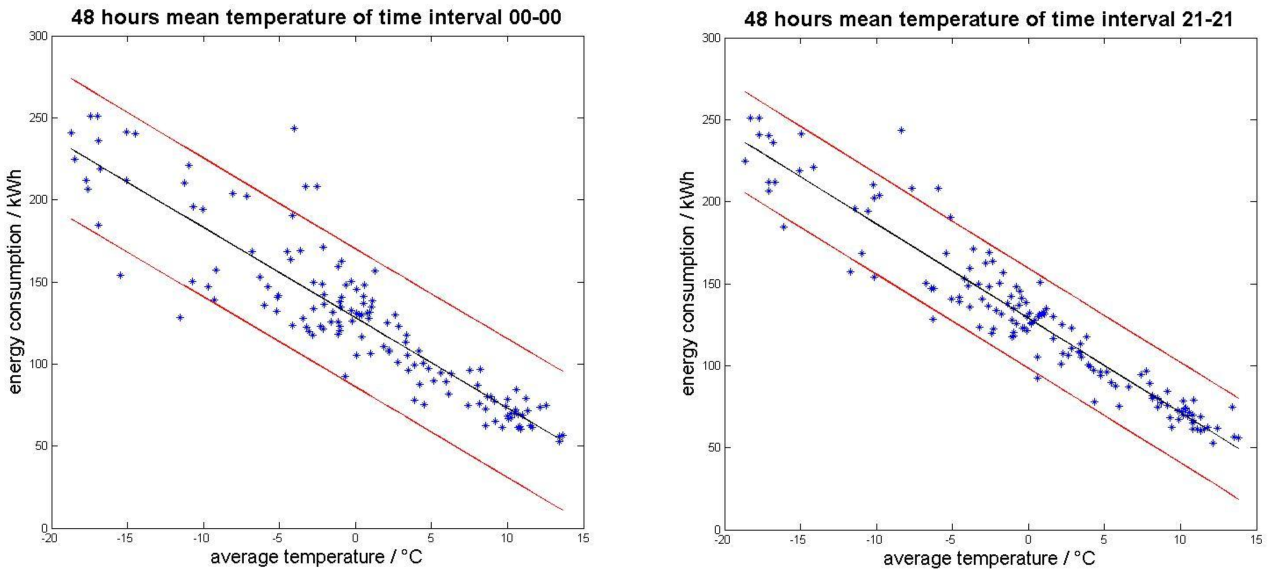


Figure 6: Energy consumption in kWh for House 2 plotted against average temperatures of 48 hour time intervals 00-00 with 69 hour time shift and 21-21 with 48 hour time shift.

Figures already show that for House 2 better correlation between the model and data is achieved. This is also confirmed with numerical key figures in Tables 3 and 4. Interestingly, the correlation is not so different from House 1 with longer time shifts but gets increasingly better with decreasing time shift. With 48 hours mean temperatures the correlation is even worse with longer time shifts and the difference doesn't become so noticeable with decreasing time shift than in the case of only 24 hours mean temperatures. This raises interesting questions about the heat dynamics of House 2 and the role of history in the dynamics. Potential for further investigation in properties of a house affecting the prediction and more thorough modelling and study of the heat dynamics is noticed.

time interval	time shift in hours	1 day mean temperature	time shift in hours	2 days mean temperature	2 days weighted mean temperature
00-00	45	$5,28 \cdot 10^4$	69	$6,23 \cdot 10^4$	
13-13	32	$3,16 \cdot 10^4$	56	$4,23 \cdot 10^4$	$3,21 \cdot 10^4$
17-17	28	$2,75 \cdot 10^4$	52	$3,79 \cdot 10^4$	$3,66 \cdot 10^4$
20-20	25	$2,44 \cdot 10^4$	49	$3,42 \cdot 10^4$	
21-21	24	$2,35 \cdot 10^4$	48	$3,30 \cdot 10^4$	$3,26 \cdot 10^4$
22-22 (overlap)	23	$2,25 \cdot 10^4$	47	$3,17 \cdot 10^4$	
00-00 (overlap)	21	$2,10 \cdot 10^4$	45	$2,92 \cdot 10^4$	

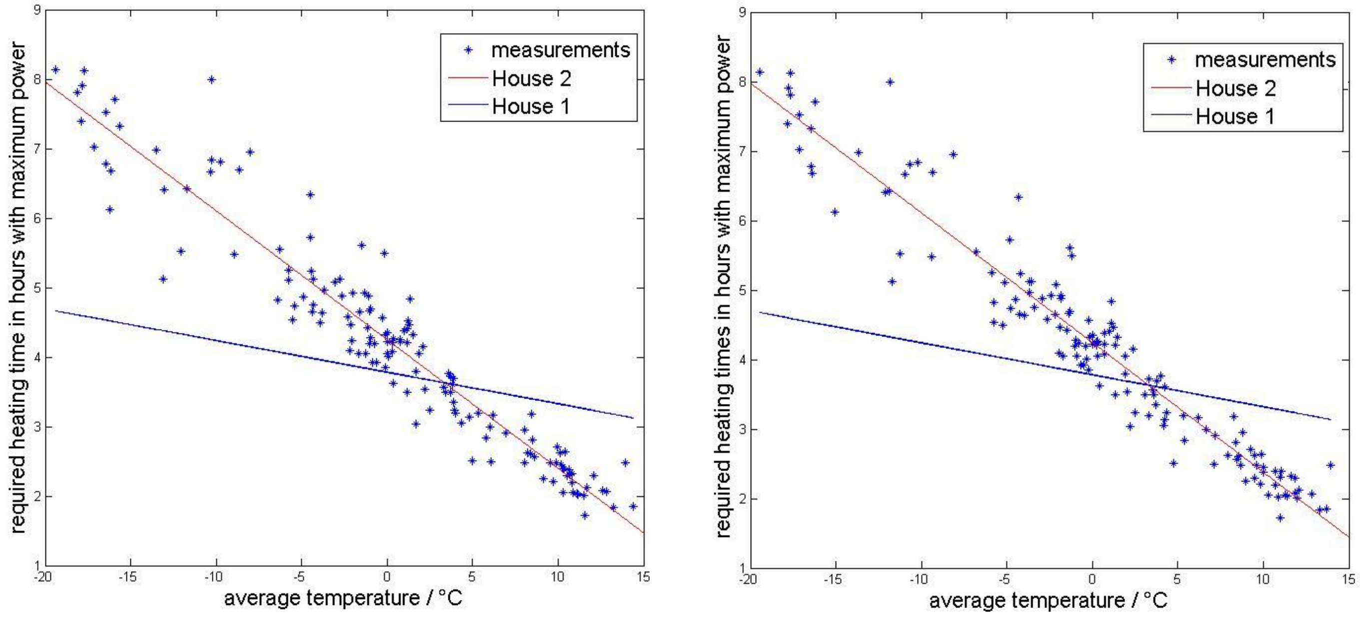
Table 3: Minimums of sums of squares at the best fit for different methods.

temperature interval	time shift in hours	1 day mean temperature	time shift in hours	2 days mean temperature
00-00	45	77,06	69	84,32
13-13	32	59,34	56	69,23
17-17	28	55,36	52	65,48
20-20	25	52,18	49	62,24
21-21	24	51,18	48	61,10
22-22 (overlap)	23	50,07	47	59,89
00-00 (overlap)	21	48,22	45	57,28

Table 4: Dispersions in prediction of energy consumption (in kWh) for different methods.

3.3 Comparison of models for Houses 1 and 2

We also compared the models of the two houses in order to see if the model for one house could be used to predict the heating demand of the other. For this the energy consumptions of prediction intervals were normalized with peak powers of the intervals. With this practice we obtained the heating times in hours with the assumption that heating is done with the maximum power. Figure 7 shows the predicted heating times for the two houses for 24 hours temperature intervals 17-17 and 21-21. Models give very different predictions. For example, when the average temperature is -20 degrees Celsius, model for House 1 gives too small a heating time for House 2. Structural properties of single houses, such as insulation and sizing of the heating capacity, are very different so the result is not very surprising. Also additional factors that change from house to house such as extra heating systems (fireplace for example) and differences in the occupancy affect prediction accuracy (prediction of an almost empty house is more accurate than that of a house where the residents use appliances much and otherwise affect the energy consumption). But predictions of the required heating times are very practical pieces of information from the point of view of an electricity retailer.



a) b)

Figure 7: Comparison of predicted heating times in hours for the two houses for temperature intervals a) 17-17 with 28 hour time shift and b) 21-21 with 24 hour time shift.

3.4 Comparison with older data for House 1

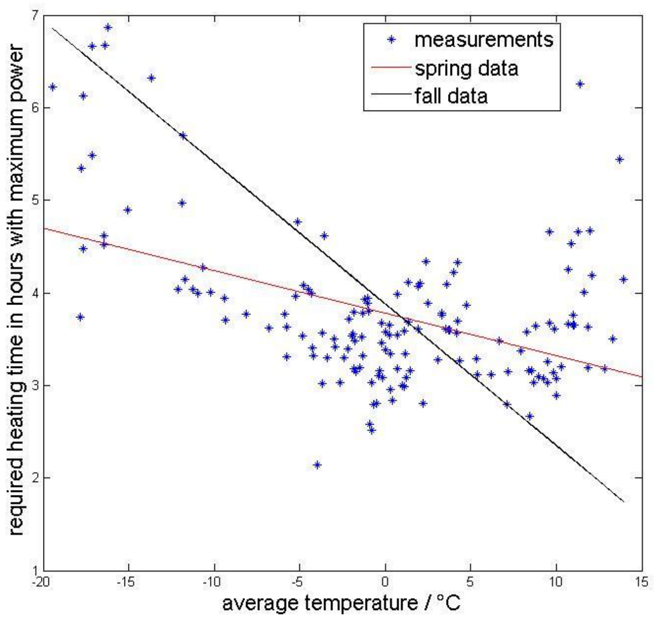


Figure 8: Comparison of predicted heating times for temperature interval 21-21 with 24 hour time shift for House 1 obtained from fall and spring measurements

Model for House 1 was also compared with a model obtained from different measurements made in fall 2010. Again the energy consumptions were normalized and we compared the predicted heating times. Figure 8 shows the comparison. Even though made for the same house, the models give rather different predictions. This can be due to variation in the sunlight, for example. In spring sun usually provides more natural heating than in fall, so heating times are expected to be smaller in spring. This may be amplified by the fact that low temperatures typically occur when the sky is clear. The occupancy of the house also varied, which may complicate the analysis.

3.5 Group of 185 houses

Investigations similar to the ones made in Sections 3.1 and 3.2 were also made for a house mass consisting of 185 individual houses. As discussed before, the behaviour of a single house as a load in a power system is normally highly stochastic. But in an aggregate of houses the random behaviour is expected to smooth out. Figures 9 and 10 and Tables 5 and 6 show this along with the already established result that better predictions are achieved with newer temperature data. Especially narrowing of confidence bound is clearly visible when temperature interval changes from 00-00 to 21-21.

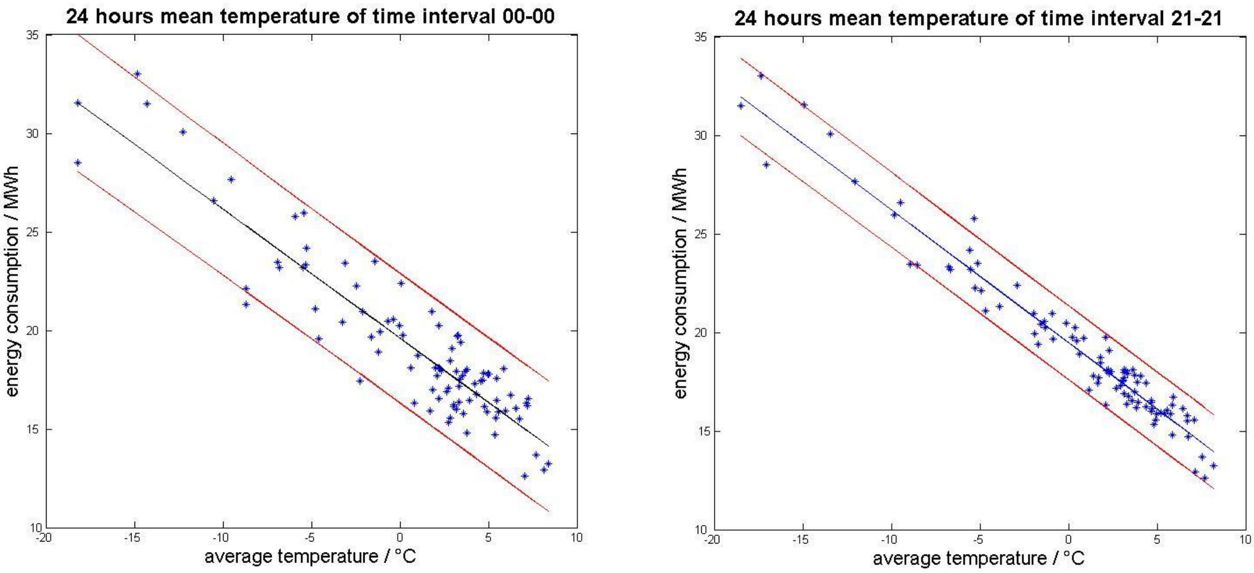


Figure 9: Energy consumption in MWh for a group of houses plotted against average temperatures of 24 hour time intervals 00-00 with 45 hour time shift and 21-21 with 24 hour time shift.

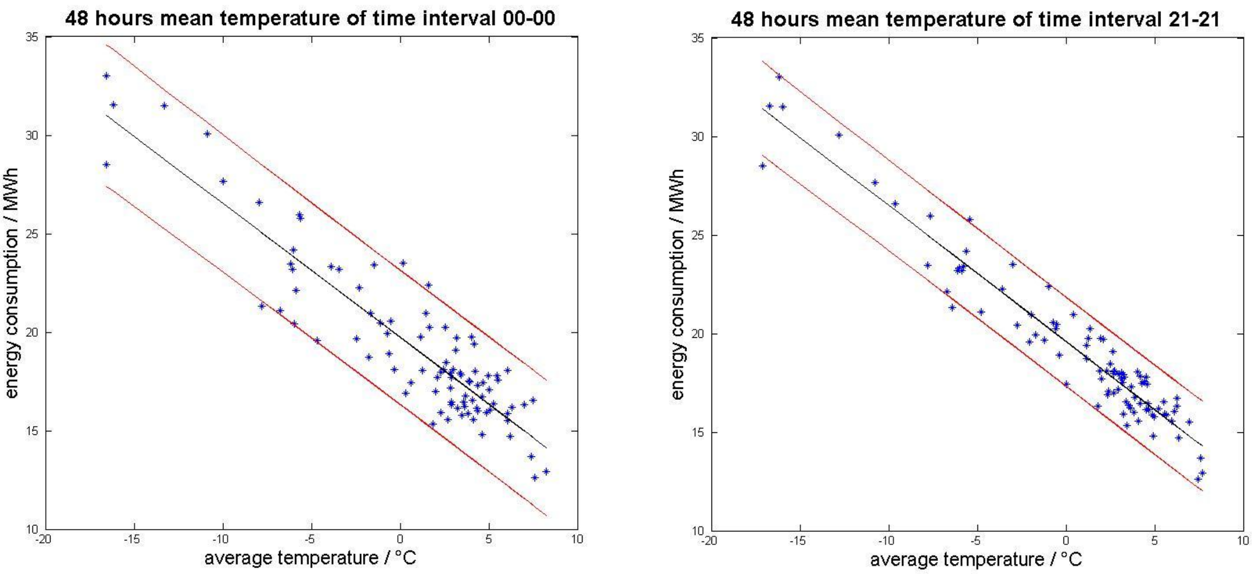


Figure 10: Energy consumption in MWh for a group of houses plotted against average temperatures of 48 hour time intervals 00-00 with 69 hour time shift and 21-21 with 48 hour time shift.

time interval	time shift in hours	1 day mean temperature	time shift in hours	2 days mean temperature
00-00	45	$2,47 \cdot 10^8$	69	$2,63 \cdot 10^8$
13-13	32	$1,33 \cdot 10^8$	56	$1,67 \cdot 10^8$
17-17	28	$1,04 \cdot 10^8$	52	$1,42 \cdot 10^8$
20-20	25	$8,55 \cdot 10^7$	49	$1,22 \cdot 10^8$
21-21	24	$8,00 \cdot 10^7$	48	$1,15 \cdot 10^8$
22-22 (overlap)	23	$7,53 \cdot 10^7$	47	$1,09 \cdot 10^8$
00-00 (overlap)	21	$6,79 \cdot 10^7$	45	$9,64 \cdot 10^7$

Table 5: Minimums of sums of squares at the best fit for different methods.

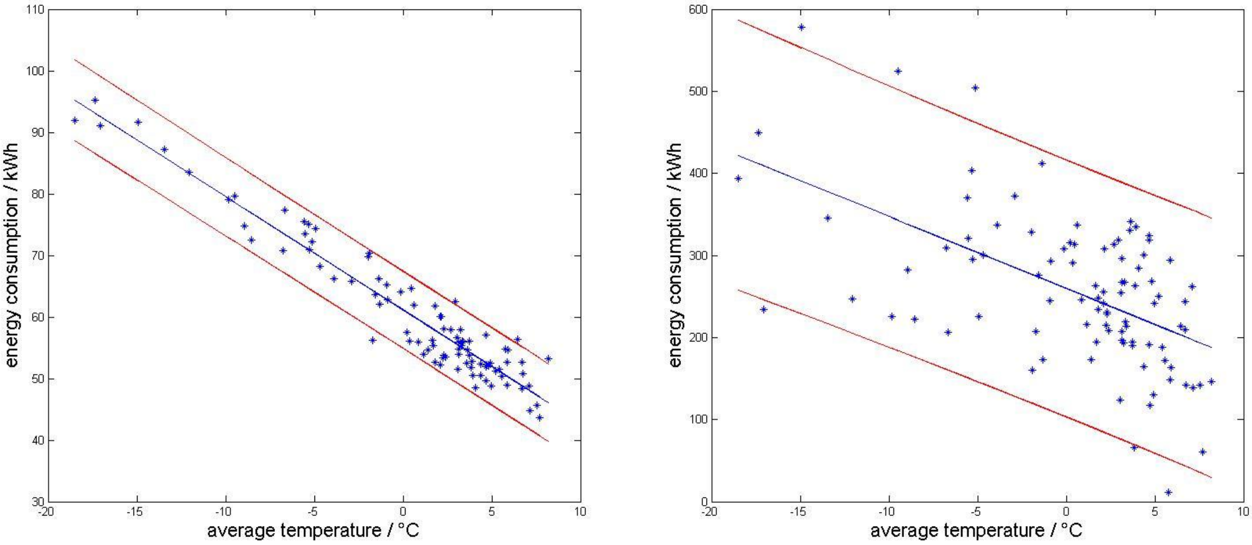
time interval	time shift in hours	1 day mean temperature	time shift in hours	2 days mean temperature
00-00	45	6,61	69	6,86
13-13	32	4,85	56	5,48
17-17	28	4,30	52	5,04
20-20	25	3,89	49	4,67
21-21	24	3,77	48	4,54
22-22 (overlap)	23	3,65	47	4,41
00-00 (overlap)	21	3,45	45	4,13

Table 6: Dispersions in prediction of energy consumption (in MWh) for different methods.

3.5.1 Variation in the behaviour of single houses in the group

We investigated how different a behaviour can occur among houses of the group. This we did by comparing the dispersions in the energy consumptions. We used temperature interval 21-21 with 24 hour time shift. First we calculated the average dispersion for the whole group. Then we used simple algorithm to search for the houses whose dispersion was 25% smaller or bigger than the average. From the results we picked up two extreme cases which are shown in Figure 11.

From the electricity retailer's point of view it is valuable information to know which houses from a larger group have poor predictability such as that shown in Figure 11 b). This information makes it possible to pay attention to houses for which load prediction doesn't work well and need some other methods of prediction. Or houses can be divided into groups with similar predictability.



a) b)

Figure 11: Two extreme cases from the group of 185 houses.
Dispersions in energy consumption are a) 12,61 kWh and b) 314,71 kWh.

Figure 12 shows dispersions in energy consumption for all the houses of the group. Predictability for around twenty houses seems to be rather poor. With information like this retailers could drop out these houses from the load predictions and increase further the prediction capability for the rest of the group. The house whose behaviour is shown in Figure 11 b) might not even be a normal residential house due to the drastically different behaviour which is shown as a peak in the graph of Figure 12.

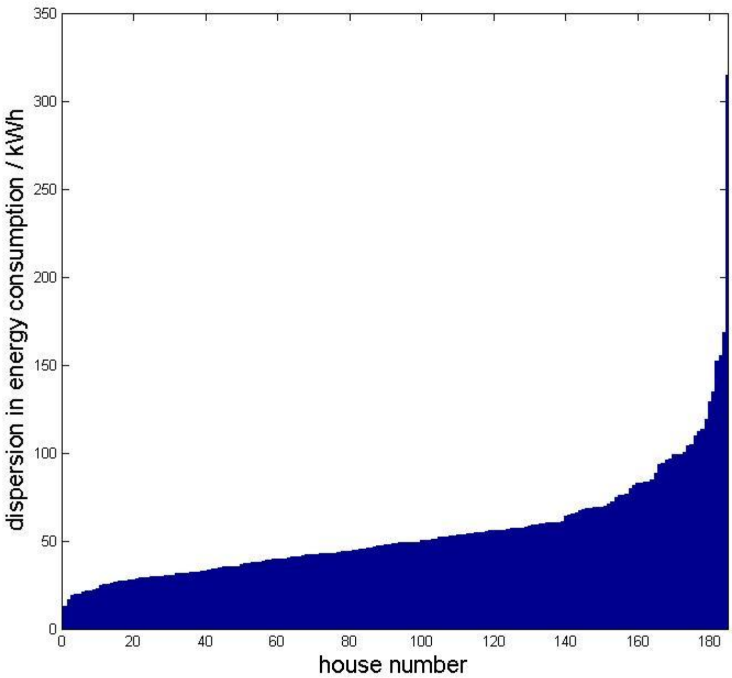


Figure 12: Dispersions in energy consumption for houses of the group.

We also investigated the predicted heating time for the group of 185 houses. Figure 13 shows the result. For comparison, the house with the longest predicted heating time on average was searched from the group. This was done by first calculating the average heating time for each of the houses. Then the maximum of the average heating times and the house corresponding to this was looked for. Figure 14 shows the prediction model for the house. Interestingly, the house was not the same as in Figure 11 b). The heating time of this house clearly exceeds the length of the night time low tariff.

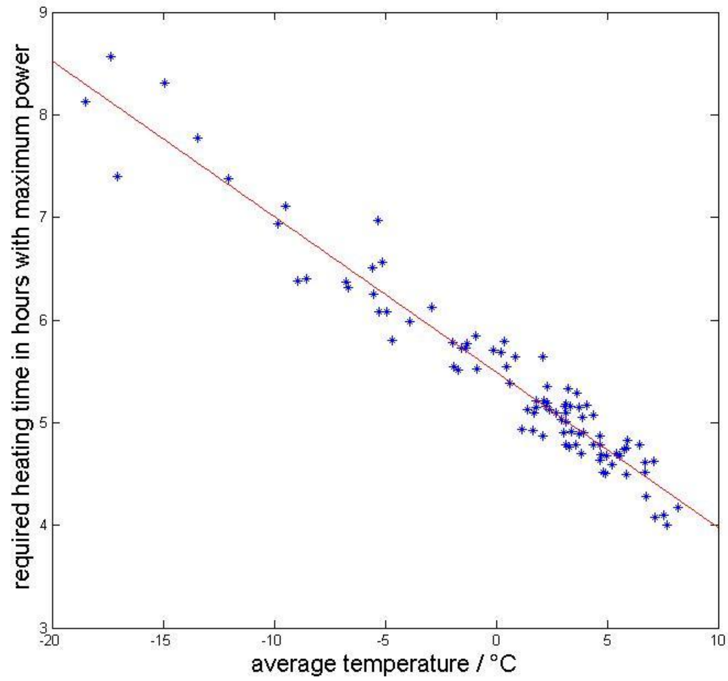


Figure 13: Predicted heating time for a group of 185 houses, temperature interval used was 21-21 with 24 hour time shift.

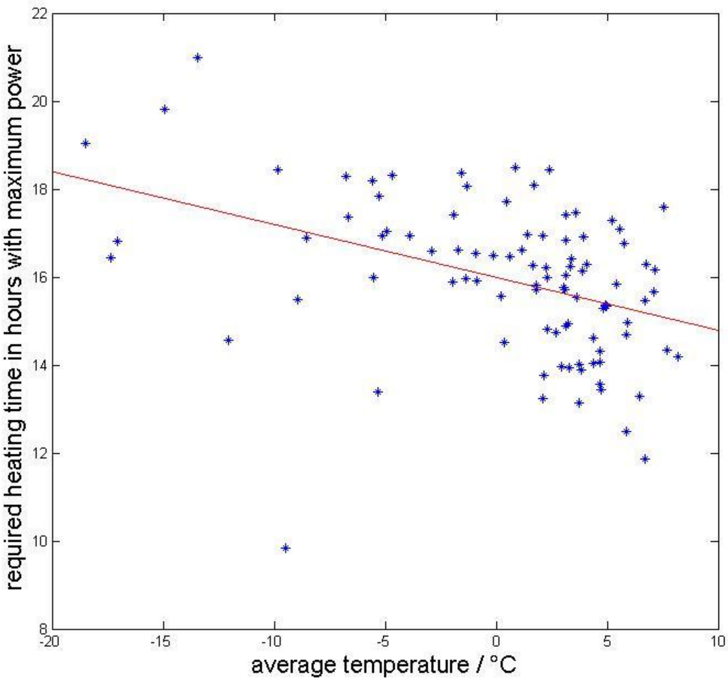


Figure 14: Predicted heating time for a house with the longest average heating time, temperature interval used was 21-21 with 24 hour time shift.

Figure 15 is from the same house as Figure 14. From Figure 15 it can be seen that when the temperature is above about -12 degrees Celsius the load in the house depends less on the temperature than in the whole group, but for cold outdoor temperatures the load is high. This suggests that the house may have an air-to-air heat pump or there are much loads that depend on something else than temperature. Figure 16 shows the time series data of that house and Figure 17 shows first five days of the time series. For this house the daytime loads are higher and night time loads lower than in houses with full storage heating. The heating is clearly not based on nigh time tariffs and the maximum heating power cannot be seen from the data.

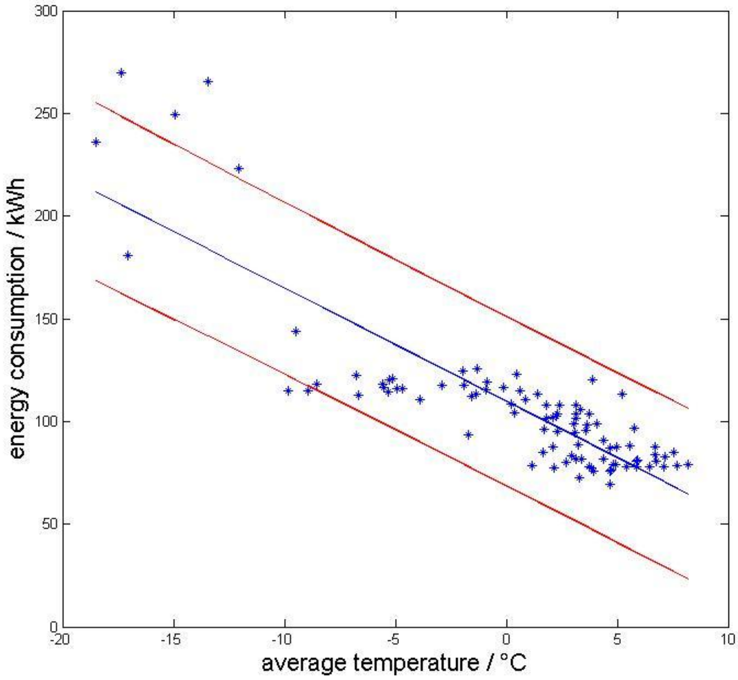


Figure 15: The temperature dependency of the energy consumption for the house of Figure 14, temperature interval used was 21-21 with 24 hour time shift

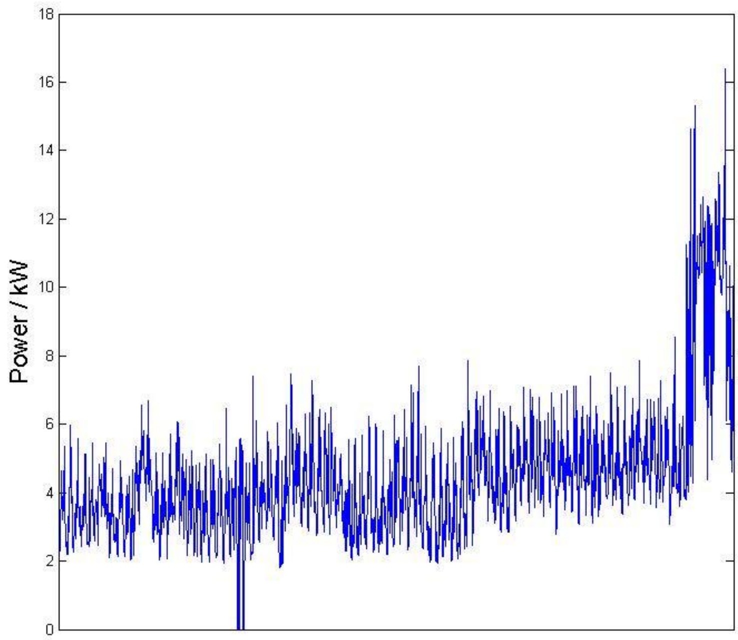


Figure 16: Time series data of power for the house of Figure 14.

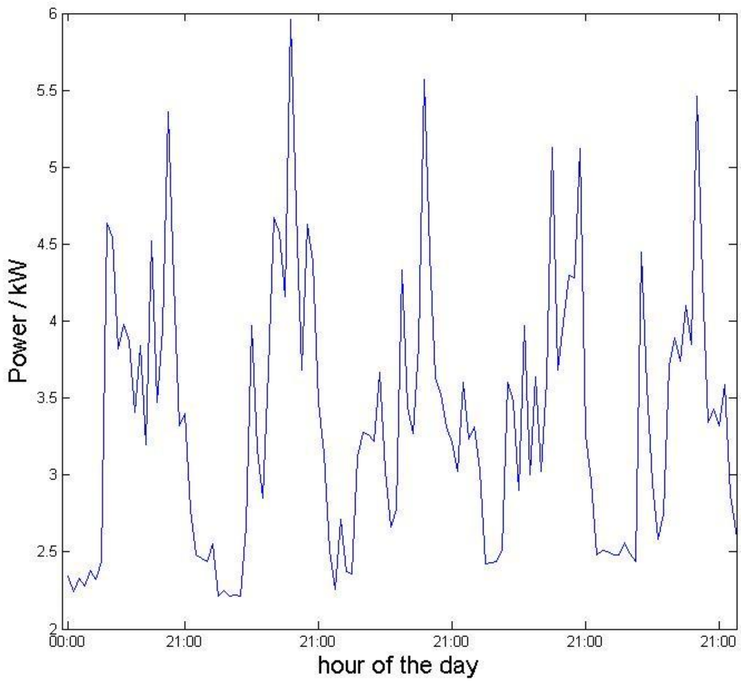


Figure 17: Time series data of power from the first five days of measurement for the house of Figure 14.

4 Possibilities for further study

Load prediction studied in this report can be enhanced by research on many aspects. As already noted, structural properties of houses affect the predictability. Question is how the prediction of energy consumption is dependent on these different properties such as age, insulation level, the heated volume of the house, heating systems and usage of the house.

Occupancy of the house also largely affects the prediction capabilities as the behaviour of the residents and the use of appliances can be very random. But for an empty house prediction can be rather accurate as desired indoor temperature level, power of the heating system to achieve this and heat losses are known or can be determined quite accurately. Houses are usually empty for many hours during daytime. And during holidays, for example, house can be devoid of residents for many days. Thus a system could be thought of that makes it possible for the residents to inform if the house is left empty for some longer period of time or if the house is uncommonly not empty during some day (in contrast to normal daytime when residents usually work or go to school). Pre-programmed in the system could be a control that lowers the indoor temperature level during the day if no information from the residents is received indicating that the house is not empty during the day. So for normal daytimes (when the house is left empty) retailers could predict the energy consumption based on the model for the house. If the house is going to be empty for a longer time, for example one week during holidays, residents could use the system to provide information about this. Providing information in advance about when the house is left empty again enhances the predictability, because at least during these times the energy consumption could be known quite accurately as noted above. System could also automatically lower the indoor temperature level for the duration of the “empty” time. Described system would require lot of research and field tests of the functionality and interface. Residents would also need some form of stimulus to use the

system, most likely money saving due to reduced electricity bill. Reference [6] describes smart metering based system that can be used to control loads and field tests of the system. Similar bi-directional approach could possibly be used to implement the system described in this section.
Also the following aspects need further studies:

- The impact of heat pumps on the load predictions.
- Clustering and removal of outlier houses to improve predictability.
- Prediction of heating needs and responses for dynamic load control.
- Prediction of loads for other types of houses, such as partially storing houses.
- The seasonal variations in the temperature dependency and the reasons for that.

5 Conclusions

The aim of this research was to study the effect of using temperature data for predicting energy consumption in electrically heated fully storing houses. The approach was kept as a simple data fitting task. The presented results clearly confirm two main conclusions. First, the prediction capability is improved by decreasing the time shift between temperature and prediction intervals to about 21 hours. As explained before this reduces the time when temperature variations can occur that we don’t take into account. Secondly, making predictions for a larger group of houses results in better prediction capability. This is in part due to the smoothing out of randomness in the behaviour of single houses and in part due to the possibility to detect differently behaving houses and remove them from the group predicted.
From the methods considered the 24 hours temperature interval gave the best correlation between mean temperature and energy consumption. Taking into account 48 hours mean or weighted mean temperatures brings no improvement to the predictions. Finding the weight coefficients for 2 days weighted mean temperature was also not an easy task for the Matlab-algorithm used and the results were sensitive to initial guess of the minimizing parameters.
Better correlation could likely be achieved with real, precise modelling based on physical principles, especially for houses that do not have full storage heating with adequate dimensioning. This of course presents itself as a challenging task because heat dynamics of a house is very complicated due to myriad of different heat capacities, nonlinearities and pathways for heat conduction. This is especially the case for partially storing or direct heating houses that lack the heat storage buffer. Multiple floors and unknown state of the doors add complexity. Also large amount of randomness is present in the heat dynamics of a house because of the changing behaviour of the residents and weather conditions, making construction of exactly accurate prediction models practically impossible.

References

[1] DOE, Smart Grids: An Introduction, available at http://energy.gov/oe/downloads/smart-grid-introduction-0, 3.4.2012.

[2] Koponen P., Seppälä J., Market price based control of electrical heating loads, CIRED 21st International Conference on Electricity Distribution, CIRED, Frankfurt (2011), Paper No 0796.

[3] Koponen, P., Porri O., Seppälä J., Plans for the second year field tests on smart metering based dynamic load control, VTT Report VTT-R-07979-11, 2012.

[4] Seppälä A., Load research and load estimation in electricity distribution, doctoral dissertation, VTT Publications 289, VTT Offsetpaino 1996.

[5] Koponen P., Measurements and models of electricity demand responses, VTT Report VTT-R-09198-11, 2011.

[6] Ljung L., Glad T., Modelling of dynamic systems, P T R Prentice Hall, 1994.