Technical University of Denmark

DTU

# Load forecasting for supermarket refrigeration

**Bacher, Peder; Madsen, Henrik; Aalborg Nielsen, Henrik**

*Publication date:*
2013

*Document Version*
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

**DTU Library**
Technical Information Center of Denmark

# Load forecasting for supermarket refrigeration

Peder Bacher, Henrik Madsen

DTU Compute, Technical University of Denmark, DK-2800 Lyngby, Denmark

Henrik Aalborg Nielsen

ENFOR A/S, DK-2970 Hørsholm, Denmark

May 15, 2013

# Contents

# 1  Introduction

This report presents a study of models for forecasting the load for supermarket refrigeration. The data used for building the forecasting models consists of load measurements, local climate measurements and weather forecasts. The load measurements are from a supermarket located in a village in Denmark. The load for refrigeration is the sum of all cabinets in the supermarket, both low and medium temperature cabinets, and spans a period of one year. As input to the forecasting models the ambient temperature observed near the supermarket together with weather forecasts are used. Every hour the hourly load for refrigeration for the following 42 hours is forecasted. The forecast models are adaptive linear time-series models which are fitted with a computationally efficient recursive least squares scheme. The dynamic relations between the inputs and the load is modeled by simple transfer functions. The system operates in two regimes: one in the closing hours during night and one in the opening hours during the day. This is modeled by a regime switching model in which some of the coefficients in the model depends on the regime. The results show that the one-step ahead residuals are close to white noise, however some dependence on the ambient temperature remains, which is caused by non-linearities in the relation between the load and the ambient temperature. Suggestions for including these non-linearities are given in the discussion of the results.

The report starts with a section in which the data and the NWPs are described. This is followed by a presentation of the modeling approach and the model identification, where a suitable forecasting model is found. Finally, the results are presented, and the method is discussed and conclusions are drawn.

# 2  Data

The measuring of load and ambient temperature were carried in EUDP-I ESO2 project. A detailed description can be found in (Fredslund, 2013). The numerical weather predictions (NWPs) are provided by the Danish Meteorological Institute (DMI). All times are in Central European Summer Time (CEST) and the time stamp for average values are set to the end of the averaging interval, i.e. for hourly values the time stamp is set to the end of the hour.

## 2.1  Refrigeration load observations

The refrigeration load is the sum all cabinets in the supermarket. The low temperature cabinets are listed in Table 1 and the medium temperature cabinets are listed in Table 2. The ambient temperature is measured at the location as described in (Fredslund, 2013). The load and temperature are measured every minute and resampled to hourly values average values. The hourly load is in kW and denoted by

$$\{Q_t;\ t = 1, \ldots, N\} \tag{1}$$

| Type | Model | Power (kW) |
|---|---|---|
| Frozen basin with glass lid | Malmoe 3 | 1.05 |
| Frozen basin with glass lid | Malmoe 3 | 1.05 |
| Frozen basin with glass lid | Malmoe 3 | 0.68 |
| Frozen storage | ECO STE 32BL7 ED | 4.8 |

Table 1: Low temperature cabinets.

| Type | Model | Power (kW) |
|---|---|---|
| Milk cooling room | ECO STE 32BL7 | 5.5 |
| Milk cooling room | ECO MTE 25L7 | 2.81 |
| Cooling rack | Lisbona | 5.77 |
| Cooling rack | Lisbona | 5.77 |
| Cooling rack | Lisbona | 3.84 |
| Vegetable cooler | Lisbona | 4.6 |
| Cool basin with glass lid | Malmoe 3 | 1.27 |

Table 2: Medium temperature cabinets.

and the ambient temperature is in °C and denoted by

$$\left\{ T_t^{\text{a,obs}} \; ; \; t = 1, \ldots, N \right\} \tag{2}$$

$$\tag{3}$$

where $N = 9024$.

Time series plots of the load and ambient temperature observations can be seen in Figure 1. Quite large gaps where the observations are not available can be seen. Clearly the load is lower in the winter period.

A time series plot of the observations from January can be seen in Figure 2. It is very clearly seen that the system operates in two modes: at nighttime the load is lower than at daytime. One of the reasons for this is that at nighttime the supermarket is closed and the open cabinets are covered. Furthermore, some systematic peaks can are found mostly in the beginning of the opening hours, which are related to defrost of the low temperature cabinets. The defrost occur automatically two times a day at the same two times of the day, except for weekend where the evening defrost occur later, for information about the system see Fredslund (2013). A times series plot of the observations over July can be seen in Figure 3. The difference between the load at night and day is here even higher, since the air temperature in the supermarket is varying more in the opening hours. The opening hours are all days 8 to 21.

## 2.2 Numerical weather predictions

The numerical weather predictions (NWPs) used for the forecasting are provided by the Danish Meteorological Institute. The NWP model used is DMI-HIRLAM-S05, which has a 5 kilometer grid and 40 vertical layers (DMI, 2011). The NWPs consist of time series of hourly values for climate variables, which are updated four times per day and have a 4 hour calculation delay (e.g. the forecast starting at 00:00 is available at 04:00). Since a new two-day load forecast is calculated every hour, then
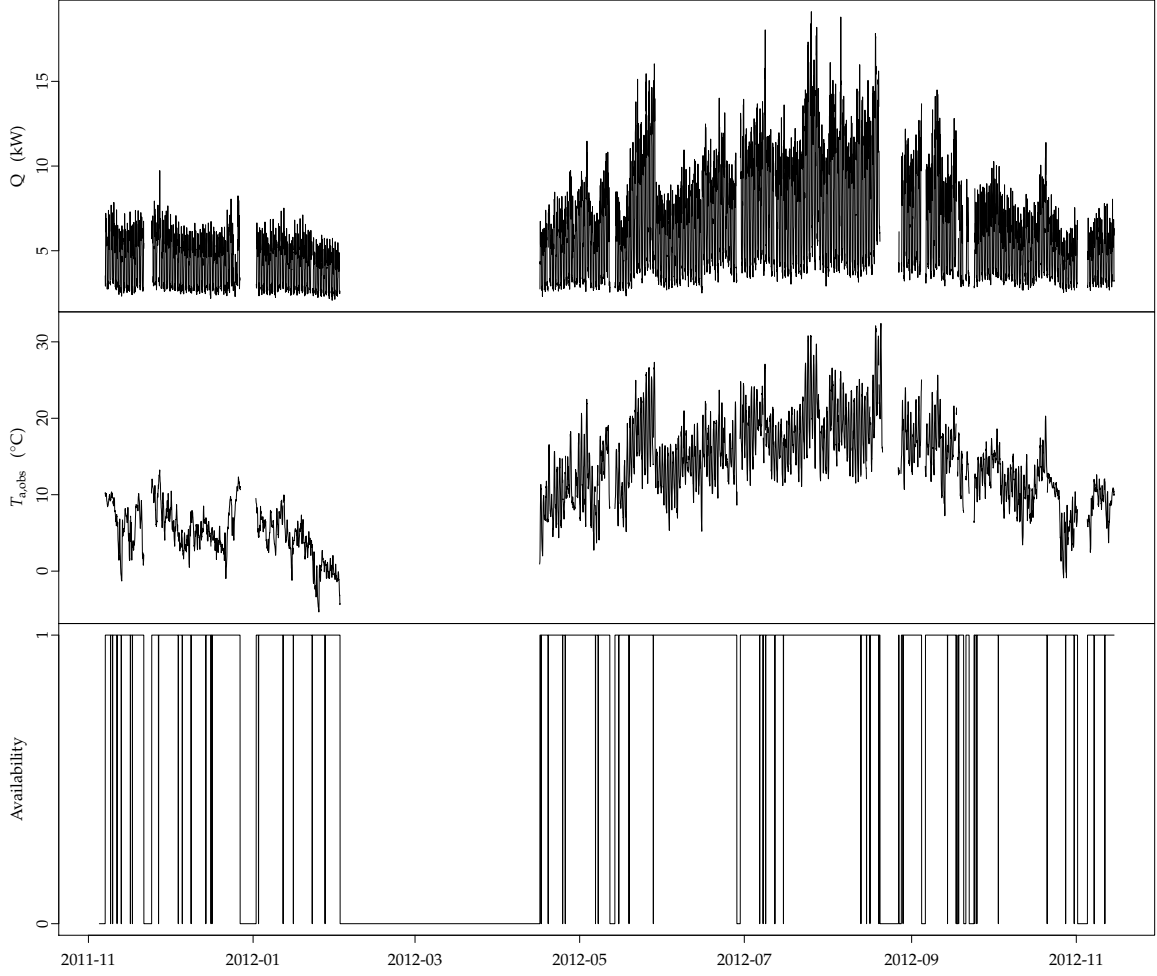
Figure 1: Times series plots of the hourly measurements for the entire period. The availability is 1 if a sample is available and otherwise 0. The ticks on the time axis are 00:00 on Mondays.

- in order to use the latest available information - every hour the latest available NWP value for the $k$'th horizon at time $t$ is picked as

$$\text{Ambient temperature } (^\circ \text{C}): \quad T^{\text{a,nwp}}_{t+k|t} \tag{4}$$

$$\text{Global radiation } (\text{W}/\text{m}^2): \quad G^{\text{nwp}}_{t+k|t}$$

$$\text{Wind speed } (\text{m}/\text{s}): \quad W^{\text{s,nwp}}_{t+k|t}$$

## 2.3 Combining local observations with NWPs

To include the dynamics of the system in an efficient way, the inputs are low-pass filtered as explained in Section 3.3. Hence, for the forecast calculated at time $t$, past values of the inputs are being used. In order to use the information embedded in the local measurements they are combined with the NWPs. The combining is carried out by forming the time series for each of the inputs at time $t$, for a specific horizon
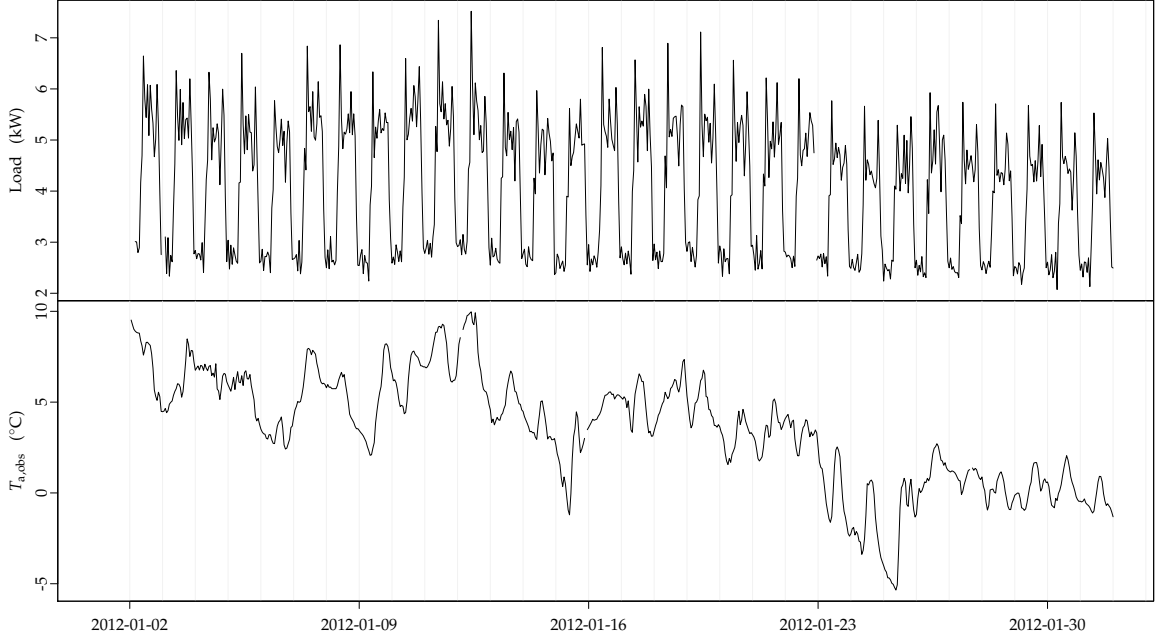
5

Figure 2: Times series plots of the hourly observations for January 2012.

$k$, by

$$\left\{ T^{\mathrm{a}}_{t+k|t} \right\} = \left\{ \dots, T^{\mathrm{a,obs}}_{t-1}, T^{\mathrm{a,obs}}_{t}, T^{\mathrm{a,nwp}}_{t+1|t}, T^{\mathrm{a,nwp}}_{t+2|t}, \dots, T^{\mathrm{a,nwp}}_{t+k|t} \right\}$$

$$\left\{ G_{t+k|t} \right\} = \left\{ \dots, G^{\mathrm{nwp}}_{t-1}, G^{\mathrm{nwp}}_{t}, G^{\mathrm{nwp}}_{t+1|t}, G^{\mathrm{nwp}}_{t+2|t'}, \dots, G^{\mathrm{nwp}}_{t+k|t} \right\} \tag{5}$$

$$\left\{ W^{\mathrm{s}}_{t+k|t} \right\} = \left\{ \dots, W^{\mathrm{s,nwp}}_{t-1}, W^{\mathrm{s,nwp}}_{t}, W^{\mathrm{s,nwp}}_{t+1|t}, W^{\mathrm{s,nwp}}_{t+2|t}, \dots, W^{\mathrm{s,nwp}}_{t+k|t} \right\}$$

Notice that local observations are available only for the ambient temperature and that for the others the most recent NWPs are used for past values instead.

## 3   Models

The applied models are similar to the models used in (Nielsen and Madsen, 2006) for forecasting of the summed total heat load for many houses and in (Bacher et al., 2012) for forecasting of the heat load for single family houses. The models are based on prior physical knowledge of the heat dynamics, which in combination with statistical time series models, forms a grey-box modeling approach. This allows for inclusion of heat transfer effects related to the climate variables in combination with a time adaptive estimation scheme applied to meet changing condition. Furthermore, in order to describe of systematic patterns in the load a diurnal curve and regime switching models are applied. The forecasting models are fitted by optimizing a few parameters to minimize the root mean square error (RMSE) in an off-line setting. The fitting is carried out separately for each horizon $k$, which means that the same model formulation - i.e. same inputs and model structure - is used, only the parameter values for each horizon vary.
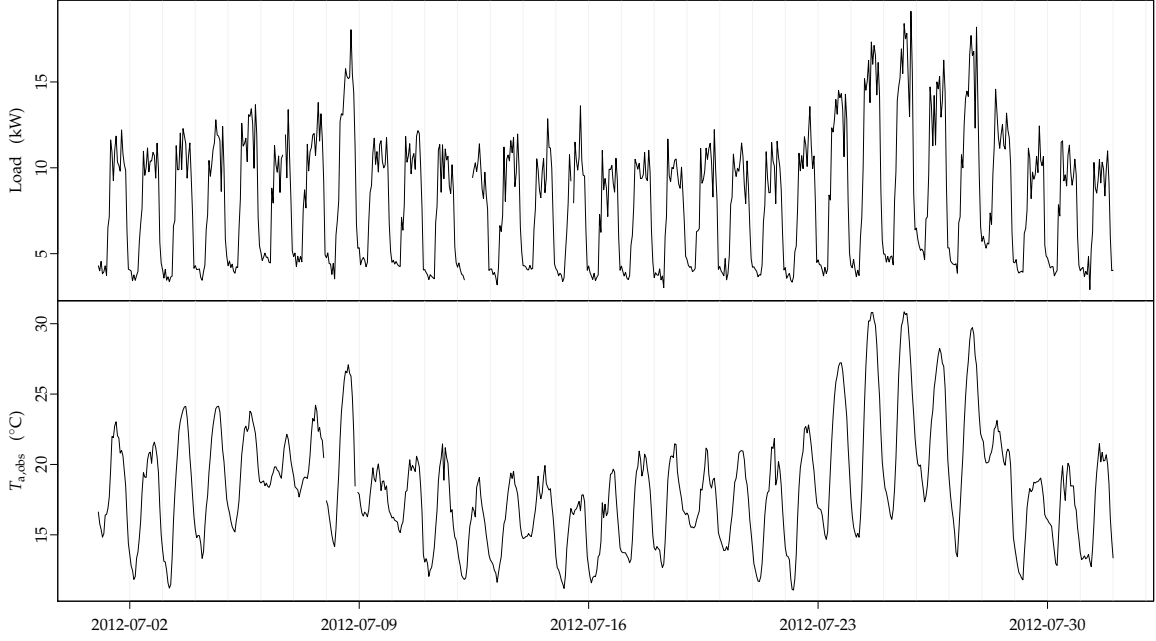
6

Figure 3: Times series plots of the hourly observations for July 2012. The ticks on the time axis are 00:00 on Mondays.

## 3.1 Time adaptive models

The models are fitted with the *k*-step recursive least squares scheme described in Bacher et al. (2009). This means that the coefficients in the model can change over time and adapt optimally, in a least squares sense, to changing conditions. The coefficients are recursively updated, which means that only a few matrix operations are required to compute an updated forecast, hence the scheme is computationally very fast. It is a recursive implementation of a weighted least squares estimation, where the weights are exponentially decaying over time. A single parameter is required, the forgetting factor $\lambda$, which determines how fast input data is down-weighted. The weights are equal to

$$w(\Delta t) = \lambda^{\Delta t} \tag{6}$$

where $\Delta t$ is the age of the data in hours. This implies that for $\lambda = 0.98$ the weights are halved in 34 hours, for $\lambda = 0.995$ they are halved in 138 hours ($\sim$ 6 days) and for $\lambda = 0.999$ in 693 hours ($\sim$ 29 days).

## 3.2 Diurnal curve

A diurnal curve is included in the models for describing systematic diurnal patterns in the load. The curve is modeled as a harmonic function using a Fourier series

$$\mu\left(t_{\text{tod}}, n_{\text{har}}, \alpha_{\text{diu}}\right) = \sum_{i=1}^{n_{\text{har}}} \alpha_{i,1}^{\text{diu}} \sin\left(\frac{t_{\text{tod}} i \pi}{12}\right) + \alpha_{i,2}^{\text{diu}} \cos\left(\frac{t_{\text{tod}} i \pi}{12}\right) \tag{7}$$

where $t_{\text{tod}}$ is the time of day in hours at time $t$, $n_{\text{har}}$ is the number of harmonics included in the Fourier series and $\alpha_{\text{diu}}$ is a vector consisting of the coefficients for

7

the included harmonics. For all the applied models a curve is fitted for working days and another curve for weekends.

## 3.3 Low-pass filtering for modeling of heat dynamics

The heat dynamics of a thermal system can be described by lumped parameter RC-models, see for example (Madsen and Holst, 1995), (Braun and Chaturvedi, 2002) and (Jiménez et al., 2008). As described by Nielsen and Madsen (2006) the response in the load to changes in the climate variables can be modeled with a rational transfer function, which is a description with an RC-model of the low-pass filtering effect through the system. In the applied models the simplest first order low-pass filter, with a stationary gain equal to one, is used. This is a model of the heat dynamics formed by an RC-model with a single resistance and a single capacitor. As an example the transfer function from the ambient temperature to the heat load is

$$Q_t = \alpha_a H_a(q) T_t^a \tag{8}$$

where

$$H_a(q) = \frac{1 - a_{T_a}}{1 - a_{T_a} q^{-1}} \tag{9}$$

and where $q^{-1}$ is the backward shift operator ($q^{-1}x_t = x_{t-1}$) (see (Madsen, 2007)), $\alpha_a$ is the stationary gain from the ambient temperature to load and $a_{T_a} \in [0, 1]$ is a parameter, which is corresponding to the time constant for the part of the system affected by changes in ambient temperature. A system with a high thermal mass and good insulation will have a relatively high $a_{T_a}$, hence the filter parameter needs to be tuned for each system in order to describe the dynamics properly. First order low-pass filters are also applied for wind speed and global radiation, with the filter parameter tuned to match the response of the system to each effect separately.

## 3.4 Parameter optimization

As described above several parameters need to be optimized for each horizon. The optimization is carried out in an off-line setting by minimizing the RMSE for each horizon $k = 1, \ldots, 42$ separately. A simple bisectioning scheme is applied for the optimization, since this allows for performing a filtering of the inputs only once for parameter values in a given range. Then these series can be used for optimization for all the horizons. The properties of the optimization is not studied in detail in this work.

The following parameters are optimized:

- The forgetting factor: $\lambda$,

- The number of harmonics in the diurnal curve: $n_{\text{har}}$,

- The coefficients for input low-pass filters: $a_{T_a}$, $a_G$ and $a_{W_s}$.
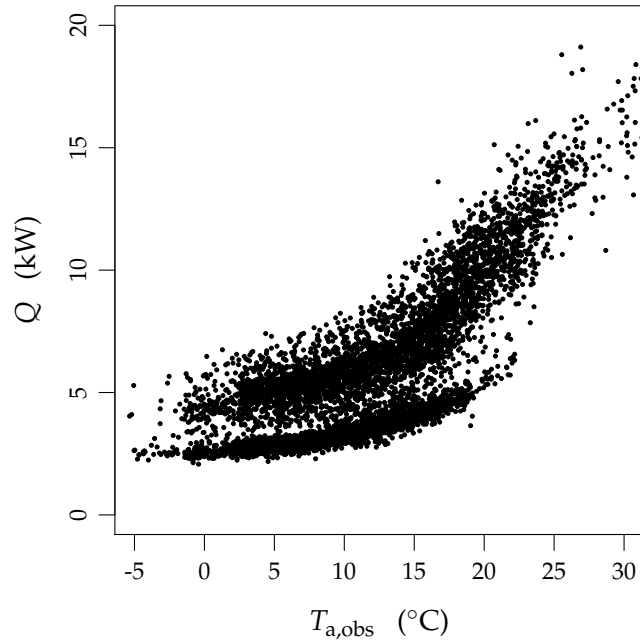
Figure 4: Scatter plot of the load versus the ambient temperature for the entire period.

# 4 Closing and opening hours operation regimes

In this section the differences between the operation of the system in the closing hours (at night time) and the opening hours are outlined. The result of the analysis is that it is found that the effect of the ambient temperature on the load is different in the two regimes and therefore that a regime-switching model should be applied. A scatter plot of the load versus the measured ambient temperature is shown in Figure 4. A clear dependence of the ambient temperature is found and furthermore it seems to be non-linear. The difference between the opening and closing hours operation can also clearly be seen, as two point clouds above each other. In Figure 5 a similar scatter plots are found, divided into the winter period and into the summer period. This reveals that the dependence is more linear when considered local in time, however the large gap from 1'st of February 2012 to around 15'th of April could hold a period where this separation is less clear. 5.

## 4.1 Separation into periods of closing and opening hours

In order to use a regime switching model it is necessary to predict which regime the system operates in at a given time $t + k$. In the present work this is simply carried out by using the time of day, since the opening and closing hours are known for the supermarket. The opening hours are all days 8 to 21. The best separation is achieved by separating the period from 7 to 22 as the opening hours. The scatter plots in Figure 6 show the result of the separation. It can be seen that some overlap does occur, i.e. some measurements from opening hours have a low level corresponding to the prevailing level of the opening hours, and the other way around.
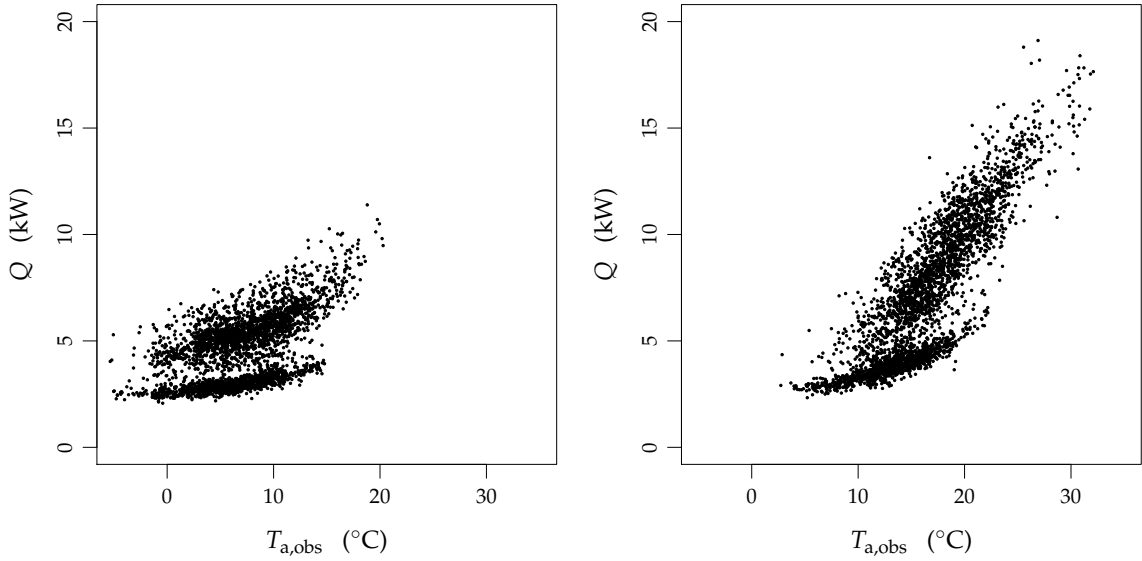
9

Figure 5: Scatter plots of the load versus the ambient temperature. The left plot is the winter period from 1'st of October to 1'st of May. The right plot is the summer period.

# 5 Model identification

In this section the model identification is presented. A forward model selection approach is applied, by first fitting a simple model not using the any NWP inputs and then expanding this until no clear improvement is found.

Forecasting models, which include different types of effects related to the climate variables, are applied in order to identify which of the inputs are important to include. Furthermore, it is tried whether the diurnal curve should be different in weekends from workdays. The model is defined as

$$Q_{t+k} = \widehat{Q}_{t+k|t} + e_{t+k} \tag{10}$$

where $\widehat{Q}_{t+k|t}$ is different among the model evaluated in the model selection.

## 5.1 Root mean square error evaluation

To evaluate the models the root mean square error (RMSE) for the $k$'th horizon

$$\text{RMSE}_k = \left( \frac{1}{N} \sum_{t=1}^{N} e_{t+k}^2 \right)^{\frac{1}{2}} \tag{11}$$

is used.

It is noted that the period before the 15'th of November 2011 is used as a burn-in period and this period is excluded from the $\text{RMSE}_k$ calculation. And furthermore the period from 2011-12-24 to 2012-01-02 are removed due unusual operation and errorfull data, the period from 2012-04-10 to 2012-04-18 is removed since this is immediately after the large gap spanning 2 months.
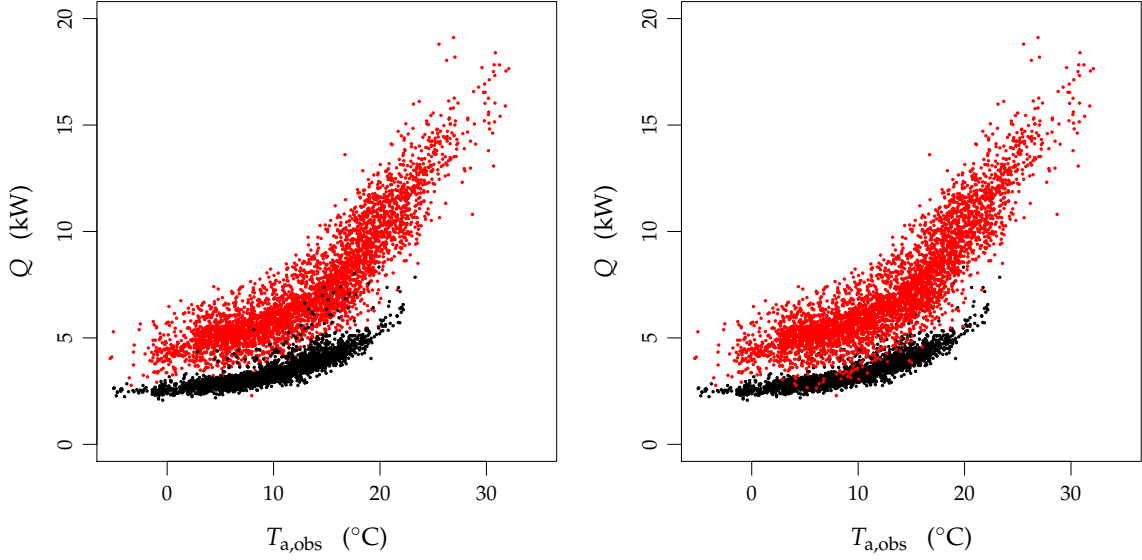
10

Figure 6: Scatter plots of the load versus the ambient temperature for showing the splitting with opening hours. The black points marks the load values which are in the closing hours and the red points marks values which are in the opening hours. In the left plot the closing hour values are plotted on top of the opening hour values and in the right plot it is the opening hour values are on top.

## 5.2 Simplest model

The simplest model considered, denoted by $Model_D$, includes no external input, it simply has an intercept and a diurnal curve

$$\widehat{Q}_{t+k|t} = \alpha_i + Q_d \tag{12}$$

where

$$Q_d = \mu(t_{tod}, n_{har}, \alpha_{diu}) \tag{13}$$

## 5.3 First step in model selection

In the first step the simplest model is expanded by adding the ambient temperature as input, which leads to $Model_{D.A}$

$$\widehat{Q}_{t+k|t} = \alpha_i + Q_d + Q_a \tag{14}$$

where

$$Q_a = \alpha_a H_a(q) T^a_{t+k|t} \tag{15}$$

The $H_a(q)$ is the low-pass filter describing the heat dynamics of the system, i.e. the response of in load to changes in ambient temperature. In Figure 7 the $RMSE_k$ for $Model_D$ and $Model_{D.A}$ is plotted. The improvement from using the ambient temperature as input is very clear for all horizons, hence $Model_{D.A}$ is preferred over $Model_D$ and kept for further expansion in the following step.
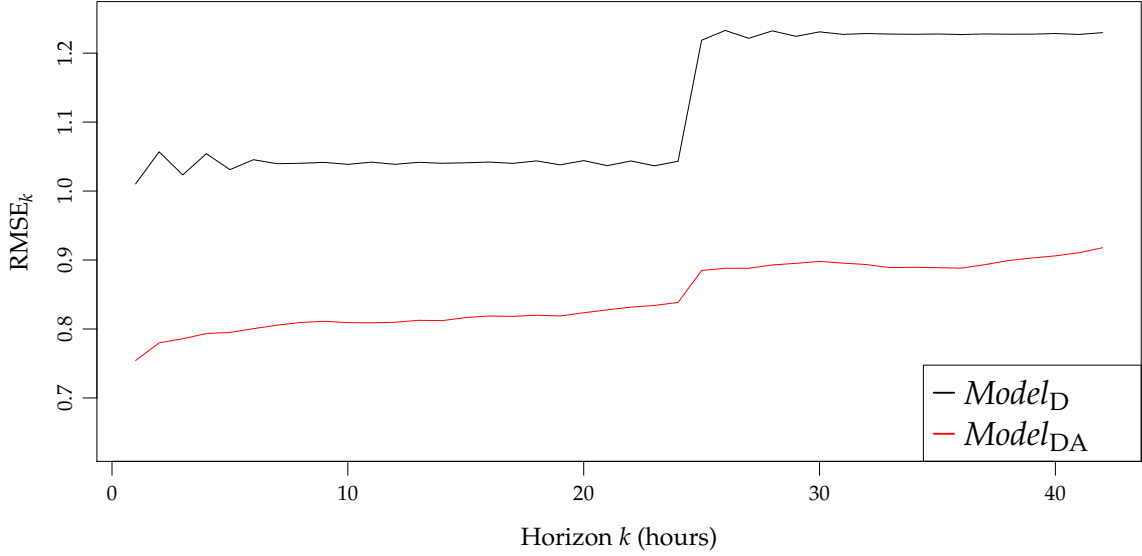
Figure 7: The RMSE$_k$ for the simplest $Model_D$ and for $Model_{D.A}$, the latter using the ambient temperature as input, which are evaluated in step 1 of the model selection.

## 5.4 Second step in model selection

In the second step of the model selection several extensions to $Model_{D.A}$ are evaluated:

- $Model_{D.A.G}$

$$\widehat{Q}_{t+k|t} = \alpha_i + Q_d + Q_a + \alpha_g H_g(q) G_{t+k|t} \tag{16}$$

  where the effect from solar radiation is included by letting the global radiation enter through a low-pass filter, which describes the dynamic response from the global radiation to the load.

- $Model_{D.A.Ws}$

$$\widehat{Q}_{t+k|t} = \alpha_i + Q_d + Q_a + \alpha_{ws} H_w(q) W^s_{t+k|t} \tag{17}$$

  where the effect from wind speed is included by letting the wind speed enter through a low.pass filter.

- $Model_{WD.A}$

$$\widehat{Q}_{t+k|t} = \alpha_i + Q_{wd} + Q_a \tag{18}$$

  where one diurnal curve is used for workdays and another for weekends

$$Q_{wd} = \begin{cases} \mu(t_{tod}, n_{har}, \alpha_{workday}) & \textit{for } \text{work days} \\ \mu(t_{tod}, n_{har}, \alpha_{weekend}) & \textit{for } \text{weekends} \end{cases} \tag{19}$$

12
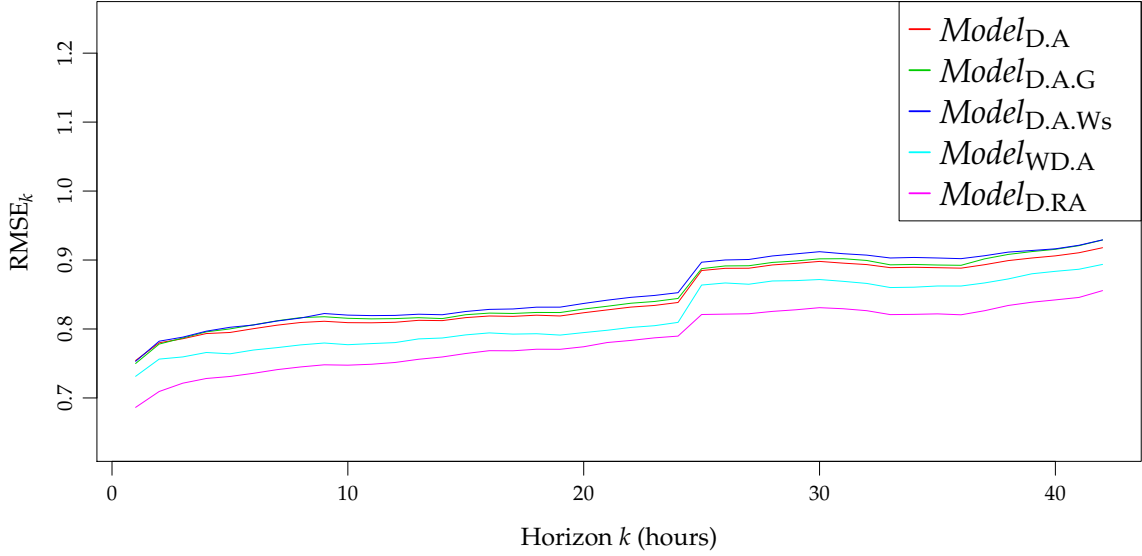
Figure 8: The $\text{RMSE}_k$ for the models evaluated in step 2.

- $Model_{\text{D.RA}}$

$$\widehat{Q}_{t+k|t} = Q_{\text{d}} + Q_{\text{ra}} \tag{20}$$

where two regimes are used for the ambient temperature effect

$$Q_{\text{ra}} = \begin{cases} \alpha_{\text{i,close}} + \alpha_{\text{a,close}} H_{\text{a}}(q) T_{t+k|t}^{\text{a,close}} & \textit{for} \;\; \text{closing hours} \\ \alpha_{\text{i,open}} + \alpha_{\text{a,open}} H_{\text{a}}(q) T_{t+k|t}^{\text{a,open}} & \textit{for} \;\; \text{opening hours} \end{cases} \tag{21}$$

note that the intercept is here included in $Q_{\text{ra}}$ since it is also different in the two regimes. The periods for opening and closing hours are separated as described in Section 4.

The $\text{RMSE}_k$ for these extensions are plotted in Figure 8. It is seen that the inclusion of global radiation and wind speed decreases the performance of the model, but that the use of different diurnal curves for workdays and weekends gives some improvement and that the use of two regimes for the effect of ambient temperature gives the highest improvement. Therefore $Model_{\text{D.RA}}$ is selected for further extension.

## 5.5 Third step in model selection

In the third step of the model selection the following extensions to $Model_{\text{D.RA}}$ are evaluated:

- $Model_{\text{D.RA.G}}$

$$\widehat{Q}_{t+k|t} = Q_{\text{d}} + Q_{\text{ra}} + \alpha_{\text{g}} H_{\text{g}}(q) G_{t+k|t} \tag{22}$$

where the effect from solar radiation is included. by letting the global radiation enter through a low-pass filter, which describes the dynamic response from the global radiation to the load.
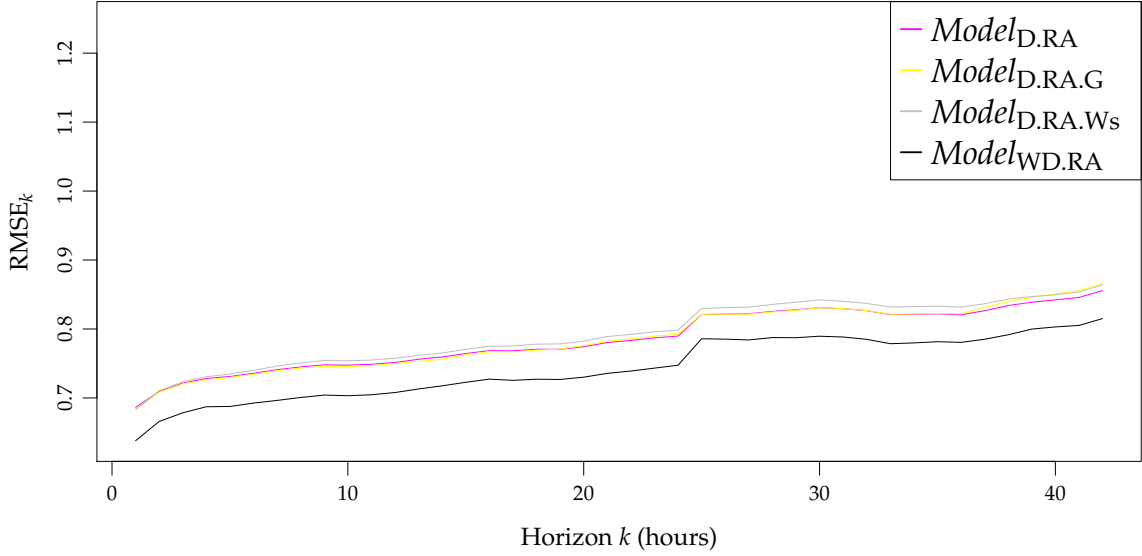
13

Figure 9: The RMSE$_k$ for the models evaluated in step 3.

- $Model_{\text{D.RA.Ws}}$

$$\widehat{Q}_{t+k|t} = Q_{\text{d}} + Q_{\text{ra}} + \alpha_{\text{ws}}H_{\text{w}}(q)W^{\text{s}}_{t+k|t} \tag{23}$$

where the effect from wind speed is included by letting the wind speed enter through a low.pass filter.

- $Model_{\text{WD.RA}}$

$$\widehat{Q}_{t+k|t} = Q_{\text{wd}} + Q_{\text{ra}} \tag{24}$$

where one diurnal curve is used for workdays and another for weekends as described in Equation (19).

The RMSE$_k$ for these extensions are plotted in Figure 9. It is seen that neither inclusion of global radiation nor wind speed improves the forecasting performance, but that the inclusion of a weekend diurnal curve clearly provides an improvement. Therefore $Model_{\text{WD.RA}}$ is kept for further extension.

## 5.6   Fourth step in model selection

In the fourth step of the model selection the following extensions to $Model_{\text{WD.RA}}$ are evaluated:

- $Model_{\text{WD.RA.G}}$

$$\widehat{Q}_{t+k|t} = Q_{\text{wd}} + Q_{\text{ra}} + \alpha_{\text{g}}H_{\text{g}}(q)G_{t+k|t} \tag{25}$$

where the effect from solar radiation is included. by letting the global radiation enter through a low-pass filter, which describes the dynamic response from the global radiation to the load.
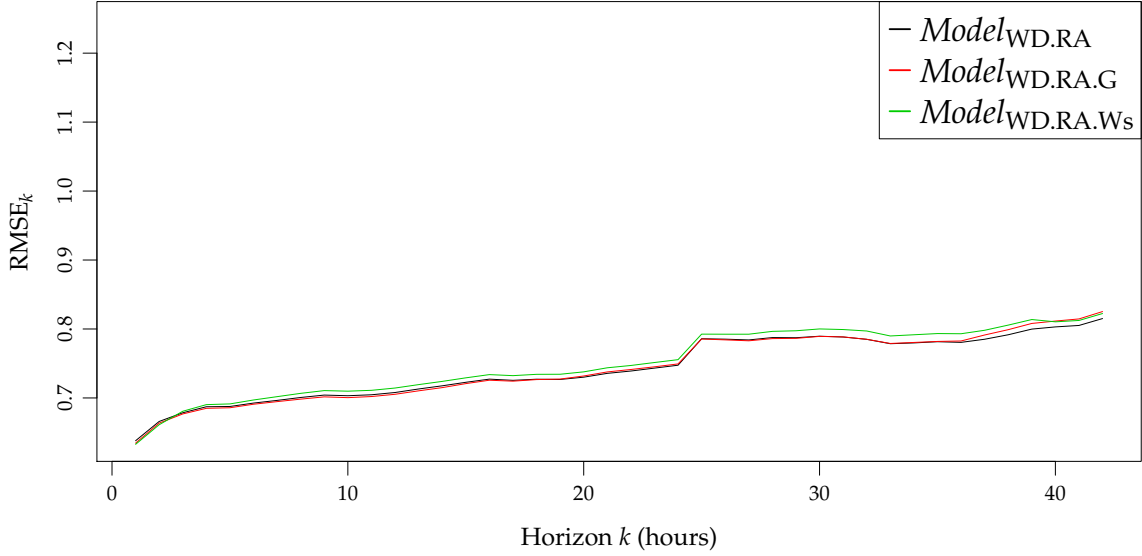
14

Figure 10: The $\text{RMSE}_k$ for the models evaluated in step 4.

- $Model_{\text{WD.RA.Ws}}$

$$\widehat{Q}_{t+k|t} = Q_{\text{wd}} + Q_{\text{ra}} + \alpha_{\text{ws}}H_{\text{w}}(q)W^{\text{s}}_{t+k|t} \tag{26}$$

where the effect from wind speed is included by letting the wind speed enter through a low.pass filter.

None of the extensions improves the performance, hence the model selection is ended and $Model_{\text{WD.RA}}$ is selected as the most suitable model and the forecasting results achieved with this model are analyzed in the following.

# 6  Results

In this section the results from forecasting with the selected model are presented and discussed. First the parameters, which are fitted in an off-line setting, are reported and then the forecasts are analyzed and discussed.

## 6.1  Model parameters

The parameters, which are fitted in an off-line setting, are listed in Section 3.4. The optimized values for each horizon are plotted in Figure 11. The optimized $\lambda$ values are between 0.993 and 0.995 which correspond to halving time of the weights from 99 hours ($\sim$ 4 days) to 138 hours ($\sim$ 6 days) respectively. The number of harmonics are for all horizons 11, which indicates that high frequencies are needed in the diurnal curve. The coefficient of the low-pass filter for the ambient temperature $a_{T_{\text{A}}}$ varies between 0.2 to 0.6, which indicates a fast dynamical response of the system with a time constant below one hour.
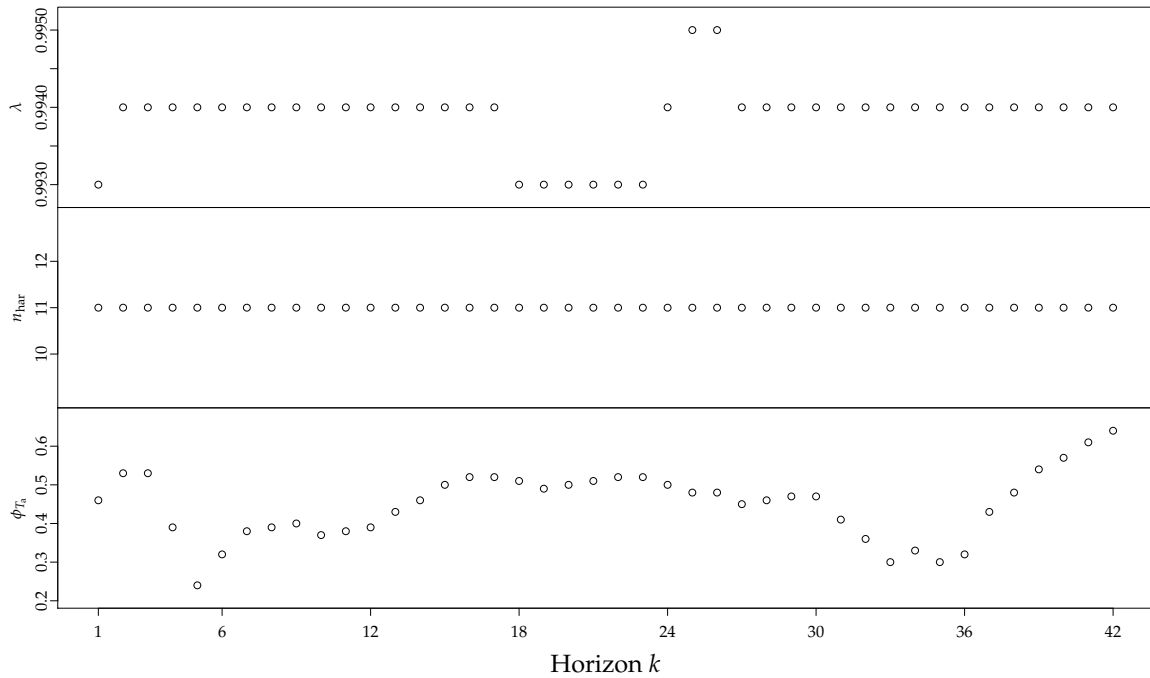
Figure 11: The model parameters fitted in an off-line setting as a function of horizon $k$.

## 6.2 Forecasting performance

In this section the forecasts are evaluated. This is carried out by analyzing the residuals for one-step (i.e. $k = 1$ hour) ahead forecasts and the coefficients from the selected model. The one-step ahead residuals are furthermore compared with the $k = 36$ hours ahead residuals to analyze for forecast performance for the longer horizons.

Plots of the one-step ahead residuals, cumulated residuals and the forecasted ambient temperature used as input are found in Figure 12. It is seen that the level of the residuals are generally lower in the winter and in the summer the level of the residual are different in some periods. The cumulated residuals have no trends in the winter period indicating that the forecasting performance is close to optimal in winter conditions. During the summer a clear increasing trend is indicating that some effects are not described by the model and hence there is a potential for further improvement of the model.

The auto-correlation function (ACF), histogram and QQ-normal plot of the one-step ahead residuals are shown in Figure 13. It is clearly seen from the ACF that some correlation is left for the first lags. This correlation could be removed with a noise model as applied in Nielsen and Madsen (2006), which would improve the forecasting performance for the shortest horizons (up to around $k < 6$). The histogram shows a non-skewed bell formed distribution of the residuals, indicating that the residuals are normal distributed, however the QQ-plot reveals that the tails of the distribution are a bit heavy compared to the normal distribution.
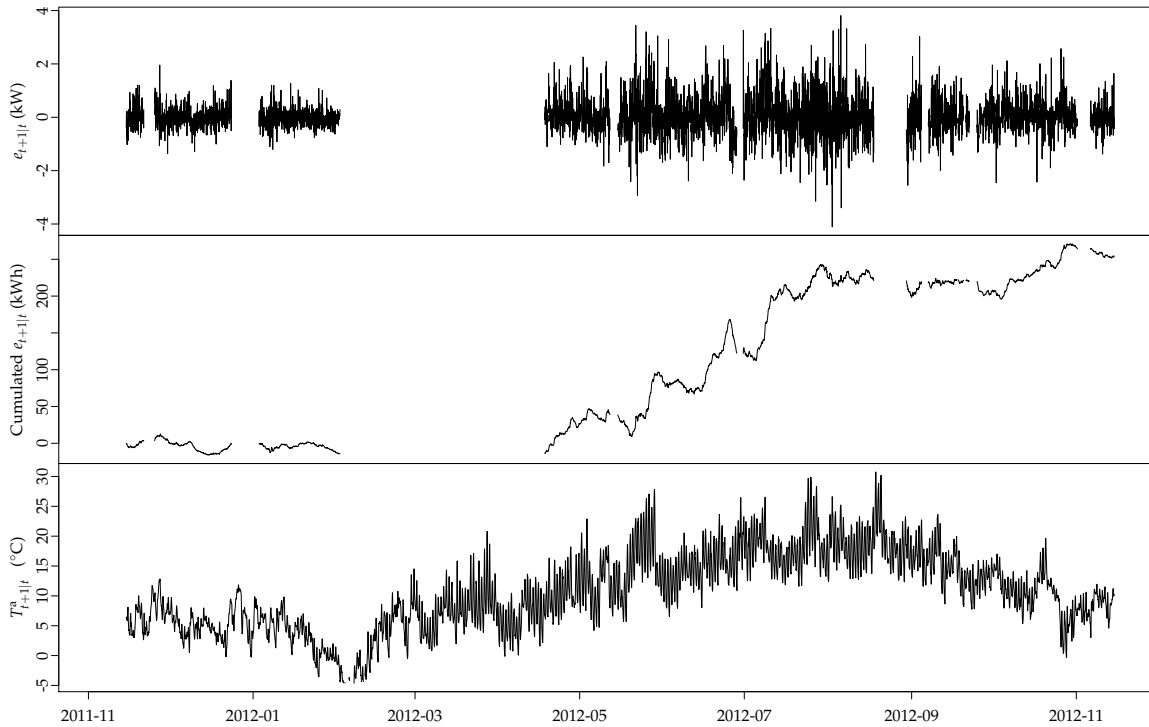
16

Figure 12: Plot of the one-step ahead residuals for the entire period. The upper plot is the residuals, the middle plot is the cumulated residuals and the lower plot is the one-step forecasted ambient temperature used as input.

In order to focus on the performance in the summer period the one-step ahead residuals for July is plotted in Figure 14 together with the regime switching inputs and fitted coefficients for the ambient temperature inputs. The plots reveal that the coefficient estimates of the inputs varies depending on the changing conditions related to the ambient temperature. This is most clearly seen for coefficient estimates for the opening hours (the two lower plots), where the intercept decrease and the slope increase for warm days compared to the colder days, for example on the 8'th of July. This shows how the model adapts to changing conditions over time, hence is a linear approximation local in time of the non-linear effects in the ambient temperature, however it also clear that the conditions change much faster than described by the model.
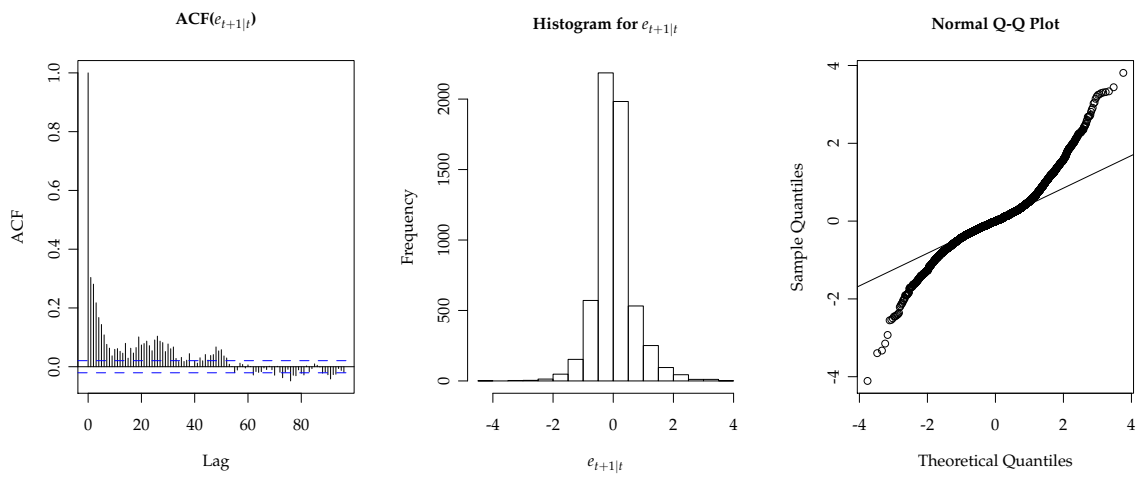
Figure 13: The auto-correlation function, histogram and QQ-normal plot of the one-step ahead residuals.
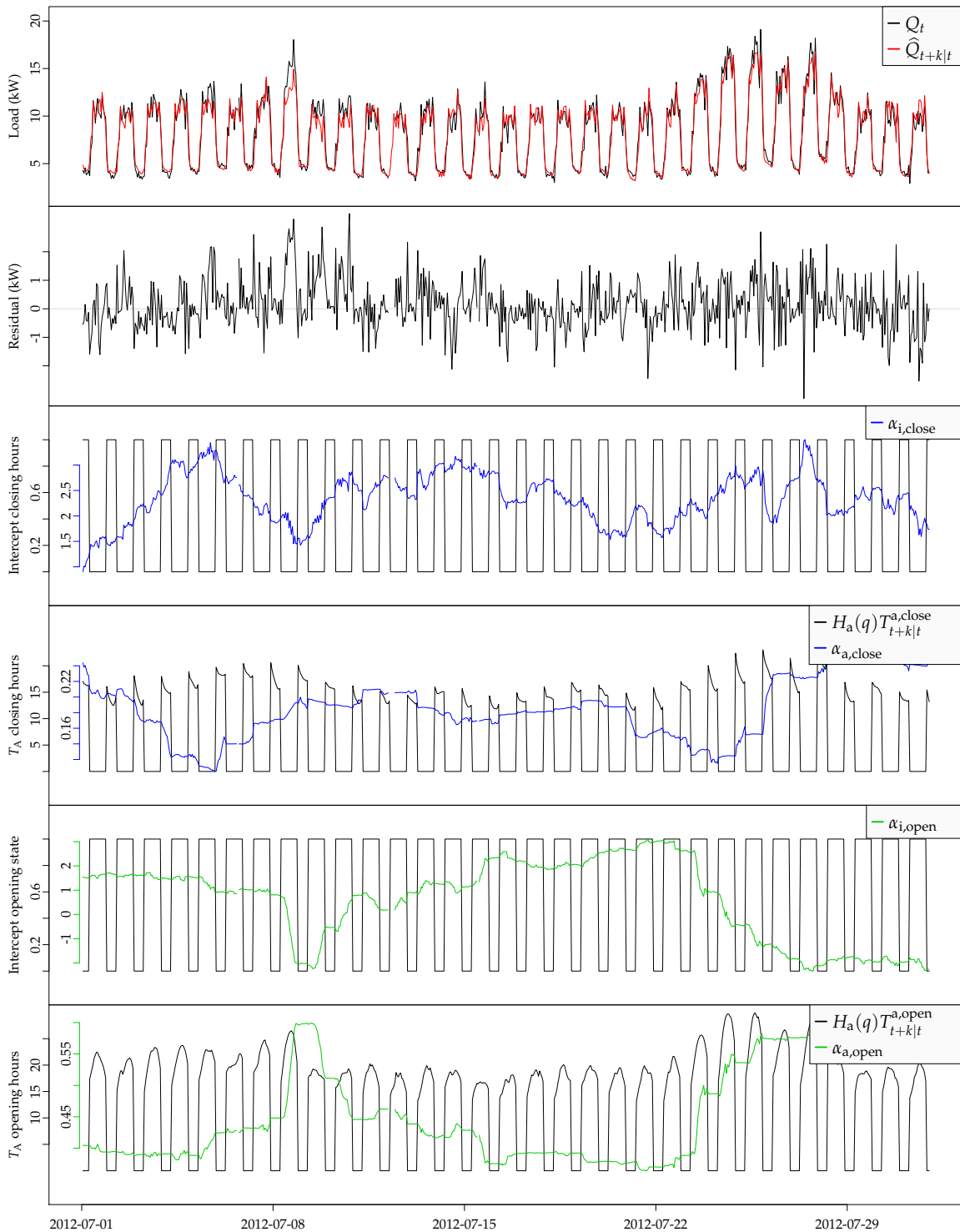
Figure 14: The one-step ahead residuals for July. The upper plot is the measured and forecasted load and below this a plot of the residuals. In the two plots in the middle are first a plot of the intercept input and estimated coefficient and below this a plot of the ambient temperature input and estimated coefficient, both for the closing hours. The lower two plots are similar, but for the opening hours.

In order to analyze the forecasting performance for the longer horizons the $k = 36$ hours ahead residuals are plotted in Figure 15 together with the regime switching inputs and fitted coefficients for the ambient temperature input. The plots are similar to the plots for the one-step ahead residuals in Figure 14 and the same conclusions are drawn, namely that the non-linear effect in the ambient temperature input are not described sufficiently with the model, which is here even more clear. Furthermore on this longer horizon it is clearly seen that the model can only adapt to the changing conditions with a delay, equivalent to the horizon, hence for the longer horizons the non-linear effects not included in the model are more clearly seen.

The analysis of the residuals are finalized with a plot in Figure 16 of the $k = 1$ and $k = 36$ ahead residuals versus the ambient temperature input for the warm period from 1'st of May to 1'st of September. Smoothed local kernel regression estimates calculated with `loess()` in R is added on top as red lines. The plots reveals that the prediction error increase with the ambient temperature and the smoothed estimates reveals that the residuals are biased for levels of the ambient temperature above approximately $20\,^{\circ}\mathrm{C}$. This indicates that the description of the non-linear effect from ambient temperature needs to be enhanced in the model.
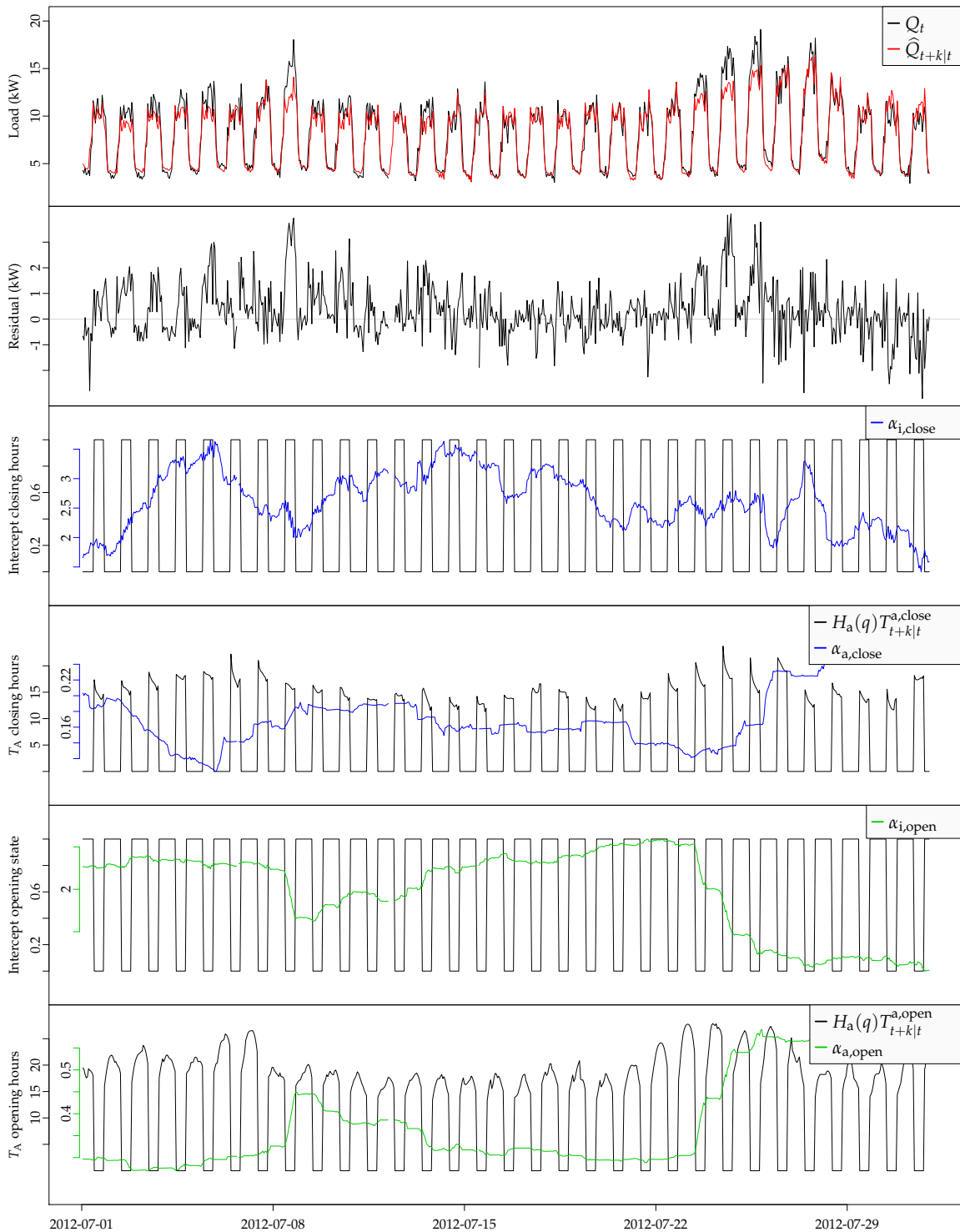
Figure 15: The $k = 36$ hours ahead residuals for July. The upper plot is the measured and forecasted load and below this a plot of the residuals. In the two plots in the middle are first a plot of the intercept input and estimated coefficient and below this a plot of the ambient temperature input and estimated coefficient, both for the closing hours. The lower two plots are similar, but for the opening hours.
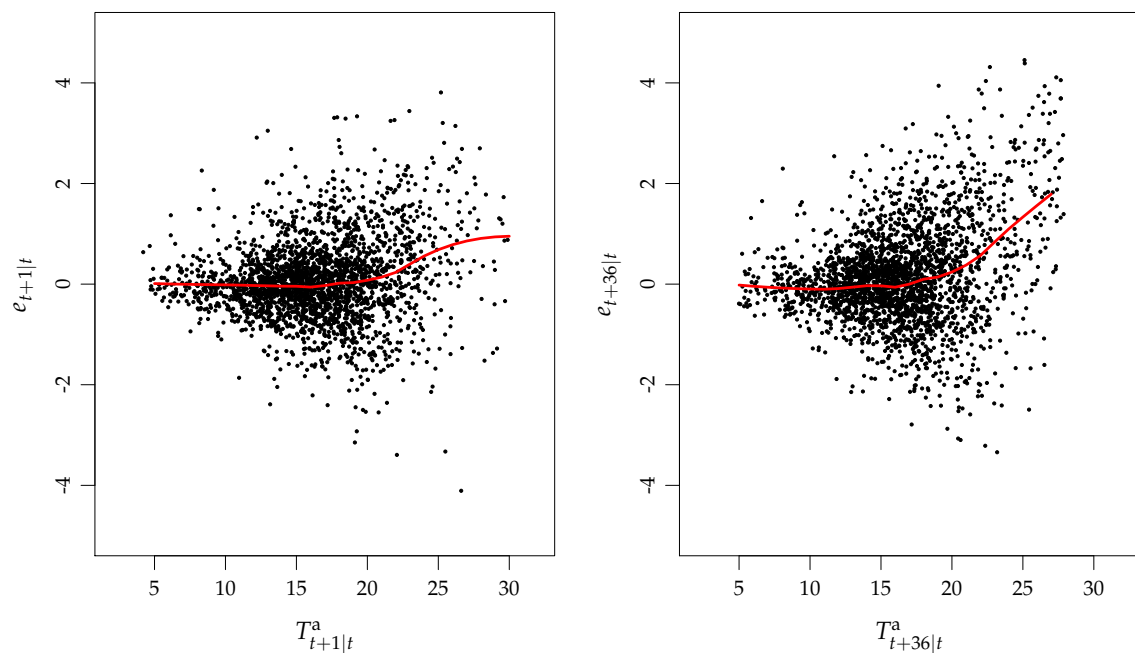
Figure 16: The $k = 1$ and $k = 36$ hours ahead residuals versus the forecasted ambient temperature used as input for the warm period. The red lines are smoothed local regression estimates.

# 7 Discussion

It is found that the forecasting method is very well suited to load forecasting for supermarket refrigeration and that further improvement can be achieved. In this section the method and the results are discussed, and ideas for further improvement are outlined.

A possible physical explanation to bias of the residuals (i.e. non-linearities in the ambient temperature input) above around 20 °C, could be that if the indoor air temperature in the supermarket normally is around 20 °C, but increases on warm sunny days (with ambient temperatures above 20 °C) then the heat transfer of the cabinets to the surroundings increases, which needs to be accounted for in the models. It could also be a decrease in the coefficient of performance (COP) of the heat pump, which converts the electrical power into heat.

Further work should focus on inclusion of the non-linear effects in the ambient temperature into the model. One approach is to use an off-line regression model as carried for wind power forecasting, where a power curve is fitted to account for the non-linear effects, see e.g. . Another approach would be to include the non-linear effects directly into the model by using piece-wise linear or spline function approach, or a kernel based method could applied.

Another improvement would be, for the shorter horizons, to apply an auto-regressive noise model, as mentioned previously.

The separation into closing and opening periods could also be carried out using a forecasting model which includes a diurnal curve. This could then be used to estimate the regime by a the sign of the diurnal curve. It is also noted that the coefficient of the applied low-pass filter could be optimized separately for both regimes in order to include different dynamical relations in closing and opening hours in the model.

Analysis of the effect of not using measurements of ambient temperature as input would also be relevant to investigate to evaluate the need for local measurements. Furthermore, in future experiments it will be relevant to include local measurements of also the solar radiation and wind, in order to determine if forecasting performance can be improved be using more local observations.

Further work could also focus on modeling of the uncertainties, since this will provide valuable information for the operation of energy systems based on fluctuating renewable energy production.

Future experiments and studies should include more than one supermarket to analyze the impact on the forecasting performance for different systems and conditions, and how to build a model which embraces the characteristic of many different refrigeration systems and conditions.

The forecasting method is found feasible to implement operationally and can be automatized to a high degree. Certainly, flawed data can cause problems, however schemes for identifying issues which needs manual handling can be implemented. Alarms could for example be triggered by unusual changes in coefficient estimates or unusually highly auto-correlated residuals. It is noted that the current implementation in R [1] is not compiled code and can be further optimized. However, a test shows that around 1000 forecast updates (including the recursive parameter estimation) of the 42 hours forecasts using the selected $Model_{\mathrm{A.G.W}}$ can be calculated in approximately 10 seconds on a 2.4 GHz single CPU computer. This is due to the computationally light recursive least squares scheme. If an update is needed every hour the time in-between updating can be used for data handling and off-line parameter optimization of the parameters listed in Section 3.4. The off-line optimization can be implemented with a recursive scheme and do not require updating very often, perhaps once a week. Based on this very coarse assessment it is found that operational implementation for a large number of systems can be carried out with feasible amounts of computational power.

# 8 Conclusion

A method for forecasting the load for supermarket refrigeration is presented. It is formed by adaptive linear time-series modeling techniques, using local observations and weather forecasts as input. Based on load measurements from a supermarket, a model is identified by using a forward selection approach. It is shown how the forecasting performance increases when extending the model by using the ambient temperature as input in two separate regimes: one for the closing hours and one for the opening hours. Furthermore it shown that the performance increase when a different diurnal curves are used, one for work days and another for weekends. For inclusion of dynamics of the system simple low-pass filter transfer functions are used. In the last step of the model selection it is shown that using the weather forecasts of global radiation and wind speed result in over-parameterization and decreased forecasting performance. The forecasting results are thereafter analyzed thoroughly to give insight into the error sources and it is found that non-linear effects in the ambient temperature are not fully described by the model. Finally, a discussion is given with ideas for further studies and advancements of the method.

# Acknowledgement

---

[1] www.r-project.org

# References

P. Bacher, H. Madsen, H. A. Nielsen, and B. Perers. Short-term heat load forecasting for single family houses. Submitted to Buildings and Energy, April 2012.

Peder Bacher, Henrik Madsen, and Henrik Aalborg Nielsen. Online short-term solar power forecasting. *Solar Energy*, 83(10):1772–1783, 2009. ISSN 0038092x.

James E. Braun and Nitin Chaturvedi. An inverse gray-box model for transient building load prediction. *HVAC&R Research*, 8(1):73–99, 2002. doi: 10.1080/10789669.2002.10391290.

DMI. Danish Meteorological Institute, DMI-HIRLAM-S05, 2011. URL http://www.dmi.dk/eng/index/research_and_development/dmi-hirlam-2009.htm.

Kristian Fredslund. Load profiles for supermarket refrigeration. Technical report, IPU Refrigeration and Energy Technology, 2013.

M.J. Jiménez, H. Madsen, and K.K. Andersen. Identification of the main thermal characteristics of building components using matlab. *Building and Environment*, 43 (2):170–180, 2008. ISSN 03601323, 1873684x. doi: 10.1016/j.buildenv.2006.10.030.

H. Madsen and J. Holst. Estimation of continuous-time models for the heat dynamics of a building. *Energy and Buildings*, 22(1):67–79, 1995. ISSN 03787788.

Henrik Madsen. *Time Series Analysis*. Chapman & Hall, 2007.

Henrik Aalborg Nielsen and Henrik Madsen. Modelling the heat consumption in district heating systems using a grey-box approach. *Energy & Buildings*, 38(1): 63–71, 2006. ISSN 03787788. doi: 10.1016/j.enbuild.2005.05.002.