Analysis of energy consumption in single family houses

May, 2010

Prepared by ENFOR A/S for
The Danish Electricity Saving Trust

ENFOR
Forecasting and Optimization for the Energy Sector

ENFOR A/S
Lyngsø Allé 3
DK-2970 Hørsholm
www.enfor.dk

Document ID: 03EKS0009A002-A

Contents

Summary

This report deals with estimation of the thermal characteristics of single family houses based on measurements of energy consumption and climate. The thermal characteristics includes the response of the building to changes in temperature (UA-value), solar radiation (gA-value), and wind (wA-value). The effect of the wind is characterised both in terms of the wind speed and the wind direction, implying that wA-values are estimated for different wind directions. Also, the dynamic response to climate is characterised and estimated.

It is concluded that if only the UA-, gA-, and wA-values are of interest it is sufficient to base the estimation on 24 hour averages, whereas in order to estimate the dynamic effects 4 hour averages are required. Also it is beneficial to use measurements of the total energy consumption (heat and electricity). However, in most cases it results in adequate estimates if only the heat consumption is used as the response variable. Obvious exceptions are cases where e.g. electrical floor heating is used for some periods.

The estimated thermal characteristics has been analysed with respect to background information regarding the households. The information is obtained via questionnaires and via the Danish Building Register (BBR). The significant effects are the ground area of the house, the year of construction, and the number of times per week a wood burning stove is used. Consequently, given information regarding the use of wood burning stoves which consumers share the estimates can be corrected so that adequate values are obtained even for houses where a wood burning stove are used.

The report includes a short outline of how the methods could be integrated in an interactive services such as "My Home" `www.elsparefonden.dk/minbolig`.

1 Introduction

This report is concerned with estimating thermal characteristics of single family houses based on measurements of energy consumption and climate. The thermal characteristics includes the response of the building to changes in temperature (UA-value), solar radiation (gA-value), and wind (wA-value). The effect of the wind could be characterised both in terms of the wind speed and the wind direction, implying that wA-values are estimated for different wind directions. Especially, the UA- and wA-values are directly related to the insulation and air sealing of the house. The gA-values are related to the ability of the house to passively use solar heating. A further characterisation of the building is the dynamic response to changes in climate variables. The dynamic response is characterised by time constants of the response to temperature and solar radiation. The report also considers methods of estimating these.

The outline of the report is as follows. Section 2 reflects on perspectives of application of the results found in this report for either an end user, via interactive services such as "My Home" `www.elsparefonden.dk/minbolig`, or a district heating company. The data used in the analysis is presented in Section 3 and Section 4 gives an overview the physical relations used as basis for the statistical modelling. The first part of the analysis in this report in Section 5 is based on daily averages of energy consumption. For this time resolution the heat transfer in and out of a house can with a reasonable approximation be assumed to balance. The second part of the analysis in Section 6 is concerned with analysis of energy consumptions with a sampling period of mainly 4 hours although a sampling period of 2 hours is also considered. For these lower sampling periods the model for heat transfer must also include the dynamics effects of heat transfer. The report ends with a discussion and conclusions in Section 7.

2 Perspectives

The perspectives of the work presented in this report points towards an automatic estimation of the main thermal characteristics of a building. Such calculations can be used to supply users with valuable information about their house via home-pages such as "My Home" `www.elsparefonden.dk/minbolig` or to help e.g. district heating companies determining users who might be worth-while contacting.

First, consider interactive services such as "My Home". Here measurements of electricity consumption is already available and clearly in order to estimate the thermal characteristics of a given building measurements of the heat consumption is required. The results presented in this report show that daily values of heat consumption are sufficient unless the dynamic characteristics are required. It is probably fair to claim that the non-dynamic thermal characteristics are the most important ones with relation to the overall energy consumption of the building. However, the dynamic characteristics could be valuable in order to set up the heating control system of the house and as such could be valuable to professionals installing or tuning the heating system of the house.

The information to the user could be extended with behavioural information. The *actual* heating season can be detected and if this is unreasonable long the user can be advised to turn of the heating system during e.g. summer periods. Also, it the user specify the overall indoor temperature and if data from the summer period is used the consumption used for hot tab water can be estimated.

Figure 1 below show a text-based sketch of the main elements in a user interface of a possible application based on the work presented in this report. As indicated such a user interaction should indicate clearly to the user what the estimated values mean in terms of the insulation of the house. Two parallel approaches are possible in this aspect:

Based on measurements from the heating season **2009/2010** your typical indoor temperature during the heating season has been estimated to **24** ^oC. If this is not correct you can change it here $\boxed{24}$ ^oC.

If your house has been left empty in longer periods with a partly reduced heat supply you have the possibility of specifying the periods in this $\boxed{\text{calendar}}$.

According to BBR the area of your house is **155** m^2 and from **1971**.

Based on BBR information it is assumed that **you do not use any supplementary heat supply**. If this is not correct you can specify the type and frequency of use here:

- Wood burning stove used $\boxed{0}$ times per week in cold periods.
- Solar heating $\boxed{\text{y/n}}$, approximate size of solar panel $\boxed{0} \times \boxed{0}$ meters.

Based on the indoor temperature **24** ^oC, the use of a wood burning stove **0** times per week, and **no** solar heating installed, the response of your house to climate is estimated as:

- The response to outdoor temperature is estimated to **200** $W/^oC$ **which given the size and age of your house is expectable**[a].
- On a windy day the above value is estimated to increase with **60** $W/^oC$ **when the wind blows from easterly directions. This response to wind is relatively high and indicates a problem related to the air sealing on the eastern side of the house.**
- On a sunny day during the heating season the house is estimated to receive **800** W as an average over 24 hours. **This value is quite expectable.**

[a] Many kind of different recommendations can be given here.

Figure 1: Main elements in a possible user interaction. ***Bold*** *entries indicates information specific for the user and* $\boxed{\textit{boxed}}$ *fields indicates information which the user has the ability to enter. Assuming measurements is available on a time scale of 4 hours the above could be supplemented with the dynamic characteristics of the response.*

(i) The values are related to a database of expert knowledge regarding what energy class the building belongs to.

(ii) The values are related to values estimated for other users and the system can automatically inform the user about the thermal performance of the particular house compared to other houses handled via “My Home”.

The information sharing implied by (ii) can be handled anonymously and is also required in order to e.g. adjust estimates for use of a wood burning stove. As the number of users increase this information sharing ability of the application will be increasingly valuable.

Besides the interactive use as outlined above the methods can also be used by e.g. district heating companies in order to screen for households with an unusual high consumption. This is of interest to district heating companies since these are obliged to implement energy savings. From 2010 also energy savings in the network can be formally included in the total energy savings of a given district heating company. In the report it is shown that the thermal characteristics of the building can often be well estimated based on measurements of the heat consumption alone. This is the case when the electricity consumption is not too large as it e.g. would be if electrical floor heating is used. Possibly, as a supplement to the BBR information, the district heating companies could collect information regarding e.g. wood burning stoves, electrical floor heating, and solar heating using questionnaires.

3 Data

The data used in the present report consists of heat and electricity consumption data for the period from ultimo September 2008 to primo December 2009 from 56 households connected to the district heating system in Sønderborg, Denmark. Also climate data and additional data describing individual households are used in the analysis. All data is described in more detail in the following.

3.1 Energy consumption data

The energy consumption data is described in detail in ENFOR/03EKS0007A001-A[2]. In this report four variables are used, namely `ID` giving identification number for the household, `ENERGI` giving accumulated district heating consumption [GJ], `EFFEKT` measuring heat load every 10 minutes [kW] and finally `P2` with accumulated electricity consumption [kWh]. Based on this data four measures of energy consumption is derived

- `heat`, `heat.EF`: heat consumption in sampling period [kWh],
- `total`, `total.EF`: sum of heat and electricity consumption in sampling period [kWh].

The heat consumption in `heat` is derived from the variable `ENERGI` measuring the accumulated district heating consumption given in GJ with a unit resolution of 0.01GJ. The heat consumption in `heat.EF` is derived from the variable `EFFEKT` measuring the heat load every 10 minutes with a unit resolution of 0.1kW. The total energy consumption in `total` and `total.EF` are found by adding the electricity consumption to `heat` and `heat.EF`, respectively.

Based on questionnaire data which is described in Section 3.3 40 households are selected to be used in the analysis and of these 26 have valid electricity measurements available. The

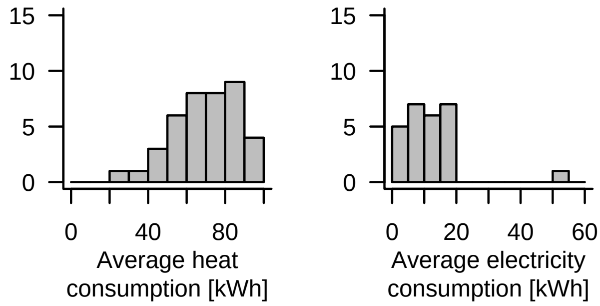

Figure 2: Distribution of daily average heat and electricity consumption in winter 2008/09.

distribution of the average of daily heat and electricity consumptions for these households are shown in Figure 2 for the winter period from 01-10-2008 to 15-04-2009. Appendix A lists a summary of the data for all 40 households.

1 day sampling period

For the analysis using a sampling period of 1 day it is chosen to work with data based on the accumulated heat consumption in the variable `ENERGI`. With a unit resolution of 0.01GJ this gives the daily consumption a unit resolution of 2.78kWh/day, which is found to be sufficient for the analysis of daily values. Alternatively, the daily consumption could be derived from the variable `EFFEKT`. With a unit resolution of 0.1kW this gives the daily consumption a maximal unit resolution of 0.00069kWh/day if all 144 observations of 10 minutes values are present. However, the `EFFEKT` observations are sometimes missing for parts of a day, which makes the estimate of daily heat consumption more unreliable in these cases. To avoid reducing the number of observations it is chosen to base the estimate of the daily heat consumption on the accumulated measurement.

The daily consumptions in `heat` is estimated based on the difference in the accumulated consumption from midnight to midnight. If the value at midnight is missing, an interpolated value is used if the accumulated value is observed at some point within 30 minutes from midnight. This again helps reducing the number of missing values. Any value of `ENERGI` equal to zero is assumed missing and daily consumptions larger than 1GJ (278kWh) are considered as outliers and removed.

The electricity consumption is treated in the same way as `ENERGI` to derive the electricity consumption and added to `heat` to give the total consumption in the variable `total`. For electricity daily consumptions larger than 100kWh are considered outliers.

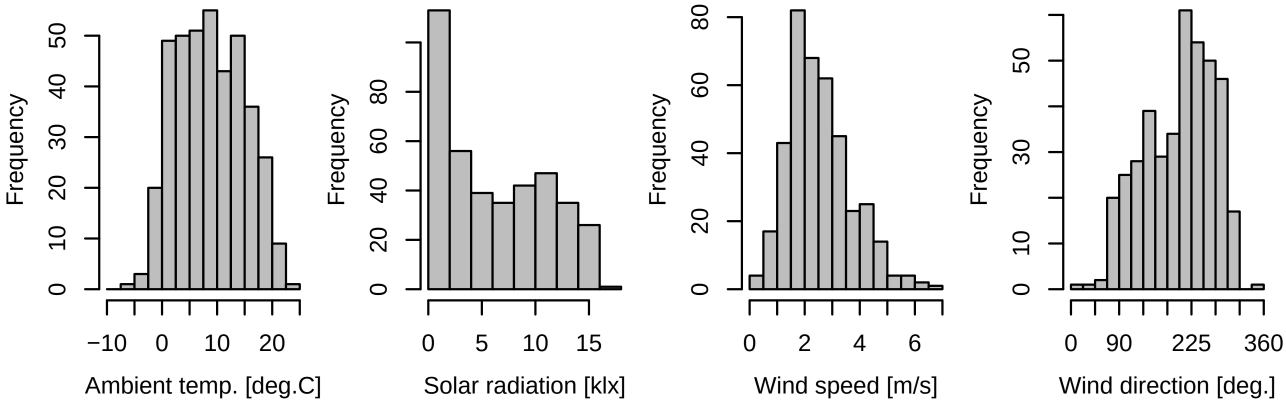

Figure 3: Daily averages of climate variables.

4 hour sampling period

For the 4 hour sampling period the heat consumption is based on the variable `EFFEKT` to give sufficient unit resolution. The sampling periods over a day starts from midnight giving the limits 00h-04h-08h-12h-16h-20h-24h for the 6 periods.

During 4 hours a maximum of 24 10-minute measurements can be available. The heat consumption is estimated by taking average of the available 10-minutes values and multiplying by the sampling period. If fewer than 24 measurements are available, this estimate is may become significantly inaccurate and thus only estimated consumptions based on at least 12 10-minutes measurements are used in the analysis. To remove remaining extreme values power consumptions larger than 50kWh, 30kWh and 18kWh were removed for households with meters 4569574, 5193768 and 5219101, respectively.

3.2 Climate data

Climate data are available from in the period from 2008-10-06 to 2009-11-18 with a 10 minute sampling interval. The available variables are ambient temperature in °C, solar radiation in lux, wind speed in m/s and wind direction in degrees.

All climate data are downsampled to averages for the necessary sampling period. The resulting distribution of the variables for a sampling period of 1 day are shown in Figure 3.

It is noted that the maximal average wind speed of about 7m/s seen in Figure 3 is relatively small. This may have been caused by a placement of the measurement equipment which did not give free wind. The wind speed measurement has been checked by comparing it with a daily averages of wind speeds predicted by DMI/VEJR2 for Sønderborg for the same period.

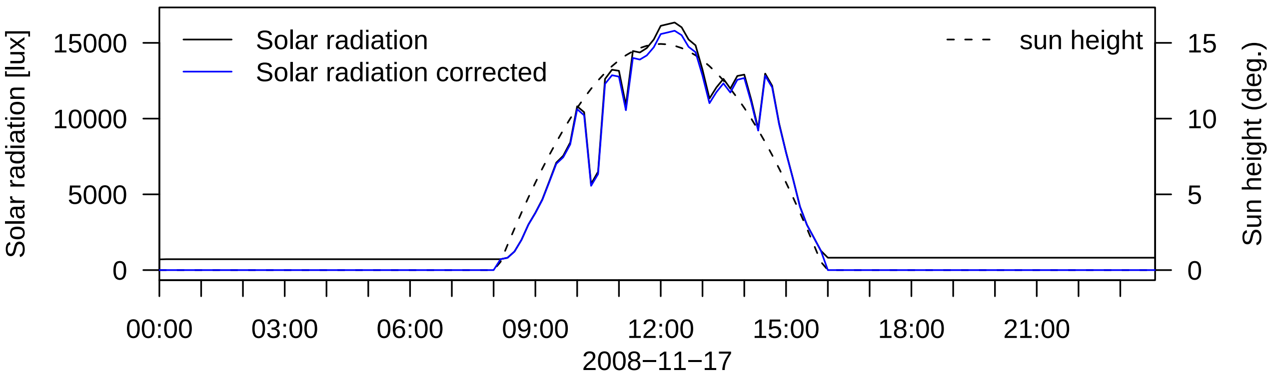

Figure 4: Measured and corrected values of solar radiation for November 17, 2008.

Although these predictions have a higher range (from approx. 1 to 10 m/s) they have a correlation of 0.92 with the wind speed from the climate data. Due to this relatively high correlation the wind speed from the original climate data will be used in the analysis.

It is also noted that the wind direction data seems to have too few days with an average wind from north. A northern average wind direction (from 315 deg. to 45 deg.) are only observed three times during a the period covering more than a year and this could again indicate a problem with free wind for the measurement equipment.

The measurement of solar radiation is assumed to be dominated by direct sunlight and thus to be proportional with the effect of the direct sunlight. The models considered here are based on the effect of solar radiation orthogonal to the building walls. This value can be approximated by projecting the raw 10 minute values of solar radiation to the direction orthogonal to building walls based on the sun height at each time point. Strictly this approach assumes that the measurements of solar radiation are proportional to the flux hitting a plane orthogonal to the direct solar radiation. The sun height depends only on the latitude of Sønderborg and the time of day and year and can be found using formulas given in [3, pp. 45-46]. The sun height is also used to truncate measurements to zero after sunset, since the original measurement for some reason indicates a constant level of about 700 lux during night time. Figure 4 shows an example of the corrected measurement together with the original measurement and the sun height. A daily average of the corrected sun measurement is used for this analysis.

3.3 Questionnaire and BBR data

The questionnaires have been completed by 47 consumers and 40 of these have been used in the analysis. The remaining 7 have been excluded for the following reasons: 3 are missing a meter number, 2 are not matched by a meter number in the power consumption data, and 2 have indicated that the household in empty for periods longer than 1 month.

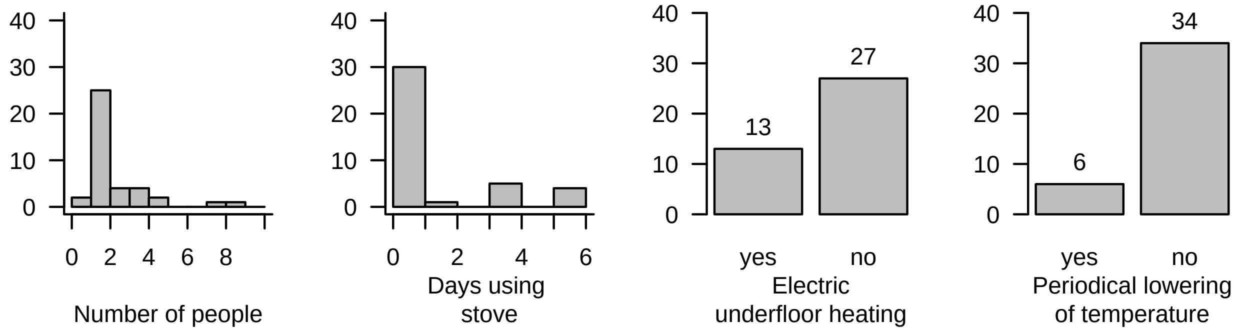

Figure 5: Histograms and bar plots with distribution of questionnaire data.

A summary of data collected from the questionnaires to characterise the households are shown in Figure 5 and are defined as:

- `bebr`: number of people,
- `stove`: number of days per week where fireplace or wood burning stove is used in winter season,
- `elflo`: electric underfloor heating used,
- `lowt`: periodically lowering of temperature (e.g. nightly).

Solar heating and supplementary heating has also been considered, but all 40 consumers have answered no and the variables were thus discarded.

The Danish building register "Bygnings og boligregisteret" (BBR) has been used to obtain further 2 variables describing the households. The data has been obtained from the web site *"Den Offentlige Informationsserver"* at `http://www.ois.dk` where the data is made freely available for public access. The distribution of the BBR data is shown in Figure 6 and are defined as

- `sqmt`: square meters of living area (Danish: *"Samlet boligareal i bygningen"*),
- `year`: year of construction.

4 Stationary relations for heat transfer

The stationary heat transfer for a house is for the main part assumed to be comprised by three ways of heat transfer, namely through walls, windows and by ventilation. Here heat transfer

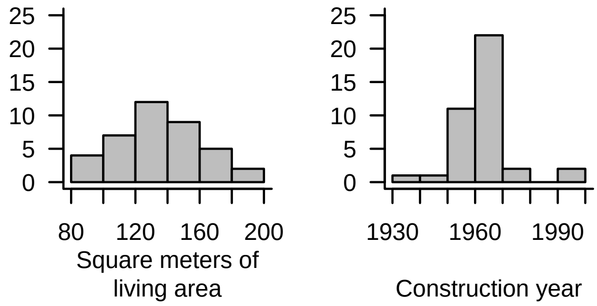

Figure 6: Histograms with distribution of BBR data.

through the roof is assumed to be included as part of the model for the walls. By considering stationary models for heat transfer trough walls and windows and via ventilation a model is derived:

- Responses on the temperature difference are collected into one term for which the coefficient is the UA-value.
- Responses on the solar radiation are are collected into one term for which the coefficient is the gA-value.
- Responses on the product of the temperature difference and the wind speed are collected into one term for which the coefficient is the wA-value.

The wA-value may be modelled as a function of the wind direction θ, in which case it is called $v(\theta)$.

The indoor temperature T_i is unknown but assumed to be constant which is also a reasonable assumption for a typical house where the indoor temperature is automatically controlled. By subsampling energy consumption and climate variables to daily averages they can be assumed to represent near stationary conditions, and coefficients in the model can thus be estimated based on a sufficiently large sample of average daily energy consumptions together with corresponding averages of the climate variables.

The model can only be used during the time period where the house is heated to maintain a constant indoor temperature, such that the heat transfer $\dot{Q}$ from the house can be measured based on the amount of energy supplied to the household. During summer time this is not the case, and part of the study here will thus also be concerned with automatically identifying the heating season for individual houses.

4.1 Statistical methods

The estimation of UA and gA values and wind dependence is based on the outlined above. The unknown parameters in the model are UA, gA, $v(\theta)$ and T_i. The function $v(\theta)$ is modelled either as a constant $v(\theta) = c_w$ or as piecewise constant for the major wind directions. As noted in Section 3.2 there are only the three days with an average wind direction from the northern quarter, and it is chosen to keep only three major wind segments, namely east (E) 0-135 deg., south (S) 135-225 deg. and west (W) 225-360 deg. The piecewise constant approximation to $v(\theta)$ is given as

$$v(\theta) = \sum_{j=E,S,W} I(\theta \in j) c_{wj} \tag{1}$$

where I is an indicator function equal to 1 when the argument is true and otherwise 0. The three coefficients c_{wj} gives wind dependence in the model and is interpreted as 'wA' values such that $\mathrm{wA}_j = c_{wj}$.

5 Analysis of daily values of power consumption

The first part of the analysis in this report is carried out using daily power consumption values, i.e. using a sampling period of 1 day. The variable `total` including both district heating and electricity is used as response in order to have the best data for all energy entering the house, including energy used for electric underfloor heating. Households using a stove are not excluded, but instead it will be analysed how estimates of UA values etc. are affected when this additional energy is not accounted for. The analysis of daily power consumption is thus based on all the 26 households where electricity data is available. A later part of the analysis will consider the effect of omitting the electricity consumption.

5.1 Time varying estimates

Initial investigation of the energy consumption data is done by estimating the time variations of the coefficients in a linearized and simplified version of the model described above To reduce the number of parameters to be estimated the last term $c_w w T_a$ is not included giving the model

$$Q_t = b_0 - \mathrm{UA} \cdot T_{a,t} - \mathrm{gA} \cdot R_{0,t} + b_1 w_t + e_t \tag{2}$$

where b_0 and b_1 are constants, $T_{a,t}$ is the ambient air temperature, $R_{0,t}$ is the measure of solar radiation, and w_t is the wind speed. The coefficient b_1 cannot be interpreted in relation to the physical model, but it still gives an indication of wind speed dependence in the energy consumption.

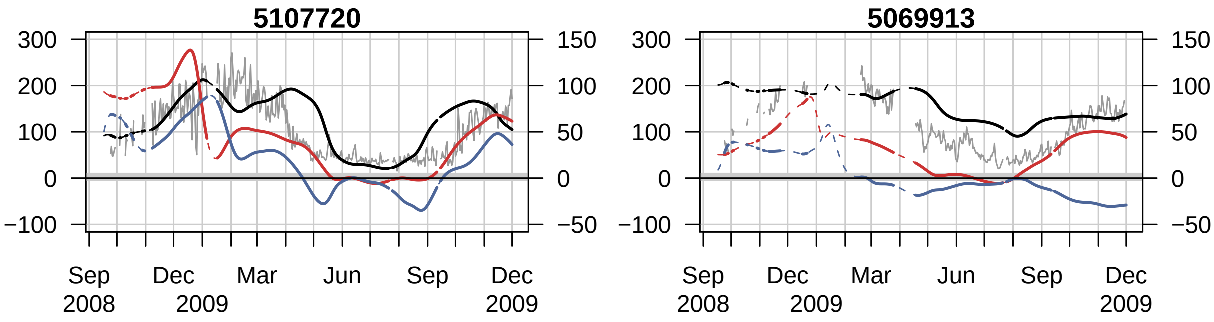

Figure 7: Time varying estimates of coefficients in Eq. (2)*. Black is UA [W/°C], red is gA [W/klux] and blue is b_1 [W/m/s] all measured on the left side axis. The underlying gray curve is daily total energy consumption in kWh measured on the right side axis.*

The time variations are estimated using locally weighted estimation of the linear model. The method is described in [4] and gives local estimates in time of the model coefficients by only considering observations within a limited time window. This makes it possible to see if they are constant over time, e.g. to look for variations during the heating season and how they change during the summer period. In Figure 7 shows these time varying estimates for two households and estimates for all 26 households can be found in Appendix B.

For the most of the households the UA, gA and b_1 values are relatively stable during the winter period which is also seen for the two households in Figure 7. However, in the household with meter no. 5107720 the UA value drops close to zero during the summer period, but this is not the case for meter no. 5069913. This appears to be caused by the fact that heat is not turned off during summer such that the house still reacts to low outdoor temperatures. This effect is seen in an increased energy consumption in the first half of June 2009, which was an unusually cold period with daily average temperatures below $10°C$. For meter no. 5069913 there is a clear 'bump' in the total energy consumption during this period which is not found for the other household no. 5107720. Looking at the plots in Appendix B this effect generally holds, such that houses where UA is high during summer also have an increase in energy consumption during July.

5.2 Parametric modelling of time-variations

Based on the estimates of time variations of the coefficients in the model in Eq. (2) it seems reasonable to assume that the coefficients can be modelled with a constant level for each of the two winter and one summer periods, giving three levels in total. Estimating when the changes in level occur will indicate the exact extent of the heating season for each individual house and

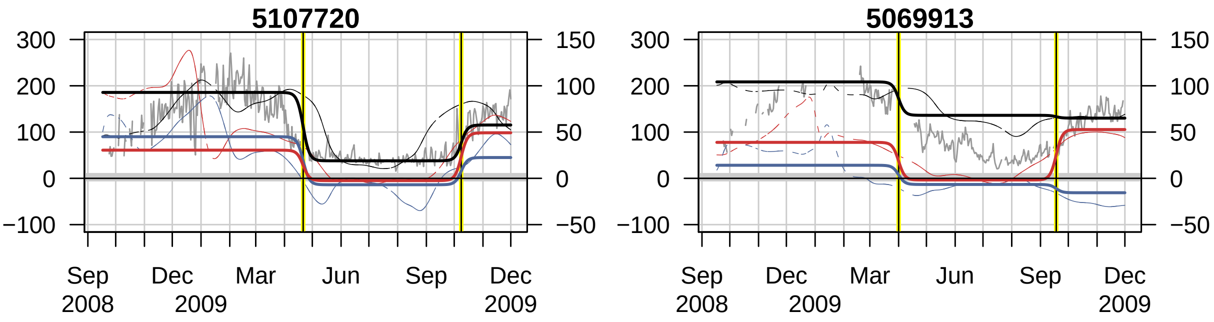

Figure 8: Thick lines are estimates of parametric models for time varying coefficients in Eq. (2)*. Black is UA [W/°C], red is gA [W/klux] and blue is b_1 [W/m/s] all measured on the left side axis. Thin lines are estimates based on local regression from Figure 7. Vertical black-yellow lines are estimated time point of change. The underlying gray curve is daily total energy consumption in kWh measured on the right side axis.*

this information can then be used to select the longest possible period of the actual heating season for further analysis.

The time variations of each parameter $p_k(t)$, $k \in \{b_0, \mathrm{UA}, \mathrm{gA}, b_1\}$ in Eq. (2) is modelled by a level for the summer plus an offset for the two winter periods and can thus be defined as

$$p_k(t) = p_{k0} + \sum_{i=1}^{2} p_{ki} f(t, t_i, n_i) \tag{3}$$

with $f(t, t_i, n_i) = 1/[1 + \exp((t - t_i)n_i)]$. The parameters t_1 and t_2 defines the two time points when the levels change, and n_1 and n_2 defines the rate of change. The model is estimated by means of partial linear estimation techniques and results are shown for the two selected households in Figure 8 and the remaining households are found in Appendix C.

The method works well for households where heating is not used during the summer period resulting in clear estimates of changes in levels for summer and winter as can be seen in the left plot in Figure 8. In particular the estimate of the time of change is more accurate than judging the local estimates of the time variations. The latter are based on a window around the time point of interest which gives rise to some amount of bias. This effect can be seen in the left plot in the estimate of the UA value (black lines) where the thin line based on local estimation decreases later in the spring and increases earlier in the fall compared to the parametric estimate of the time change in the thick black line.

For the households which also use heating during the summer it is more difficult to estimate the change in level between seasons. In order to make convergence more robust in these situations it is necessary to fix the rate of change prior to estimation such that $n_1 = n_2 = 0.3$. In cases where there is only a small change in estimated levels for summer and winter as in the right plot in Figure 8 the estimate of the time point of change is more uncertain, which can also be seen from an increased standard error of the estimate. However, this may just be an indication that the heating system is not shut down during the summer.

5.3 Estimation based on single heating season

The last part of the analysis of daily values of energy consumption is done based on a single heating season. The first part of the analysis has shown that the parameters can be considered constant throughout the heating season and this allows the use of the model described in Section 4. In previous work for estimation of UA values alone based on daily averages of energy consumption in ENFOR/03EKS0001A003-A[1] it has been found that there is significant dependence on the ambient temperature one day back for the heat consumption. This dynamic effect is also included here. The dependence on wind speed is estimated with a coefficient for each of the major wind direction east, south and west as given in Eq. (1).

In order to be able to get a good estimate of the effect of the solar radiation the heating season must extend into the spring where there is a significant contribution from the sun. Based on the results from the previous section found in Appendix C it is chosen to define the heating season from 1st of October to 15th of April common for all households.

Results from estimation are shown in Table 1. Residual plots for all models are found in Appendix D together with estimated auto-correlations. The plots indicate a quite good fit of the model in general and the auto-correlations show only a small amount of correlation for some of the households. For some households (e.g. 5107720, 5183232 and 5223036) there seem to be a tendency to a small peak in auto correlations for 7 and 14 days back. This indicates a possible weekly changing pattern in energy consumption which has not been included in the model.

To aid the interpretation of the estimates the estimates of gA are multiplied by the 99% fractile of observed daily average of solar radiation and denoted $\mathrm{gA}^{\max}$. This parameter determines the average effect in W that is absorbed during a day with maximal average solar radiation. Similarly the direction dependent estimates of the wind c_{wj} are multiplied by the 99% quantile of observed daily average wind speed and denoted $\mathrm{wA}_E^{\max}$, $\mathrm{wA}_S^{\max}$, and $\mathrm{wA}_W^{\max}$. These parameters determine the absolute change in UA value in W/°C due to ventilation for a day with a maximal average wind from the particular direction. As example, household no. 4836681 has a UA value of 155.3W/°C and during a very windy day with wind coming from east the UA value is increased by 39.5W/°C.

In Table 1 the UA value and indoor temperature T_i are shown together with their estimated standard error. These parameters are significant in all models and also individually testing c_1 was found significant in all models. The $\mathrm{gA}^{\max}$ and $v_*^{\max}$ parameters are shown together with a p-value testing the significance of the parameters using a likelihood ratio test, since these parameters are not always statistically significant. For households that have autocorrelated residuals these p-values are too small, since the test assumes uncorrelated residuals.

	UA W/°C	σ_{UA}	$\mathrm{gA}^{\max}$ W	$\mathrm{wA}_E^{\max}$ W/°C	$\mathrm{wA}_S^{\max}$ W/°C	$\mathrm{wA}_W^{\max}$ W/°C	T_i °C	σ_{T_i}	p_{gA}	p_v
4218598	211.8	10.4	597.0	11.0	3.3	8.9	23.6	1.1	0.00	0.51
4218600	98.7	10.8	-96.2	23.6	10.1	13.0	22.3	2.3	0.62	0.08
4381449	228.2	12.6	1012.3	29.8	42.8	39.7	19.4	1.0	0.00	0.00
4711160	155.4	6.3	518.8	14.5	4.4	9.1	22.5	0.9	0.00	0.03
4711176	178.5	7.3	800.0	1.9	-7.6	8.5	26.4	1.0	0.00	0.00
4836681	155.3	8.1	591.0	39.5	28.0	21.4	23.5	1.1	0.00	0.00
4836722	236.0	17.7	1578.3	4.3	3.3	18.9	23.5	1.6	0.00	0.06
4986050	159.6	10.7	715.7	10.2	7.5	7.2	20.8	1.4	0.00	0.68
5069878	144.8	10.4	87.6	3.7	1.6	17.3	21.8	1.5	0.57	0.03
5069913	207.8	9.0	962.5	3.7	8.6	10.6	22.6	0.9	0.00	0.01
5107720	189.4	15.4	657.7	41.4	29.4	16.5	21.0	1.6	0.00	0.08
5127784	264.7	16.6	1364.5	18.4	-10.0	-20.0	27.0	1.7	0.00	0.00
5159799	204.8	5.5	614.2	-1.9	-2.9	3.9	26.0	0.7	0.00	0.01
5164474	173.4	14.3	68.4	8.2	8.2	-4.8	23.4	1.8	0.59	0.33
5164485	196.2	6.6	931.3	14.6	23.8	30.6	22.6	0.8	0.00	0.00
5164523	148.3	8.5	758.1	-6.9	1.1	7.0	26.0	1.4	0.00	0.03
5168264	169.6	7.7	554.1	25.8	8.4	2.1	21.7	0.9	0.00	0.00
5183206	177.7	14.3	429.0	-4.3	-26.2	5.6	24.2	1.7	0.00	0.00
5183228	208.9	7.8	724.7	23.1	19.4	31.9	21.6	0.7	0.00	0.00
5183232	128.8	14.6	608.7	18.4	2.5	8.4	25.0	2.7	0.00	0.42
5191179	63.3	5.4	186.9	0.3	-1.1	0.4	50.0	4.2	0.00	0.56
5194940	221.5	13.2	246.3	8.0	2.1	30.5	17.5	1.0	0.35	0.04
5194965	132.3	9.6	407.5	-7.4	-2.4	7.3	26.5	1.9	0.00	0.02
5197381	182.3	13.9	1038.8	31.6	19.7	23.5	24.4	1.8	0.00	0.00
5223030	206.2	17.8	841.3	6.3	-42.1	-8.7	27.3	2.3	0.00	0.00
5223036	171.4	15.2	522.3	2.8	-6.5	12.7	22.7	1.9	0.00	0.06

Table 1: Estimates based on total energy consumption.

Histograms of from estimated values from Table 1 are found in Figure 9 (with the outlier removed). In Figure 10 the values are normalised by the size of the house using the variable `sqmt`. This narrows the distribution, especially for the UA values. The estimates of the UA and $\mathrm{gA}^{\max}$ values are all positive as they are expected to be except for one case of $\mathrm{gA}^{\max}$ for household no. 4218600. The estimates of the $\mathrm{wA}_*^{\max}$ values are mostly positive, although there are some negative estimates indicating a reduced UA values these wind directions. However, overall the estimates of the $\mathrm{wA}_*^{\max}$ values gives a picture of the wind dependence for each house, and it is seen that some are clearly more wind sensitive than others.

The estimated indoor temperature to some extent serves to validate the model. T_i is, except for one outlier no. 5191179, estimated in the range from $17.5°C$ to $27.3°C$ which is within range of realistic values.

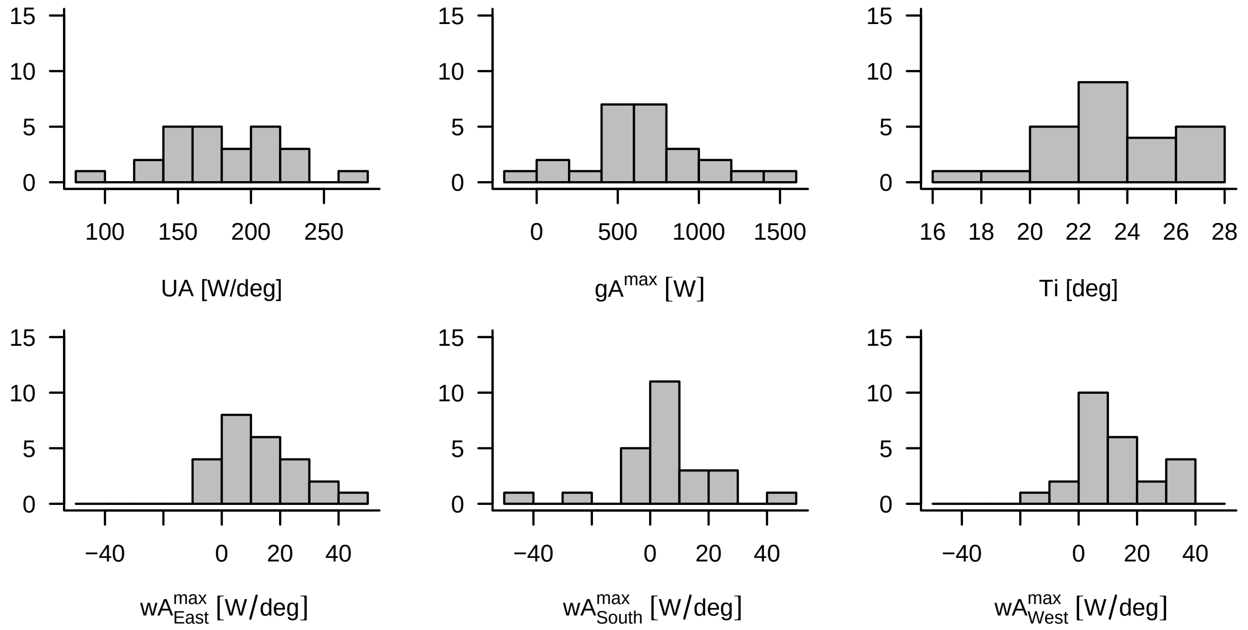

Figure 9: Histograms of estimates.

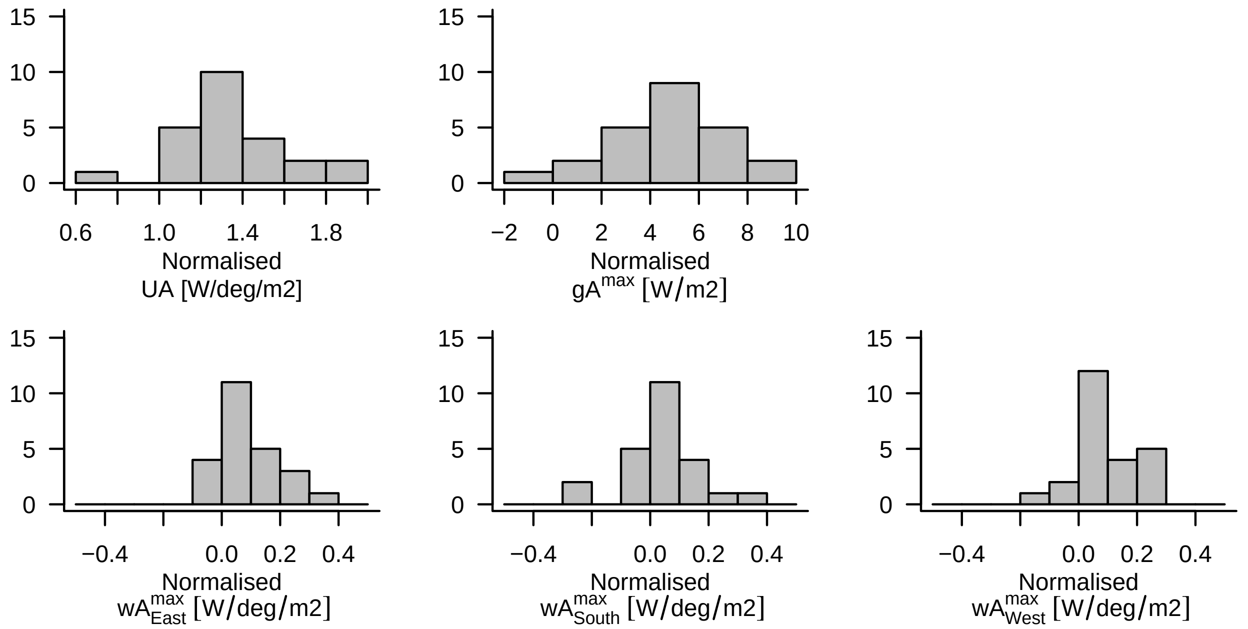

Figure 10: Histograms of estimates normalised by sq. meters.

5.3.1 Analysis of estimated coefficients

The estimates of UA and gA values for each household are analysed based on all explanatory variables described in Section 3.3 in order to investigate systematic differences between the households. The analysis is done using a standard analysis of variance where all variables are initially included in the model and following removed one by one if they are not statistically significant at a 5% level. The outlier no. 5191179 is not included in the analysis.

For the UA values the resulting model includes the variables `sqmt`, `year` and `stove` and is found to explain 48% of the variation in UA values. The estimated coefficients are found in Table 2. It is seen that the UA value increases 0.94 per square-meter and overall the UA values decreases by 2.6 per year the later the house is constructed. Of particular interest is the effect of the use of a wood burning stove, which is estimated to reduce the estimate of the UA value by 8.2 per day the stove is used per week. This indicates that the UA values are underestimated if there is a contribution to heat consumption from a stove in the household, since this is not included in the measured energy consumption. However, using the model it is possible to correct for this underestimation, and thus still provide a reasonable estimate of the UA values if information about frequency of the use of a stove can be obtained.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	56.8012	33.8290	1.68	0.1087
sqmt	0.9456	0.2426	3.90	0.0009
year	-2.5792	1.1342	-2.27	0.0341
stove	-8.2303	3.6987	-2.23	0.0377

Table 2: Analysis of UA values based on explanatory variables.

For the gA values the resulting model includes the variables `sqmt` and `stove` and is found to explain 62% of the variation in gA values. The estimated coefficients are found in Table 3. Again it is seen that the size of the house is significant and also that the use of a stove leads the underestimates of the gA value. There is no significant effect on the gA values related to construction year as it was found for the UA values. This is reasonable since the UA value is reduced by improved methods for insulation during the years, but the effect of solar radiation will not be affected by this.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-317.1228	317.8738	-1.00	0.3298
sqmt	7.9011	2.2759	3.47	0.0023
stove	-85.1062	26.8447	-3.17	0.0046

Table 3: Analysis of gA values based on explanatory variables.

5.3.2 Estimation using only district heating consumption

In this section the effect of using only heat consumption for estimation of UA and gA values is investigated. The model is thus based on the variable `heat` instead of the variable `total`. The results using `heat` as response are shown in Table 4 together with the change from using `total` as response.

	UA_{heat} W/°C	ΔUA W/°C	gA^{max}_{heat} W	ΔgA^{max} W	$T_{i,heat}$ °C	ΔT_i °C	`elflo`	`elfra`
4218598	217.75	-5.92	481.03	115.97	20.95	2.61	no	0.11
4218600	88.52	10.20	-258.27	162.06	15.42	6.88	yes	0.43
4381449	220.39	7.77	799.51	212.80	17.66	1.79	no	0.16
4711160	153.88	1.49	472.95	45.87	20.62	1.93	no	0.12
4711176	173.33	5.19	611.66	188.31	23.75	2.67	no	0.12
4836681	156.47	-1.15	578.11	12.91	22.97	0.54	no	0.02
4836722	230.02	6.00	1372.34	205.97	22.21	1.27	yes	0.09
4986050	154.53	5.09	664.00	51.72	19.22	1.55	no	0.13
5069878	116.03	28.72	22.68	64.88	23.56	-1.73	yes	0.12
5069913	197.87	9.93	818.97	143.54	19.72	2.91	yes	0.21
5107720	185.96	3.49	685.37	-27.65	20.07	0.97	no	0.08
5127784	243.45	21.22	422.92	941.62	18.83	8.20	yes	0.37
5159799	204.34	0.45	595.16	19.01	25.00	1.02	no	0.05
5164474	170.20	3.23	12.64	55.78	19.84	3.55	no	0.20
5164485	195.36	0.87	769.34	161.95	18.89	3.67	yes	0.20
5164523	146.59	1.70	640.39	117.75	23.43	2.57	yes	0.13
5168264	169.67	-0.03	553.99	0.13	21.70	0.00	no	0.00
5183206	142.91	34.83	566.93	-137.92	23.63	0.52	yes	0.23
5183228	209.28	-0.37	722.97	1.76	20.78	0.82	no	0.05
5183232	126.09	2.69	569.19	39.51	22.77	2.27	no	0.14
5194940	206.37	15.10	44.51	201.77	14.85	2.65	yes	0.24
5194965	131.74	0.58	356.68	50.84	24.85	1.63	no	0.07
5197381	185.28	-2.93	870.44	168.37	21.18	3.22	no	0.15
5223030	219.53	-13.34	902.27	-60.95	23.02	4.31	yes	0.13
5223036	169.96	1.41	381.04	141.23	18.84	3.83	no	0.20

Table 4: Estimates based on only heat consumption Columns with Δ-values are estimates using total consumption minus estimate using heat consumption.

In Figure 11 the changes in the estimate of UA values and indoor temperature T_i are plotted together with the fraction of the total energy consumption which comes through electricity (denoted `elfra`) and it is also indicated if the household uses electric underfloor heating or not. These two explanatory variables are also included in Table 4 for reference.

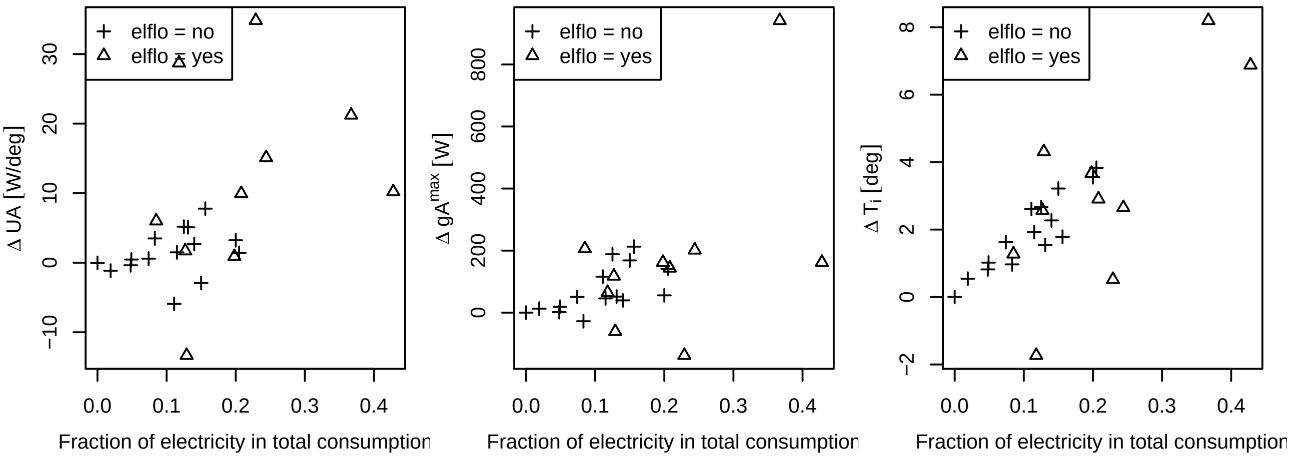

Figure 11: Illustration of increase in estimates of UA, gA^{max} and T_i using total instead of heat consumption plotted against the fraction of electricity consumption in the total consumption. Variable `elflo` *indicates use of electric underfloor heating.*

From Figure 11 it is seen that for households using underfloor heating (plotted as triangles) the UA value is for all except one estimated higher when electricity is included in the energy consumption. This is as expected, since when a house uses electricity for heating part of the response to outdoor temperature lies in the electricity consumption and the UA value is thus underestimated if electricity is not included in the energy consumption. Similarly for these households, the indoor temperature and gA^{max} values are also generally estimated higher when electricity is included in the consumption.

For households not using underfloor heating (plotted as crosses) the estimate of the UA value does not show a clear dependence of the electricity fraction when looking at Figure 11. For the gA^{max} values there is a tendency to underestimation when electricity consumption is not included. For households not using underfloor heating the most significant dependence is for the change in the estimate of the indoor temperature which appears directly proportional to the electricity fractions. This seems reasonable, since the electricity consumption is not intended for heating of the house and thus does not show up as an increased response to outdoor temperature. Instead, electricity consumed e.g. during cooking will temporarily raise the indoor temperature and thus increase the estimated average temperature. Also, if a large part of the electricity is used for lighting, this will raise the temperature in the upper part of the air in the house again raising the estimate of the indoor temperature.

From the above observations it is concluded that the UA value can be reasonably well estimated for a house without electric underfloor heating by using only the district heating consumption as response in the model. The estimate of the indoor temperature will be to low, but since this is not a parameter of interest this is not of concern. Based on this the analysis using only `heat` as response is redone for the households without electric underfloor heating. There are 27 households not using electric underfloor heating (see Figure 5) and all estimates are shown in Appendix E. The estimates of the indoor temperature are found in the range from 17.7°C to 25.7°C (the previous outlier no. 5191179 uses electric underfloor heating and is thus already excluded).

The estimated UA values are analyses with respect to the explanatory variables `sqmt`, `year` and `stove`. The result is shown in Table 5 and the model explains 57% of the variation in the UA values. The estimates for `sqmt` and `year` are similar to the ones found in Table 2 when estimating UA values using `total` as response. The estimate for `stove` is a little lower than before indicating an underestimation of the UA value of 3.1W/°C per day the stove is used per week. This estimate is a little more unreliable than before since there are only 5 households using a stove and not underfloor heating whereas there where 7 households using a stove in the analysis with `total` as response. This naturally emphasises that if the estimate of the effect of using a stove should be used to correct the estimated UA value it is necessary have a as large as possible data set available.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18.4575	31.3958	0.59	0.5626
sqmt	1.1205	0.2281	4.91	0.0001
year	-1.6489	0.5261	-3.13	0.0048
stove	-3.0834	2.8930	-1.07	0.2981

Table 5: Analysis of UA values estimated without electricity consumption.

The estimated gA^{max} values are also analysed with respect to the explanatory variables in Table 6 for comparison with Table 3. The coefficients in the two tables are more different than for the UA values, but the effect of `sqmt` is still positive and the effect of `stove` is still negative as would be expected.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	140.9287	246.8972	0.57	0.5737
sqmt	2.9344	1.7907	1.64	0.1149
stove	-47.5483	22.4965	-2.11	0.0456

Table 6: Analysis of gA^{max} values estimated without electricity consumption.

6 Analysis of energy consumption with shorter sampling time

The second part of this report deals with analysis of energy consumption on a shorter time scale than 24 hours in order to be able to model the dynamic effects relating to ambient temperature and solar radiation. The modelling in this section will only be concerned with the heating season 2008/09 which is defined as 1 October 2008 to 15 April 2009 as discussed previously. The analysis will mainly be based on a sampling time of 4 hours. This sampling time is chosen so that it is short enough to capture the dynamics caused by diurnal changes in ambient temperature and solar radiation, but still not so short that it is expected to be strongly influenced by the irregular behaviour of people living in the household. A 2h hour sampling period is also considered for a five selected households households as part of the analysis found in Appendix H.

It is chosen to base the analysis on households which are not using a wood burning stove, electric underfloor heating, or periodically lowering of temperature, and this gives a total of 17 households. This is done in order to focus on data with the best possible control of energy consumption in the household which is increasingly important for reliable estimation on shorter sampling times. The household 5197381 is excluded since there is only data for the month of December which is found to be to sparse to be considered here. Of the remaining 16 there are 8 which also have measurements of electricity consumption available.

The general structure of the applied dynamic model is based on the model described in Section 4. Additionally an intercept is included in the model to account for the energy consumed not resulting in heating of the house and which is thus not related to the climate variables. The intercept term will thereby cover e.g. warm water leaving the house when used for bathing etc. With a 4 hour sampling period the day is divided in the intervals 00-04-08-...-24. (see Section 3.1). The energy consumption covered by the intercept is assumed to be zero during night time 20-04 and vary during the day for the 4 periods 04-08-12-16-20 and is modelled with a level for each of these. The wind dependence $v(\theta)$ is again modelled a piecewise constant for the three major wind directions as given is Eq. (1). In order to account for dynamic effects the temperature difference and the solar radiation are low pass filtered using rational transfer functions $H_{T_a}(q)$ and $H_{R_0}(q)$, respectively. The UA and gA values estimated from the model are the stationary response to temperature and solar radiation.

Three different models based on the structure have been tried by varying the structure of the transfer functions $H_{T_a}(q)$ and $H_{R_0}(q)$. For model no. 1 the impulse response function for the temperature is a single exponential decaying function giving one time constant τ_{T_a}. The impulse response for solar radiation is a delta function making the solar effect instantaneous. Model no. 2 extends the first by also modelling the implulse response for the solar radiation as a single exponential decaying function with a time constant τ_{R_0}. Model no. 3 further extends the second model with a 2nd order polynomium in the denominator for $H_{T_a}(q)$. The roots in this polynomium are constrained to be real such that the impulse response function for the temperature is a double exponential decaying function with two time constants $\tau_{T_a,1}$ and $\tau_{T_a,2}$ if the corresponding poles are estimated positive. The time constant are given as the time to $1 - e^{-1} = 63.2\%$ of the maximal step response.

The models 1, 2 and 3 are estimated using a two step procedure. In the first step the indoor temperature T_i is assumed constant and estimated as a single parameter T_i in the models. In

	Response	No. of fits	No. not converged T_i const.	$T_{i,t}$ var.
Model 1	`total`	27	0	0
Model 2	`total`	27	0	1
Model 3	`total`	27	9	6
Model 1	`heat`	53	1	0
Model 2	`heat`	53	1	1
Model 3	`heat`	53	18	6

Table 7: Count of negative convergence of models 1 – 3 using 4h data.

a second step the indoor temperature estimated for each 4h period directly from the residuals from the fit in the first step. This estimated temperature is then smoothed using local constant model with a bandwidth of 4 days (i.e. the kernel ranges over 8 days). The smoothed estimate of the indoor temperature $T_{i,t}$ is then used as input in the three models, which results in updated estimates of all the remaining parameters in the models. The details regarding this method are found in Appendix F. Based on the results presented in that appendix it is chosen work with the model assuming a constant indoor temperature, and all results presented in this sections will be based on this.

The models are estimated using partial linear estimation techniques where T_i and the parameters in the transfer functions are nonlinear and the remaining are liniear. This gives 2, 3, and 5 nonlinear parameters for models 1 to 3, respectively (and one less in the second run where $T_{i,t}$ is given). This will in some cases result in ill-conditioned estimation problems which may not converge. To initially investigate the stability of the three models they are estimated on all availble households. There are in total 53 with heat consumption data available in the period 1 October 2008 to 15 April 2009 and of these 27 also have electricity data. The estimation is done with both `heat` and `total` as response and the results are shown in Table 7.

It is seen that Models 1 and 2 generally converge every time with only 4 failed convergences whereas Model 3 fails to converge 39 times. The cases of non-convergence for Model 1 and 2 are found acceptable since the data at rather strange in the cases of non-convergence. For Model 3 the numurous non-convergence issues does not seem to be related to the number of observations missing, but rather to strange temporal behaivour of the observations. Some of these could be related to holidays (consumption constant at a slightly lower level). In Appendix H a detailed analysis regarding the convergence problems for Model 3 is studied in detail where the model is also applied to 2h data. In the remaining analysis presented here Model 3 will be omitted.

Models 1 and 2 are estimated for all 16 households using both `heat` and `total` as response. Initially the estimates are analysed with respect to the fraction of missing data in the period considered. The fraction of missing data ranges form 20.1% to 59.5% with 3 out of 4 below 45%. The additional missing data for those with the highest percentage are mainly found in early 2009 where it was quite cold. For T_i and wA for all wind directions there is no clear dependence on the fraction missing. For the remaining parameters there is a weak tendency to higher estimates when a low fraction is missing, which is probably related to the noted pattern of missing data.

A graphical comparison of the estimates of model parameters for Model 1 (M1) and Model 2 (M2) using `heat` and `total` as response is performed. Figure 12 shows a comparison of UA and gA values and the comparison for the remaining model parameters (indoor temperature, time constants and wA-values) are found in Appendix G, Figure 23 and Figure 24. Both the UA and gA values are faily independent of the response `heat` or `total` and T_i is estimated higher using `total`. This coresponds well to the findings for the analysis on 24h data. For the UA vales there is a slight tendency to lower values for M2. For the gA values there are clearly higher estimates for M2 compared to M1 and this indicates that the dynamic response included in M2 for the solar radiation is important. The wA values are fairly constant between models but with some dependence on response type without any clear tendencies.

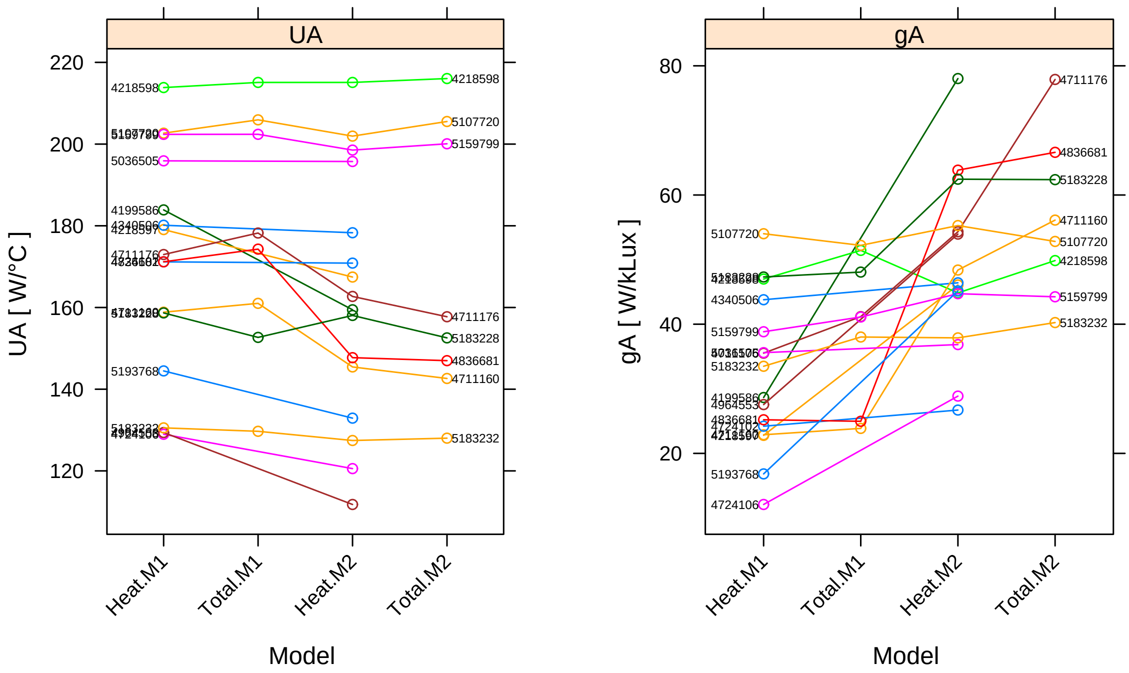

Figure 12: Estimates of UA and gA values using dynamical models on 4h data.

The comparison of estimates betweens Models 1 and 2 are mainly differentiatied by a larger estimated effect of solar radiation using Model 2. This indicates that the dynamic response for the sun is important to capture the total effect and Model 2 is therefore prefered over Model 1. The relation between estimated UA and gA values for the 4h and 24h models are illustrated graphically in Figure 13 and wA values are compared in Appendix G, Figure 24 (right). The comparison shows good agreement in UA values between 24h and 4h models, although with a tendency for the 24h values to be higher than the 4h values. The gA values seems to be estimated around 50% higher in the 4h models but otherwise the relative relation is fairly constant. These higher estimates seems reasonable, since the gA value is taken as the stationary response which is not fully reached for the sun since it is only present for around 8-12 hours during the winter season.

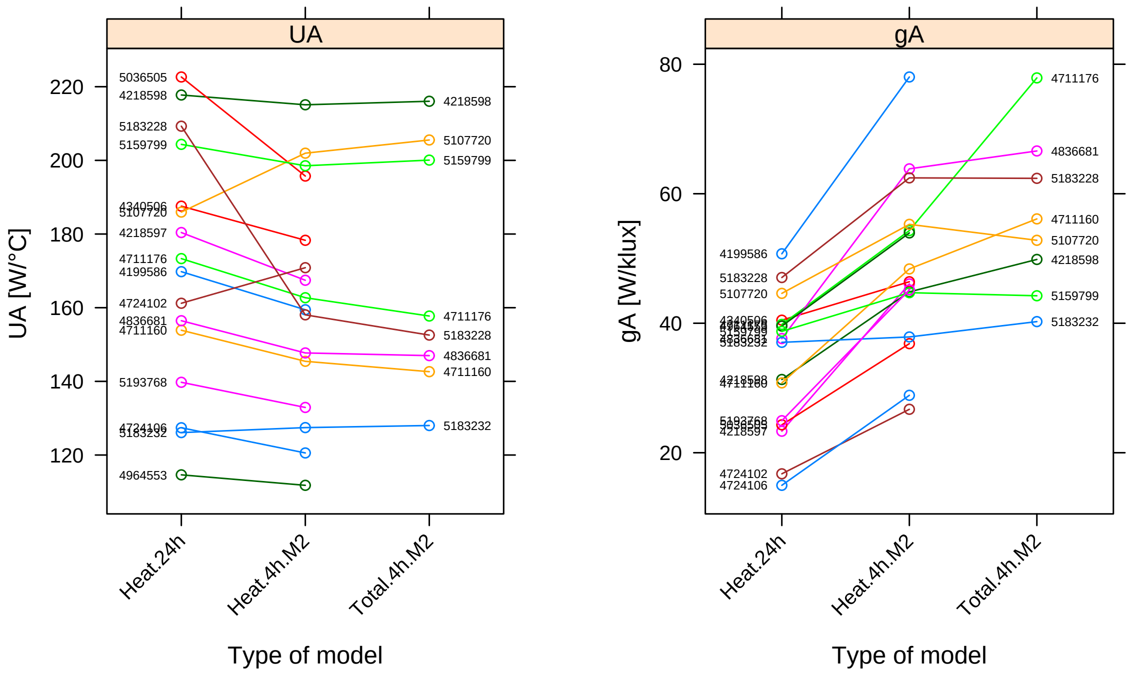

Figure 13: Comparison of UA and gA values for 24h model and 4h model.

7 Conclusion and discussion

This report deals with the estimation of thermal characteristics of single family houses based on measurements of energy consumption and climate. A fundamental assumption of the work is that measurements on a short time scale is available. During the recent years such measurements has become increasingly common, both for electricity, district heating, and natural gas consumption. Obviously, measurements performed with a high frequency imply higher demands for bandwidth and data storage. Maybe less obviously high frequency measurements also require better resolution of the basic measurement equipment. This is because often measurements are performed as the accumulated consumption since installation of the equipment. For these reasons it is desirable if the methods developed are able to estimate the thermal characteristics of the house without using high-frequency data. However, it is also clear that since the methods are based on measurements only it does not make sense to aggregate the data too much. From a practical point of view it is necessary to observe variations in the temperature in order to estimate the effect of the temperature on the consumption. For this reason the largest time scale considered is daily values, but actually measurements on a 10 minute time scale are available in this case. Also, since presumably random variations in the consumption gets more dominating as the averaging period decrease, the shorter averaging period considered is 4 hours (with a few experiments related to a 2 hour averaging period).

The data originates from 56 consumers connected to the district heating system in Sønderborg, Denmark. For a subset of these measurements of electricity consumption are available. Measurements are available covering the period from ultimo September, 2008 to primo December, 2009. However, due to problems with the measurement equipment longer periods without measurements are present in the data and these periods differ between consumers. Also, for a subset of the consumers background knowledge is available.

The main elements of the report are:

(i) Estimation based on the full data period with the aim of establishing methods for identification of the *actual* heating season.

(ii) Estimation of thermal characteristics based on daily averages for the heating season.

(iii) Analysis of the estimated thermal characteristics of the building in relation to background information about the house / household.

(iv) Estimation of thermal characteristics based on 4 hour averages for the heating season. Also, some attention is given to 2 hour averages.

(v) Comparison of estimates based on 4 and 24 hour averages.

Re. (i) methods has been established which allow the *actual* heating season to be estimated for individual households. Presumably, such a method would be applied to data covering a period from summer to summer and would then allow the dates at which the heating system is turned on and off to be estimated. Basically, the method works by assuming and estimating a different response to climate during summer and winter. If the heating system, as expected, is turned off during the summer period the response to climate practically vanish during summer. If this is not the case the consumer can be advised to turn off the heating system off during summer.

Re. (ii) a simple model based on stationary heat transfer is set up. The model allows estimation of the coefficients characterising the response of the building to changes in temperature (UA-value), solar radiation (gA-value), and wind (wA-value). The effect of the wind is characterised both in terms of the wind speed and the wind direction, implying that wA-values are estimated for different wind directions. The effect of the wind can be presented as an increase in the UA-value given a high-wind situation. Here the 99% quantile of the observed wind speed has been selected. This has the further benefit of making the quantity presented independent of spatial variations in the wind speed near the surface. A similar approach is used for the solar radiation where the 99% quantile is used to derive the near-maximum energy input from solar radiation during the heating season.

As part of the model the indoor temperature is estimated. These estimates result in quite reasonable values (values from 17.7 to 25.7 ^oC are obtained). However, the hot tap water consumption of the house which in practice does not contribute to the indoor temperature will result in some bias on the estimates. Given a user interaction, as e.g. available via "My Home" `www.elsparefonden.dk/minbolig`, the user can correct the estimate of the indoor temperature. Consequently, it should be possible to estimate the hot tap-water consumption. For low-energy houses it will probably be beneficial to measure the hot tap-water consumption separately.

It is noted that data regarding the long-wave radiation balance has not been available. This implies that it has not been possible to investigate whether an effect of clear nights on the heat consumption can be observed. However, the model could be extended to incorporate such data.

Estimation has been performed both based on the total energy consumption (district heating and electricity), as well as on the district heating consumption alone. Overall, the estimated UA-values only change marginally (up to 10 $W/^oC$) depending on the response variable used. Furthermore, the sign of the change is not systematic. However, if electrical floor heating is used the estimates of the UA-values increase with the electrical consumption as a fraction of the total energy consumption. As opposed to this the estimates of the indoor temperature increase almost linearly with this fraction. This indicates that electricity consumption, which is not directly used for heating, does not lower the district heat consumption. An investigation of the reason for this is beyond the scope of this report. With respect to the gA-values these are in general estimated somewhat higher when the total energy consumption is used as the response variable as opposed to the case where only the district heating consumption is used. The change in estimated gA-value seem to increase with the electrical consumption as a fraction of the total energy consumption.

As a conclusion it is recommended to use the total energy consumption when estimating the thermal characteristics of a given house. However, if electrical heating is not used measurements of heat consumption will give quite adequate estimates.

Re. (iii) the estimates of UA- and gA-values have been analysed with respect to the background information concerning the houses and households. With respect to the estimated UA-values the significant effects is the ground area of the house, the year of construction, and the number of days per week a wood burning stove is used. For the estimates of the gA-value only the ground floor area and the number of days per week a wood burning stove is used are significant. Assuming a platform like "My Home" allows users anonymously to share information in the sense that such analyses are performed based on available data, it is hence possible to correct the estimated UA-values for errors originating from the use of e.g. a wood burning stove.

Re. (iv) the model described under (ii) is extended to account for the dynamical response to climate. More precisely, dynamic responses to temperature and solar radiation are allowed for, whereas the effect of the wind speed is assumed instantaneous when considering 2 or 4 hour averages as it is the case here. Furthermore, diurnal variations are accounted for. The analysis is mainly concerned with 4 hour averages. This averaging interval was chosen in order to allow for estimation of dynamic effects and at the same time reduce the influence of random variations originating from the timing of the activities of the inhabitants.

On the time scale considered the dynamic effect of temperature and solar radiation are well described by a single time constant. However, for some houses, in the dynamical response of energy consumption on temperature not related to the wind speed there seems to be an instantaneous effect. This effect is presumably related to an other physical time constant somewhat lower than the averaging period. In these cases a two hour averaging period were tried, but the effect seemed to be more well described as an instantaneous effect.

The method is extended with a method for estimating the indoor temperature as a smooth function of time. The use of this method seem to result in more realistic estimates of the time constants. However, this analysis is somewhat uncertain due to the large amount of missing data. Nevertheless, it has been shown that it is possible to estimate time constants based on standard measurements of consumption and climate.

Re. (v) for the dynamic models a natural definition of the thermal characteristics is the stationary gain of the corresponding transfer functions. With this definition the estimates obtained for 4 and 24 hour averages are compared.

For the UA-values there is a good agreement although there is a tendency of somewhat smaller values being obtained when using the dynamic models. The difference is of the order 10 $W/^oC$. For the gA-values the estimates based on 4 hour averages are about 50% larger than the estimates based on 24 hour averages. However, it should be noted that given a dynamic response to solar radiation the stationary gain is never reached. Therefore, it is not expectable that the 4 and 24 hour values are comparable. This also highlights the need to define how the response to solar radiation should be quantified. One approach would be to define it as a day with clear skies, e.g. in the end of March. For such a day the solar radiation can be calculated minute by minute and using such information the effect of the solar radiation will be much more comparable between averaging periods.

A Data summary

This appendix gives an overview of the data for the 40 households used in the analysis of daily average energy consumptions.

The overview table is shown on the next page.

- Columns 1 - 7
  shows the explanatory variables described in Section 3.3 which are derived from questionaires and BBR database. The data is illustrated in Figure 5 and Figure 6. Missing values are indicated by an empty cell.
- Columns 8 - 10
  gives in the first two columns an average of the daily average of heat and electricity consumption in the winter period from 01-10-2008 to 15-04-2009. The following column denoted % is average percentage of electricity consumption out the total consumption (calculated by averaging only for days where both heat and electricity consumption is present).
- Columns 11 - 12
  shows the percentage of missing values in the daily averages of heat and total consumptions used in the analysis. The variables `heat` and `total` are described in Section 3.1.

							Summary for winter 2008/09				
Explanatory variables for households							daily avg. [kWh]			Pct. missing observations	
ID	bebr	stove	elflo	lowt	sqmt	year	heat	elec	%	heat	total
4199586	9	0.0	no	no	138	1969	87.8			26	
4199598	4	4.0	no	no	181	1957	81.2			30	
4218597	2	0.0	no	no	151	1970	89.9			40	
4218598	2	0.0	no	no	163	1969	86.8	10.3	11	30	30
4218600	2	6.0	yes	no	148	1957	26.2	18.4	43	69	69
4340506	3	0.0	no	no	145	1969	85.0			35	
4381443	2	4.0	no	no	169	1962	92.9			36	
4381449	4	0.0	no	yes	138	1964	71.3	12.4	16	42	42
4711160	2	0.0	no	no	112	1964	60.1	7.4	11	40	40
4711161		0.0	no	yes	155	1970	78.0			15	
4711176	2	0.0	no	no	140	1963	81.5	11.5	13	19	19
4724102	2	0.0	no	no	146	1954	75.3			40	
4724106	2	0.0	no	no	86	1952	46.1			26	
4836681	1	0.0	no	no	111	1966	72.5	1.3	2	37	37
4836722	3	0.0	yes	no	166	1969	91.1	8.3	9	45	45
4964512	2	0.0	yes	no	96	1991	32.1			45	
4964553	2	0.0	no	no	119	1963	42.1			46	
4986050	2	4.0	no	no	140	1964	56.2	7.6	13	26	26
5036505	3	0.0	no	no	119	1947	83.2			31	
5069878	4	0.0	yes	no	86	1970	58.0	8.4	12	30	31
5069913	2	0.0	yes	yes	152	1976	66.9	17.0	21	65	65
5107720	4	0.0	no	no	160	1965	73.5	6.1	8	17	17
5127784	2	0.0	yes	no	193	1960	88.6	50.8	37	24	24
5159799	1	0.0	no	no	173	1965	97.1	4.8	5	45	45
5164474	2	6.0	no	no	130	1957	64.4	15.4	20	42	42
5164485	2	1.5	yes	no	130	1963	76.4	17.3	20	54	54
5164523	2	0.0	yes	no	116	1965	62.8	9.0	13	43	43
5164534	2	0.0	no	yes	135	1996	51.7			33	
5168264	3	4.0	no	no			73.6	0.0	0	15	15
5183206	2	0.0	yes	no	90	1959	60.4	17.6	23	40	40
5183228	2	0.0	no	no	146	1963	85.5	4.1	5	54	54
5183232	2	0.0	no	no	122	1966	53.8	9.0	14	37	37
5191179	2	0.0	yes	no	112	1959	56.7	10.2	15	54	54
5193768	2	0.0	no	no	136	1975	60.9			20	
5194940	5	6.0	yes	no	114	1938	59.9	17.8	24	49	49
5194965	2	0.0	no	yes	123	1965	62.7	4.8	7	58	58
5197381	8	0.0	no	no	127	1953	73.6	12.3	15	67	67
5219090	5	6.0	yes	no	150	1952	44.5			39	
5223030	2	4.0	yes	no	165	1957	94.2	13.4	13	23	23
5223036	2	0.0	no	yes	137	1967	61.0	15.3	20	45	45

B Time varying estimates

Figure 14 and Figure 15 show time varying estimates of coefficients in Eq. (2) based on total energy consumption for all houses with electricity data. The curves are plotted with a dashed line when the daily energy consumption is missing.

Time variations in coefficients are estimated using first order local polynomials with fixed bandwith of 60 days for the kernel.

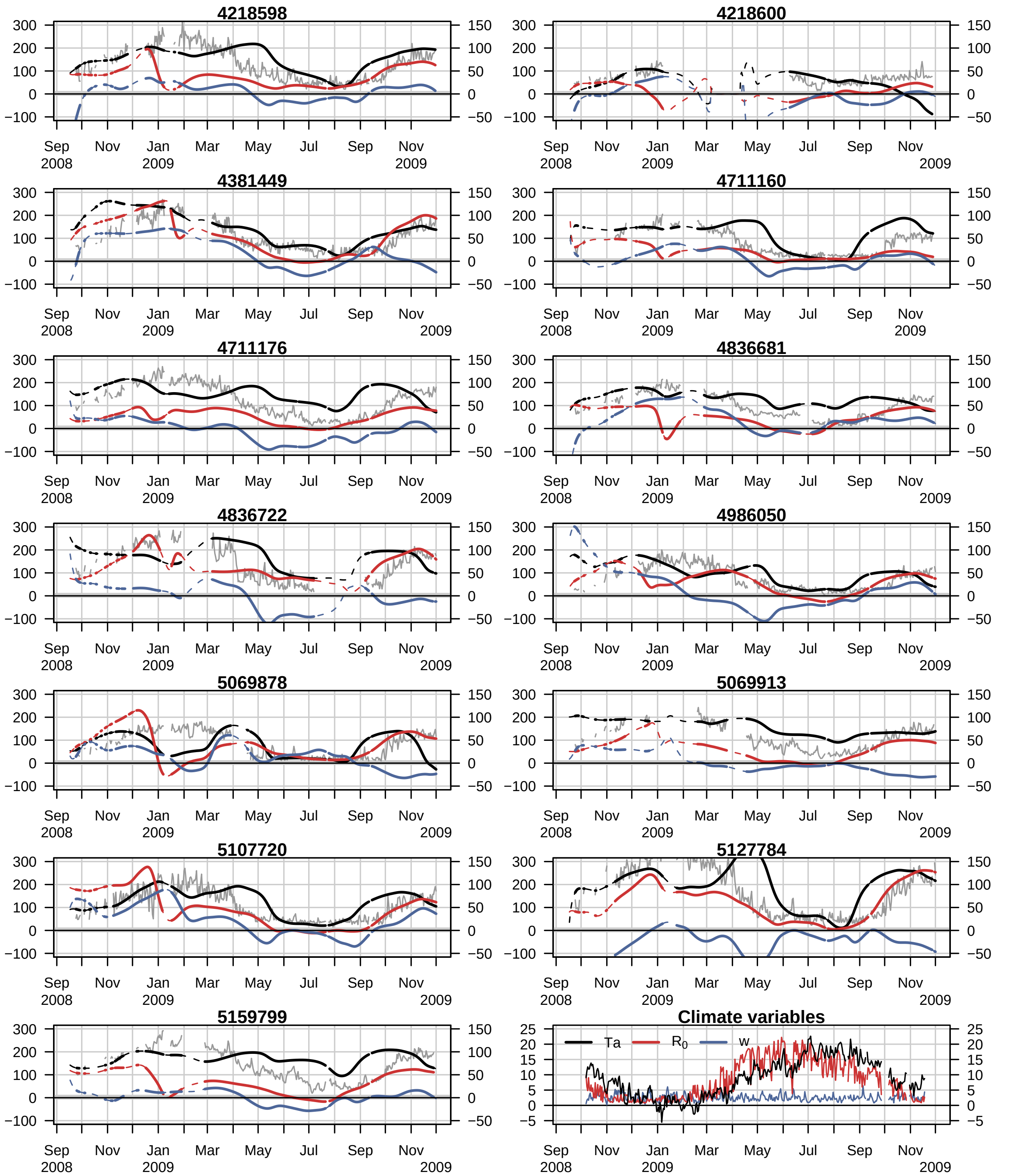

Figure 14: Time varying estimates of coefficients in Eq. (2). Black is UA [W/°C], red is gA [W/klx] and blue is b_1 [W/m/s] all measured on the left side axis. The underlying gray curve is daily total energy consumption in kWh measured on the right side axis.

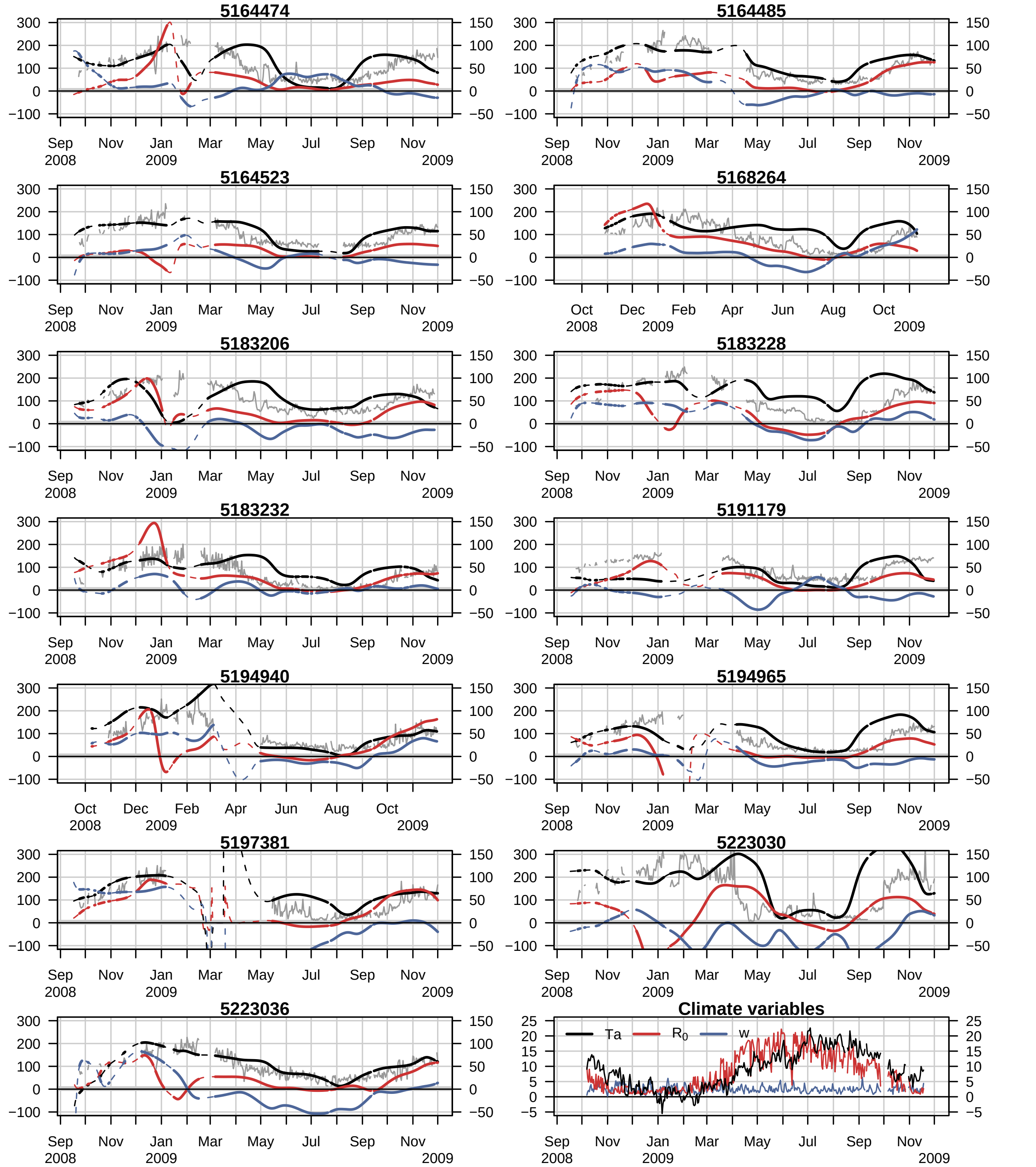

Figure 15: Time varying estimates of coefficients in Eq. (2). Black is UA [W/°C], red is gA [W/klx] and blue is b_1 [W/m/s] all measured on the left side axis. The underlying gray curve is daily total energy consumption in kWh measured on the right side axis.

C Parametric modelling of time-variations

Figure 16 and Figure 17 show estimates of a parametric model for time variations of the coefficients in Eq. (2) based on total energy consumption for all houses with electricity data. The estimation failed to converge due to a long period of missing observations from January to June for meters no. 4218600 and 5197381.

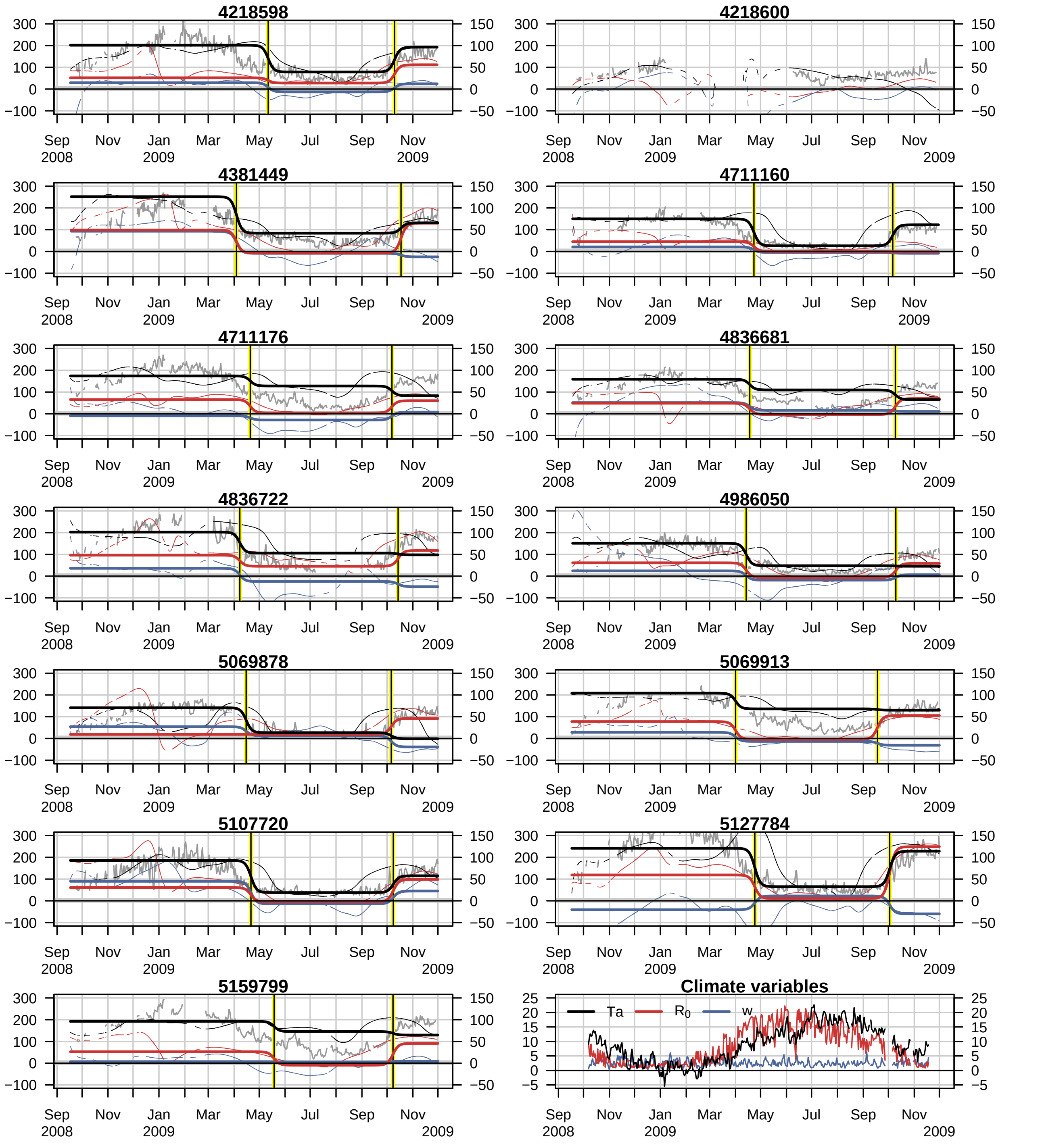

Figure 16: Thick lines are estimates of parametric models for time varying coefficients in Eq. (2). Black is UA [W/°C], red is gA [W/klx] and blue is b_1 [W/m/s] all measured on the left side axis. Thin lines are estimates based on local regression, see App. B. Vertival black-yellow lines are estimated time point of change. The underlying gray curve is daily total energy consumption in kWh measured on the right side axis.

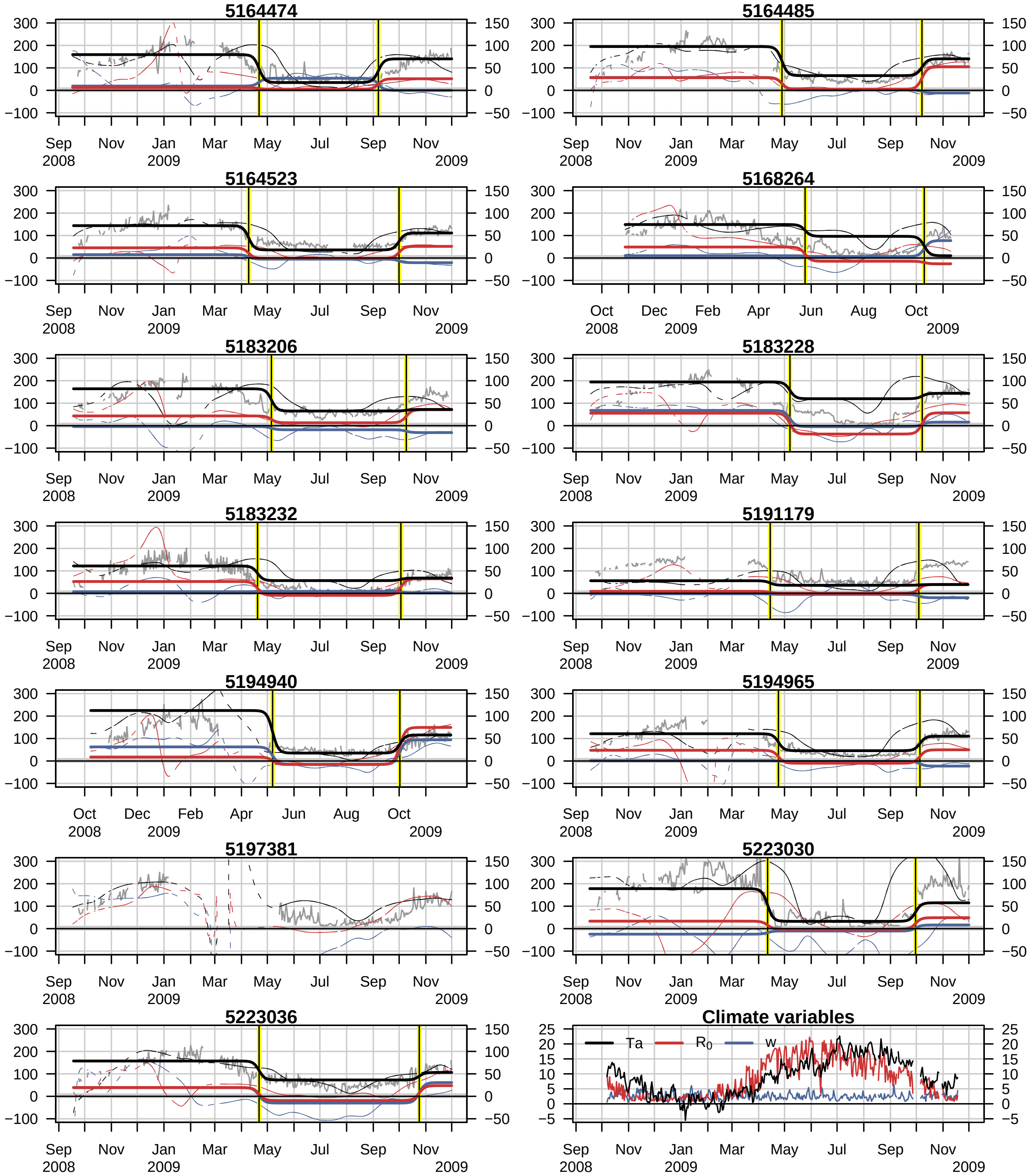

Figure 17: Thick lines are estimates of parametric models for time varying coefficients in Eq. (2). Black is UA [W/°C], red is gA [W/klx] and blue is b_1 [W/m/s] all measured on the left side axis. Thin lines are estimates based on local regression, see App. B. Vertival black-yellow lines are estimated time point of change. The underlying gray curve is daily total energy consumption in kWh measured on the right side axis.

D Residual plots

Figure 18 and Figure 19 show residual plots for all 26 households used for estimation of model in Section 5.3.

Figure 20 and Figure 21 show estimated auto-correlations the residuals.

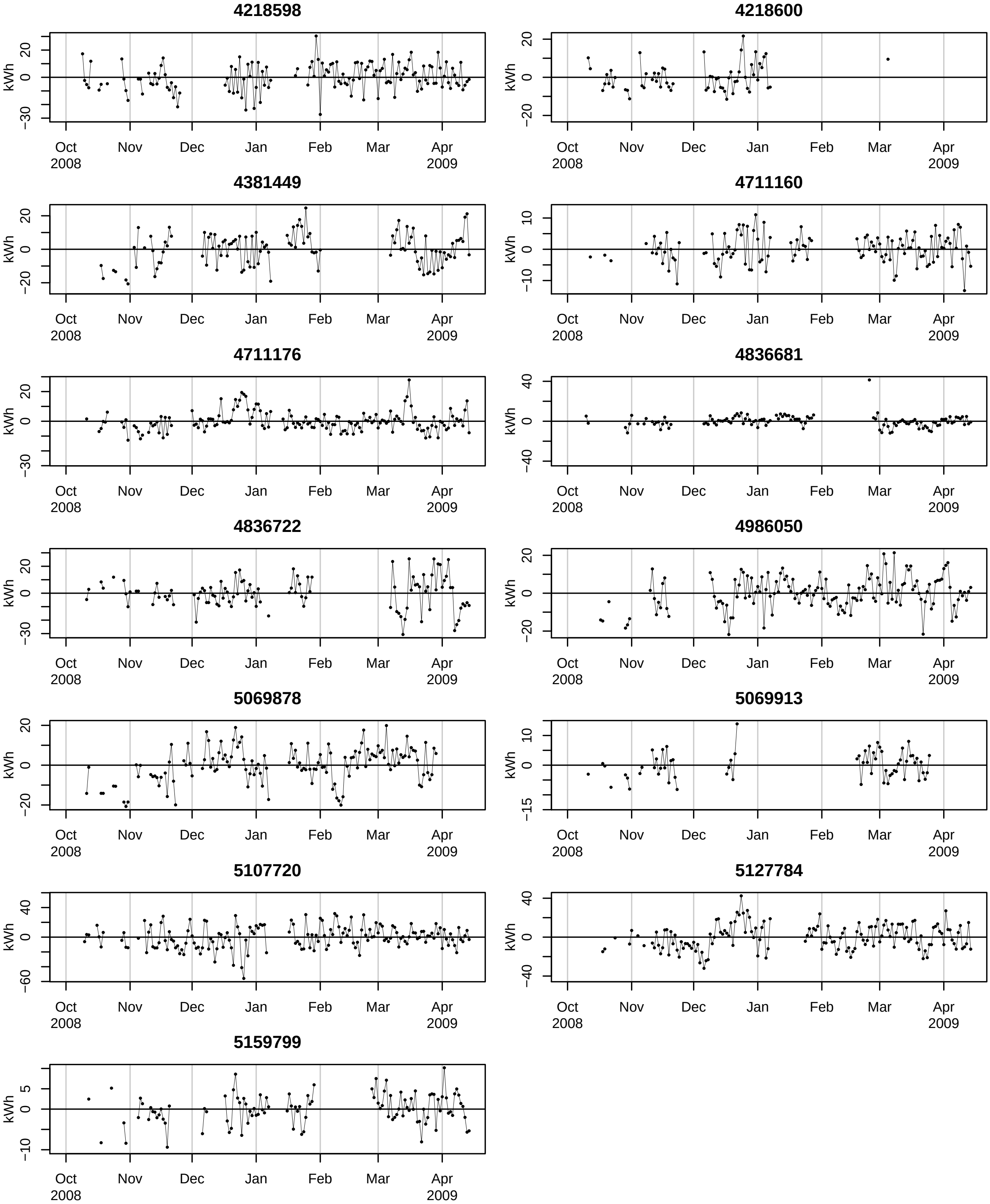

Figure 18: Residual plots for model estimates in Section 5.3.

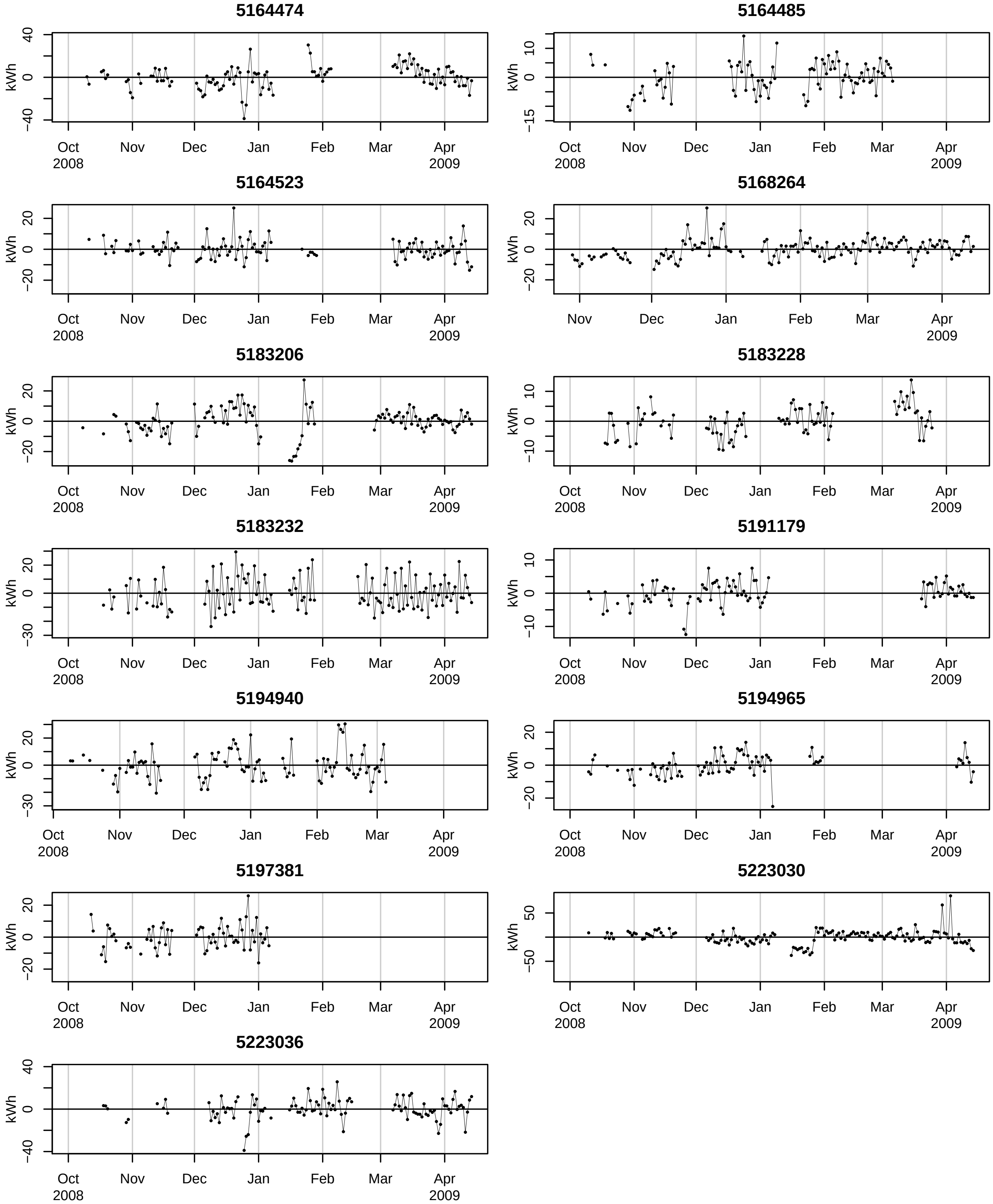

Figure 19: Residual plots for model estimates in Section 5.3.

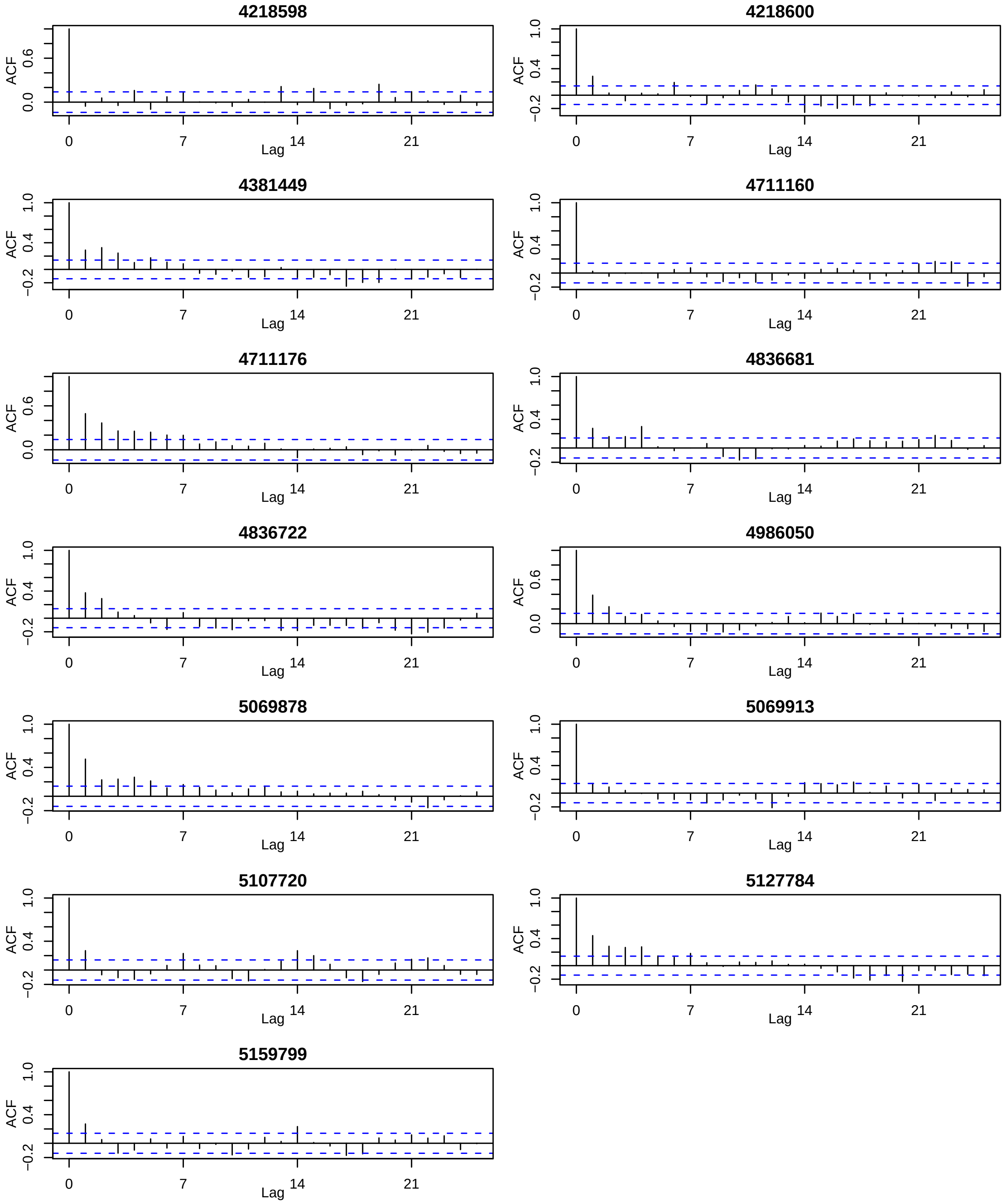

Figure 20: Auto-correlation for residuals for model estimates in Section 5.3.

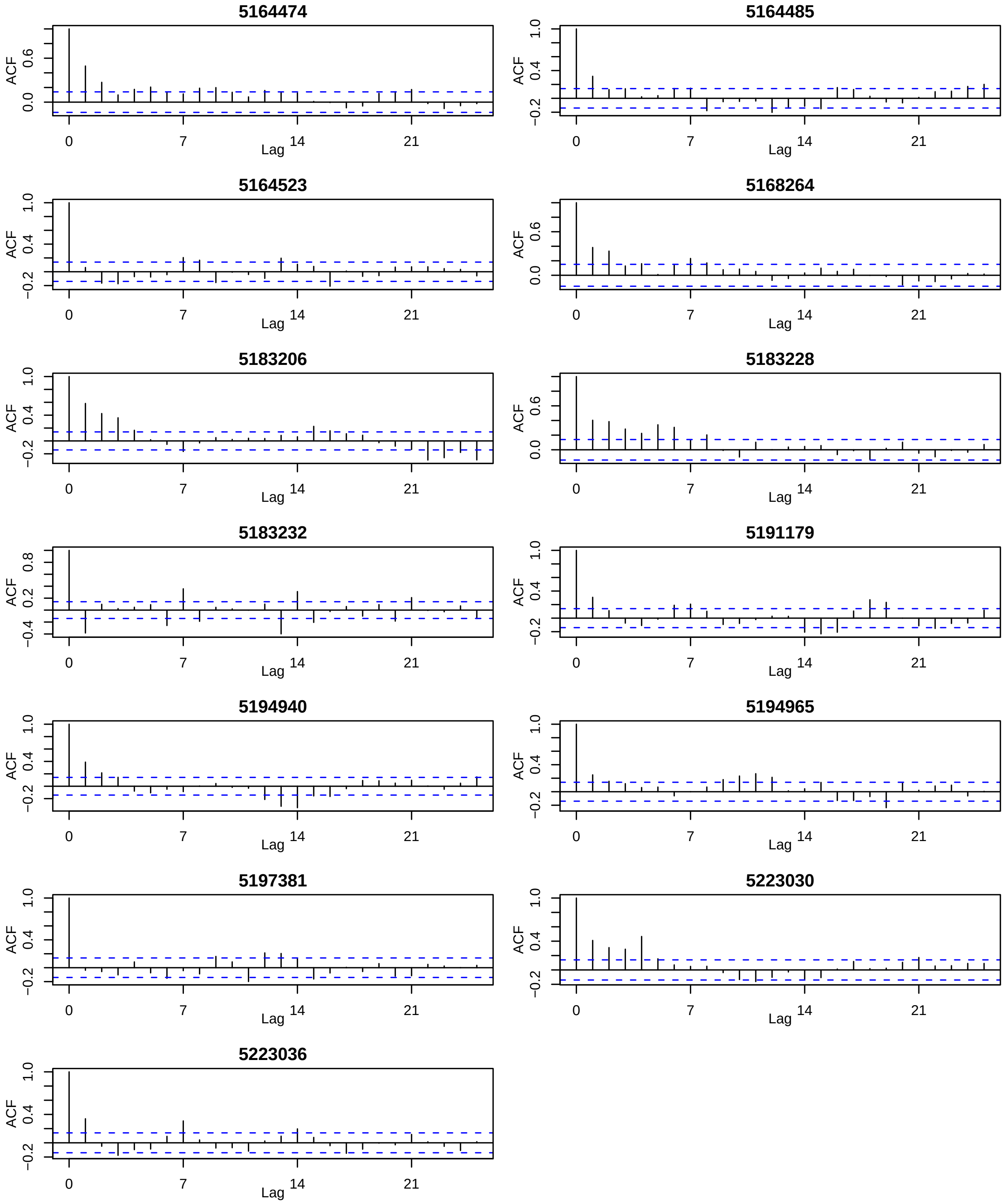

Figure 21: Auto-correlation for residuals for model estimates in Section 5.3.

E Estimation using only heat consumption

	UA W/°C	σ_{UA}	$\text{gA}^{\max}$ W	$v_E^{\max}$ W/°C	$v_S^{\max}$ W/°C	$v_W^{\max}$ W/°C	T_i °C	σ_{T_i}	p_{gA}	p_v
4199586	169.8	8.0	779.3	15.0	10.9	11.3	25.7	1.2	0.00	0.05
4199598	204.3	8.7	606.3	58.2	57.0	54.4	19.2	0.7	0.00	0.00
4218597	180.4	6.9	358.3	2.4	2.2	11.4	24.4	0.9	0.00	0.02
4218598	217.7	8.6	481.0	10.5	7.9	6.1	21.0	0.8	0.00	0.69
4340506	187.6	8.7	621.7	3.3	5.6	6.8	22.9	1.0	0.00	0.64
4381443	211.3	9.0	356.5	1.4	13.0	16.7	22.9	0.9	0.00	0.01
4381449	220.4	10.5	799.5	31.7	35.0	34.5	17.7	0.7	0.00	0.00
4711160	153.9	5.8	472.9	16.3	5.3	11.0	20.6	0.7	0.00	0.01
4711161	180.6	5.4	548.2	9.3	8.9	13.2	22.0	0.6	0.00	0.02
4711176	173.3	4.8	611.7	5.9	1.0	12.8	23.8	0.6	0.00	0.00
4724102	161.2	7.0	257.2	11.8	17.6	22.8	21.5	0.9	0.03	0.00
4724106	127.4	4.0	230.0	11.6	11.9	8.1	18.6	0.6	0.00	0.05
4836681	156.5	8.1	578.1	41.4	29.3	22.2	23.0	1.1	0.00	0.00
4964553	114.6	5.7	607.8	6.3	12.3	8.9	21.3	0.9	0.00	0.03
4986050	154.5	10.3	664.0	11.1	8.7	5.5	19.2	1.3	0.00	0.65
5036505	222.6	8.1	373.5	3.5	12.6	14.9	18.7	0.6	0.01	0.14
5107720	186.0	14.4	685.4	42.1	30.6	16.3	20.1	1.5	0.00	0.06
5159799	204.3	4.9	595.2	-1.8	-1.5	5.0	25.0	0.6	0.00	0.01
5164474	170.2	11.4	12.6	26.3	3.6	-5.4	19.8	1.2	0.90	0.01
5164534	126.4	6.2	455.2	12.0	12.5	8.2	20.9	0.9	0.00	0.02
5168264	169.7	7.7	554.0	25.8	8.3	2.1	21.7	0.9	0.00	0.00
5183228	209.3	7.7	723.0	21.7	19.0	34.2	20.8	0.6	0.00	0.00
5183232	126.1	9.6	569.2	8.6	2.7	2.5	22.8	1.6	0.00	0.85
5193768	139.7	4.1	383.2	1.8	6.6	11.6	21.9	0.6	0.00	0.00
5194965	131.7	8.6	356.7	-3.5	-1.1	8.6	24.9	1.6	0.00	0.02
5197381	185.3	9.1	870.4	38.1	27.5	27.5	21.2	1.0	0.00	0.00
5223036	170.0	9.6	381.0	8.3	-3.7	9.0	18.8	1.0	0.00	0.08

Table 8: Estimates based on only district heating consumption for households not using electric underfloor heating.

F Estimation of time-variation of indoor temperature

In the first step the dynamical models no. 1 to 3 described in Section 6 are estimated assuming a constant indoor temperature T_i. In a second step the residuals e_t from the first step are used to estimate an apparent indoor temperature $T^*_{i,t}$. The apparent temperature is then smoothed to give $T_{i,t}$. For reasons which will be explained, only the smoothed estimate can be considered an actual estimate of the variation in temperature.

The estimate of the apparent indoor temperature $T^*_{i,t}$ is found by removing the noise term from the statistical model and then, for each point in time, calculating the value of the indoor temperature which makes the result in a perfect fit to the observered heat consumption (given the estimated coefficeints from the previous time estimation step). This requires that the filtered indoor temperature can be approximated by the product of the stationary gain times the indoor temperature. The estimate is smoothed using local linear regression with a bandwidth of 4 days ($\pm$4 days) to obtain $T_{i,t}$. Due to the assumption just described only the smoothed estimate can be considered a temperature.

The estimates of the model parameters can now be updated by using the indoor temperature as $a + bT_{i,t}$, where a and b are estimated from data. Fitting the model as described gives updated estimates of all model parameters and also an updated indoor temperature given as

$$\hat{T}_{i,t} = \hat{a} + \hat{b}T_{i,t}. \tag{4}$$

Figure 22 shows a comparison of estimates for UA and gA values based on Model 2 using either a constant or variable indoor temperature. For the UA values there seems to be a tendency to down estimation for some households when the variable estimate of the indoor temperature is used whereas others remain fairly constant. Given that the indoor temperature is correctly estimated, the lower estimate of the UA value is as expected if a household has a tendency to raise the indoor temperature when it gets cold out side.

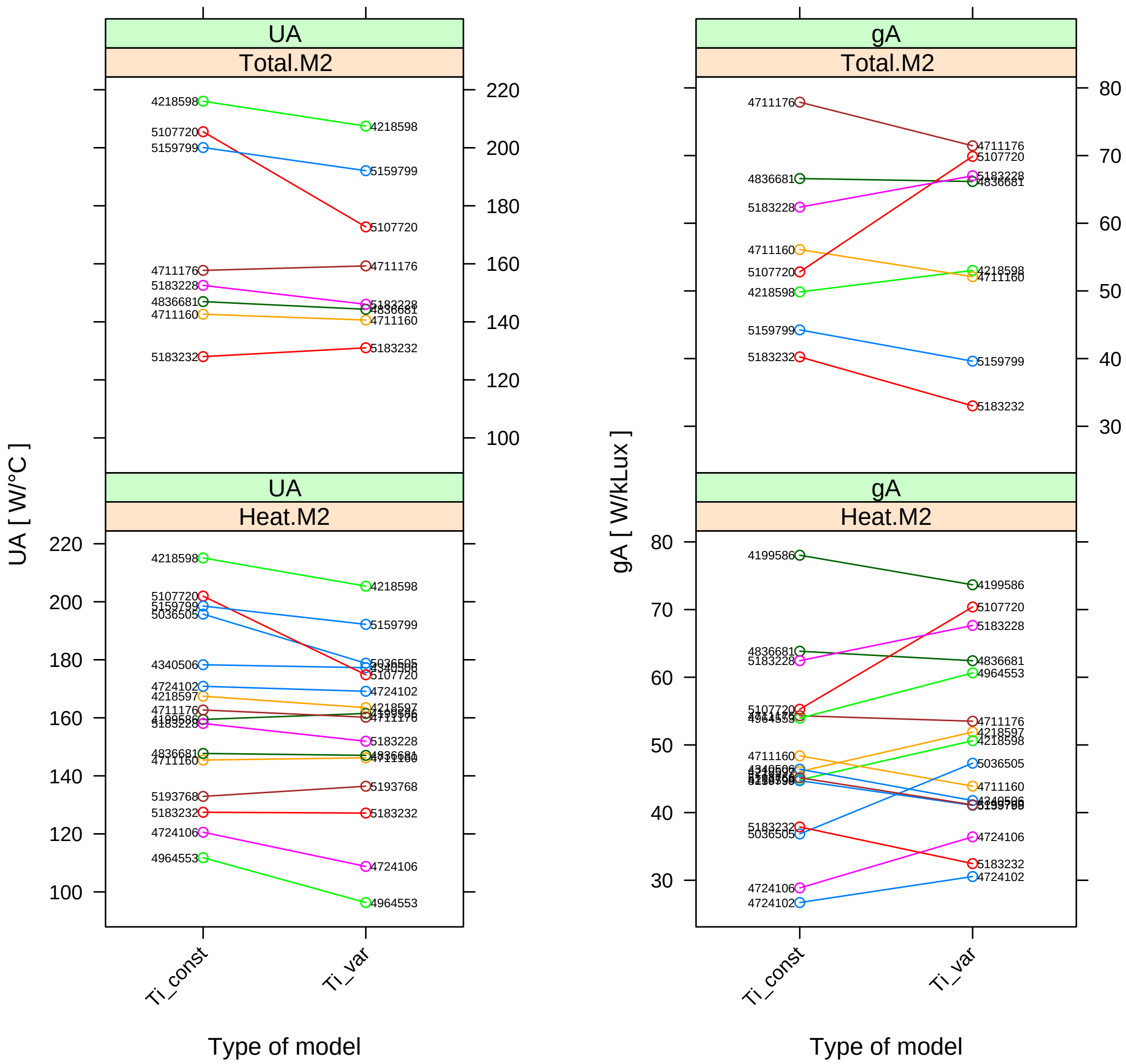

Figure 22: Comparison of estimates for UA and gA values based on Model 2 using either a constant or variable indoor temperature.

G Estimates based on 4h dynamical models

This appendix contains a comparison of estimates based on dynamical models no. 1 and 2 (see Sec. 6) using both `heat` and `total` energy consumption as response. Estimates for UA and gA values are found in Figure 12 p. 26.

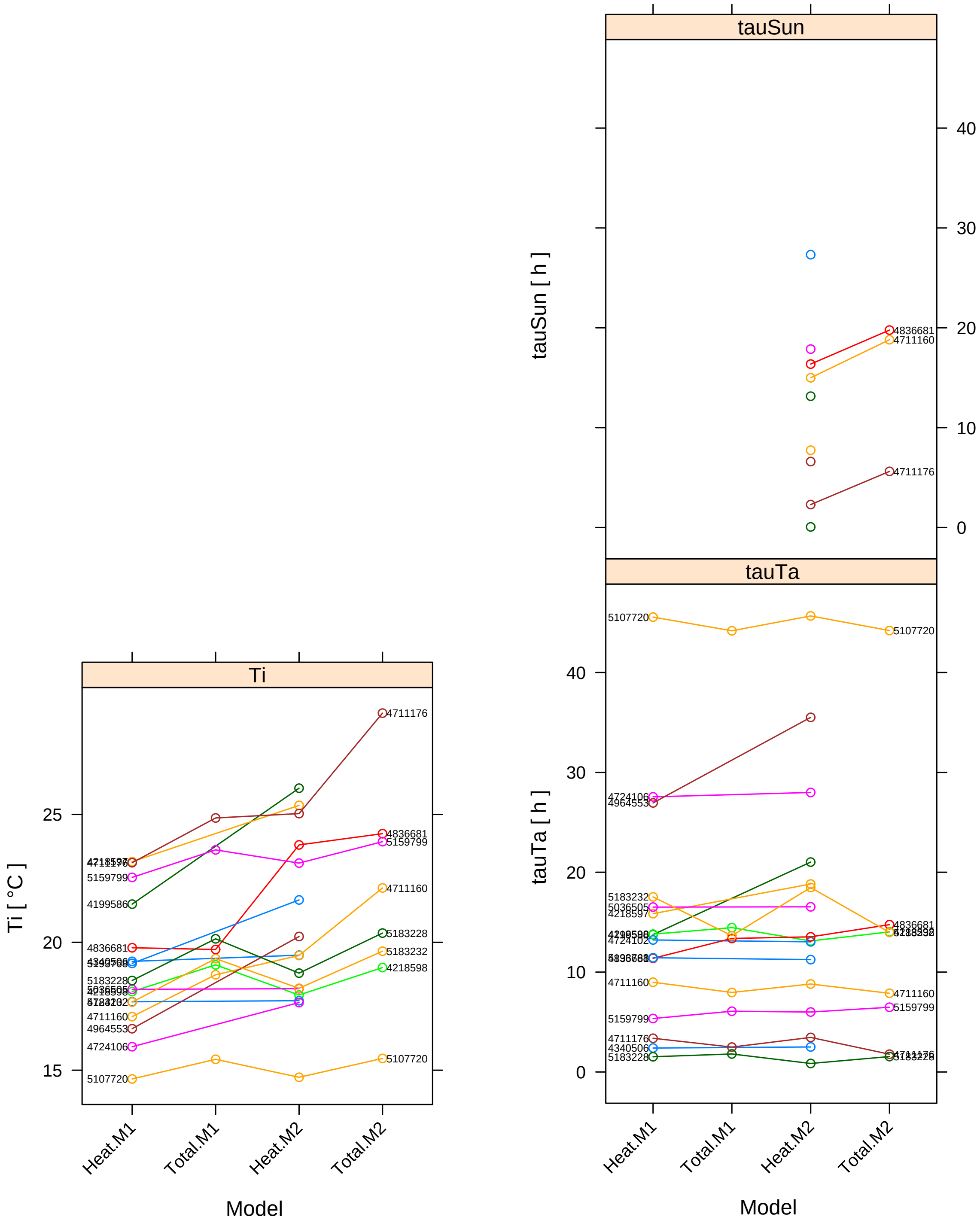

Figure 23: Estimates for T_i (left) and time constants τ_{R_0} and τ_{T_a} (right).

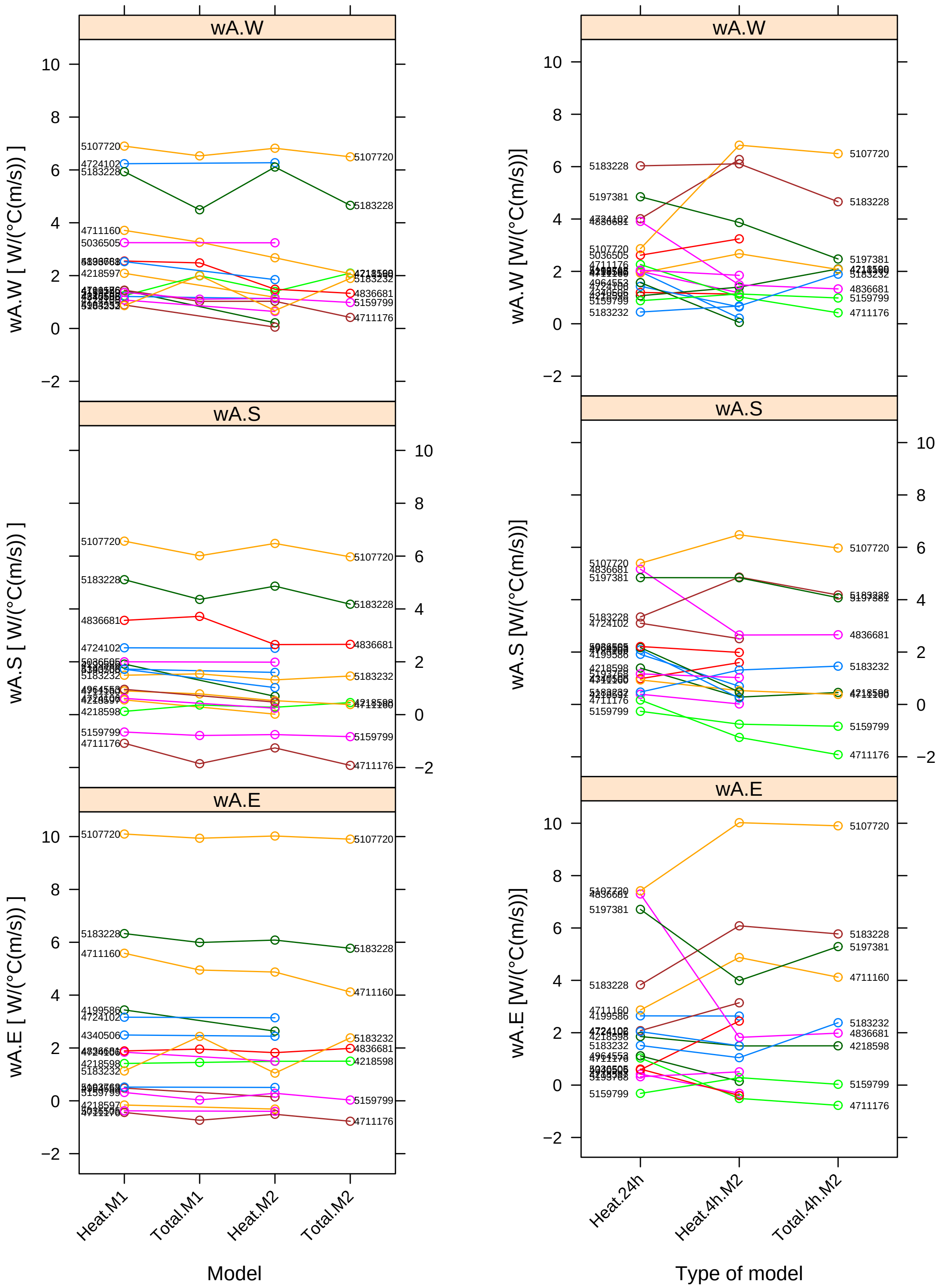

Figure 24: Estimates for wA.E, wA.S, wA.W. (left) and comparison with 24h values (right).

H Detailed modelling of the dynamic response on temperature

As described in Section 6 the estimation procedure often does not converge for model no. 3. This model is characterised by having two real poles in the transfer function related to temperature. Since non-convergence of the estimation procedure may be related to the quality and amount of the data used, five consumers with relatively good data has been selected. For three of these, measurements of electricity consumption is available also.

The data are depicted in Figures 25 and 26, placed near the end of this appendix. The consumption data are shown together with the ambient air temperature in a linear cave plot[1].

In this appendix focus is on the estimated poles and resulting time constants in the transfer function related to the temperature, i.e. $H_{T_a}(q)$, for models no. 2 and 3 described in Section 6, when applied to 4 hour averages. Furthermore, model no. 3 is applied to 2 hour averages. Finally a new model no. 4 is defined by considering an immediate response together with a simple transfer function as used e.g. in model no. 2.

- Model no. 4

$$H_{T_a}(q) = 1 + \frac{b_0}{1 + a_{T_a,1}q^{-1}} = \frac{b_0 + a_{T_a,1}q^{-1}}{1 + a_{T_a,1}q^{-1}}$$

$$H_{R_0}(q) = \frac{1}{1 + a_{R_0,1}q^{-1}}$$

This model is applied to 4 hour averages, since on that time scale it is indeed plausible that part of the response will be immediate. If the hypothesis is correct, model no. 4 should result in higher time constants than model no. 2.

In each case we also consider the estimates obtained when these are updated after estimation of the indoor temperature. These are denoted "T_i var." as opposed to the case of constant indoor temperature denoted "T_i const.".

Table 9 show the estimated poles and Table 10 show the corresponding time constants for the positive poles. Comparing the estimates obtained for model no. 3 for 2 and 4 hour averaging periods, it is seen that the reduction of the averaging period does not resolve the convergence issues. Furthermore, it is seen that for the 2 hour averaging period negative poles are estimated

[1]Linear cave plots are characterised by the fact that if the relation between the two series are linear and time-invariant the visual distance between the two series are constant over time. See also `http://www2.imm.dtu.dk/~han/pub/caveplot/`.

Identification			4199586	4218598	4711176	5107720	5193768
Heat.4h.M2	T_i const.		0.852	0.791	0.579	0.923	0.768
	T_i var.		0.752	0.656	0.546	0.796	0.763
Heat.4h.M4	T_i const.		0.968	0.972	0.759	0.928	0.862
	T_i var.		0.845	0.956	0.752	N/C	0.826
Heat.4h.M3	T_i const.	min.	0.661	0.416	-0.543	N/C	-0.115
	T_i const.	max.	0.971	0.980	0.700	N/C	0.844
	T_i var.	min.	0.674	0.451	-0.517	–	-0.252
	T_i var.	max.	0.968	0.971	0.688	–	0.803
Heat.2h.M3	T_i const.	min.	N/C	-0.260	-0.636	N/C	-0.435
	T_i const.	max.	N/C	0.986	0.863	N/C	0.917
	T_i var.	min.	–	-0.301	-0.328	–	-0.398
	T_i var.	max.	–	0.977	0.857	–	0.894
Total.4h.M2	T_i const.		–	0.800	0.479	0.920	–
	T_i var.		–	0.644	0.437	0.777	–
Total.4h.M4	T_i const.		–	0.963	N/C	0.927	–
	T_i var.		–	0.941	–	N/C	–
Total.4h.M3	T_i const.	min.	–	0.514	-0.728	N/C	–
	T_i const.	max.	–	0.977	0.604	N/C	–
	T_i var.	min.	–	0.408	-0.532	–	–
	T_i var.	max.	–	0.961	0.590	–	–
Total.2h.M3	T_i const.	min.	–	-0.139	-0.819	N/C	–
	T_i const.	max.	–	0.981	0.818	N/C	–
	T_i var.	min.	–	-0.209	-0.259	–	–
	T_i var.	max.	–	0.971	0.828	–	–

Table 9: Estimated poles, "N/C" indicated that the estimation did not converge and "–" indicates that the estimation was not attempted.

in all cases where the estimation procedure converge. Since negative poles corresponds to an "overshoot" in every second lag (starting in lag 0) of the step-response corresponding to that pole, these negative poles indicates that the data are well described by one pole. This observation is consistent with the root mean square of the in-sample model errors as shown in Table 11, which show that with respect to the quality of the fit there is only little difference between model no. 3 and 4. Please note that results for model no. 3 fitted to 2 hour averages are not included in the table since these are not directly comparable to the 4 hour values.

Following this it is concluded that for these consumers only models no. 2 and 4 should be considered further. For model no. 4 it is seen that the estimated time constants seems unrealistically high when the indoor temperature is assumed constant. However, when a slowly varying indoor temperature is accounted for the time constants are much more realistic[2]. Considering heat consumption as the response variable, the estimation procedure for model no. 4 does not converge in one case where a varying indoor temperature is accounted for. This could be cause by a misfit of model no. 4 to the particular data. E.g. if the a_{Ta_a} parameter in the numerator should be small because model no. 2 is more appropriate for the specific data, then this is

[2]With the possible exception of the house with identification number 4218598 (from year 1969) for which a time constant of 84 hours is estimated when using the heat consumption as the response variable.

Identification			4199586	4218598	4711176	5107720	5193768
Heat.4h.M2	T_i const.		21.0	13.1	3.5	45.7	11.3
	T_i var.		10.2	5.7	2.9	13.6	10.9
Heat.4h.M4	T_i const.		117.9	135.2	10.6	49.6	23.0
	T_i var.		19.7	83.9	10.2	N/C	17.0
Heat.4h.M3	T_i const.	min.	5.9	0.8	7.3	N/C	19.7
	T_i const.	max.	134.2	195.7	7.3	N/C	19.7
	T_i var.	min.	6.3	1.3	6.9	–	14.3
	T_i var.	max.	118.2	132.5	6.9	–	14.3
Heat.2h.M3	T_i const.		N/C	140.0	11.6	N/C	21.0
	T_i var.		–	83.5	11.0	–	15.8
Total.4h.M2	T_i const.		–	14.0	1.8	44.2	–
	T_i var.		–	5.3	1.1	11.8	–
Total.4h.M4	T_i const.		–	102.1	N/C	48.6	–
	T_i var.		–	62.2	–	N/C	–
Total.4h.M3	T_i const.	min.	–	2.3	4.0	N/C	–
	T_i const.	max.	–	171.5	4.0	N/C	–
	T_i var.	min.	–	0.7	3.7	–	–
	T_i var.	max.	–	95.4	3.7	–	–
Total.2h.M3	T_i const.		–	103.9	8.0	N/C	–
	T_i var.		–	65.7	8.6	–	–

Table 10: Estimated time constants in hours, "N/C" indicated that the estimation did not converge and "–" indicates that the estimation was not attempted. Compared with Table 9 this table contains fewer rows since time constants are not reported for negative poles. In the case of negative poles the same time constant may be reported for as both minimum and maximum.

Identification		4199586	4218598	4711176	5107720	5193768
Heat.4h.M2	T_i const.	470	790	551	1361	246
	T_i var.	418	746	533	1317	220
Heat.4h.M4	T_i const.	451	771	550	1362	245
	T_i var.	424	752	531	N/C	224
Heat.4h.M3	T_i const.	445	770	550	N/C	244
	T_i var.	416	749	530	–	223

Table 11: RMSE of model residuals (in-sample errors) in W as averages over 4 hours for the models fitted to heat consumption data. 'N/C" indicated that the estimation did not converge and "–" indicates that the estimation was not attempted.

problematic since (i) the same parameter occurs in the denominator and (ii) if the parameter gets close to zero both the parameter b_0 and the parameter c_{T_a} in the full model can vary freely (which result in non-uniqueness of the estimates, which in turn will result in non-convergence of the estimation procedure). It is therefore suggested to revert to model no. 2 in case the estimation procedure for model no. 4 does not converge. Also, it is suggested to allow for a varying indoor temperature since this seems most realistic.

The estimated step responses and indoor temperatures are shown in Figures 27 (page 56) and 28 (page 57), respectively. It is noted that the estimated indoor temperature for 4199586

Identification		4199586	4218598	4711176	5107720	5193768
Heat.4h.M2	T_i const.	159 (24)	215 (45)	163 (69)	202 (16)	133 (31)
	T_i var.	162 (40)	205 (71)	160 (73)	175 (36)	136 (32)
Heat.4h.M4	T_i const.	155 (64)	235 (123)	165 (92)	203 (28)	135 (52)
	T_i var.	131 (58)	209 (127)	162 (99)	N/C	133 (54)
Total.4h.M2	T_i const.	–	216 (43)	158 (82)	206 (16)	–
	T_i var.	–	207 (74)	159 (90)	173 (39)	–
Total.4h.M4	T_i const.	–	234 (116)	N/C	207 (30)	–
	T_i var.	–	212 (127)	–	N/C	–

Table 12: Estimated UA and, in parentheses, immediate response (lag 0) of the transfer function from temperature in $W/^oC$. "N/C" indicated that the estimation did not converge and "–" indicates that the estimation was not attempted.

is quite high. This house has a floor area of 138 m^2 and is inhabited by 9 persons (2 adults, 3 teenagers, and 4 children under 12 years of age). Apparently, the estimation of the indoor temperature does not supply appropriate values in such cases. Possibly, this could explained by a high fraction of the consumption not being related to the climate and a lager need for ventilation of the house.

Table 12 show the estimated UA values together with the estimate of the immediate response. As expected it is seen that, while the estimated UA values are very similar, model no. 4 results in higher estimates of the immediate response. In fact, since the model is applied to 4 hour averages this immediate response actually account for a small time constant in the physical system.

The results obtained when applying the models to the total of the heat and the electricity consumption seems to raise some issues related to the convergence of the estimation procedures. However, more reliable data are needed in order to investigate this further.

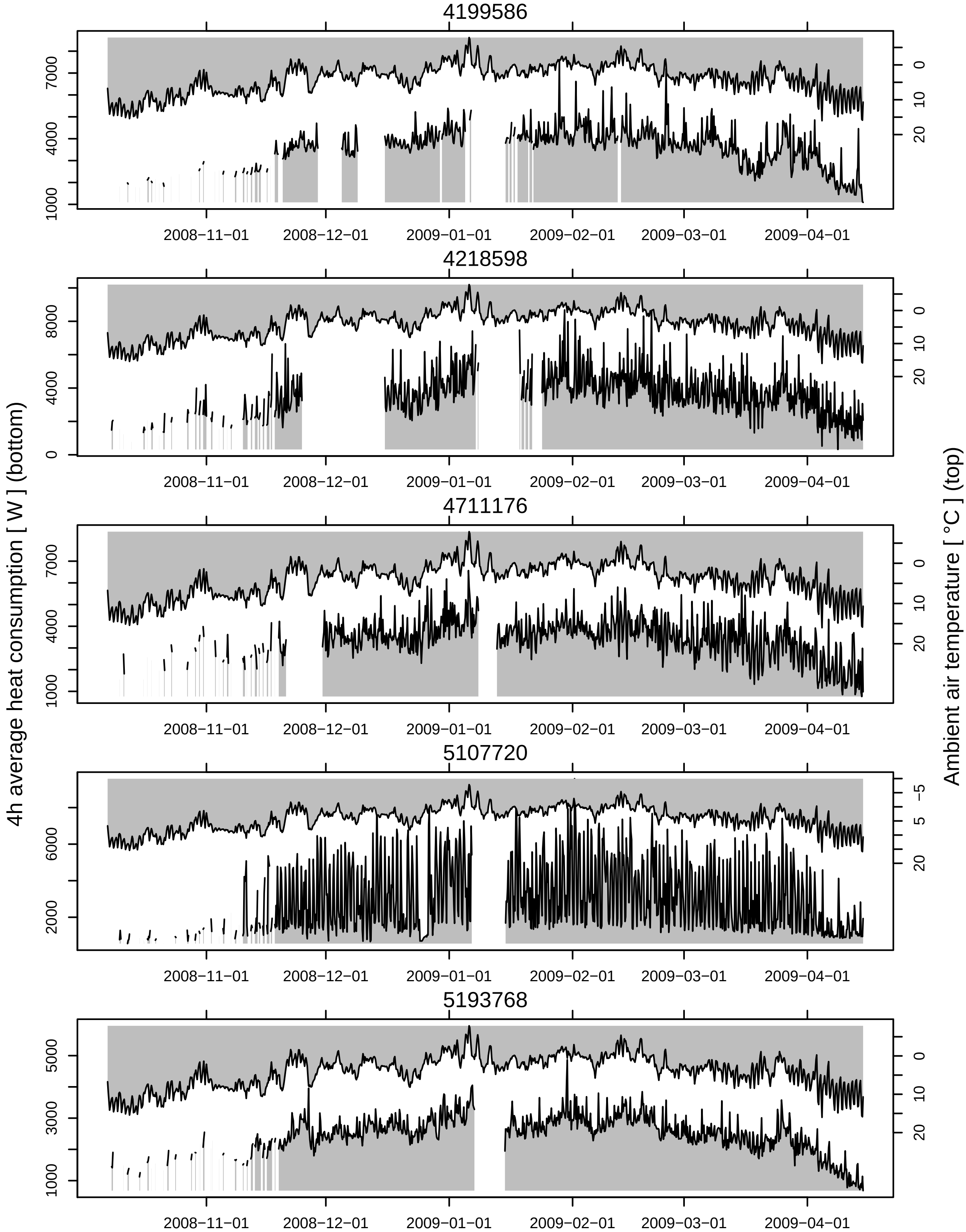

Figure 25: Measurements of heat consumption and ambient air temperature for five selected consumers.

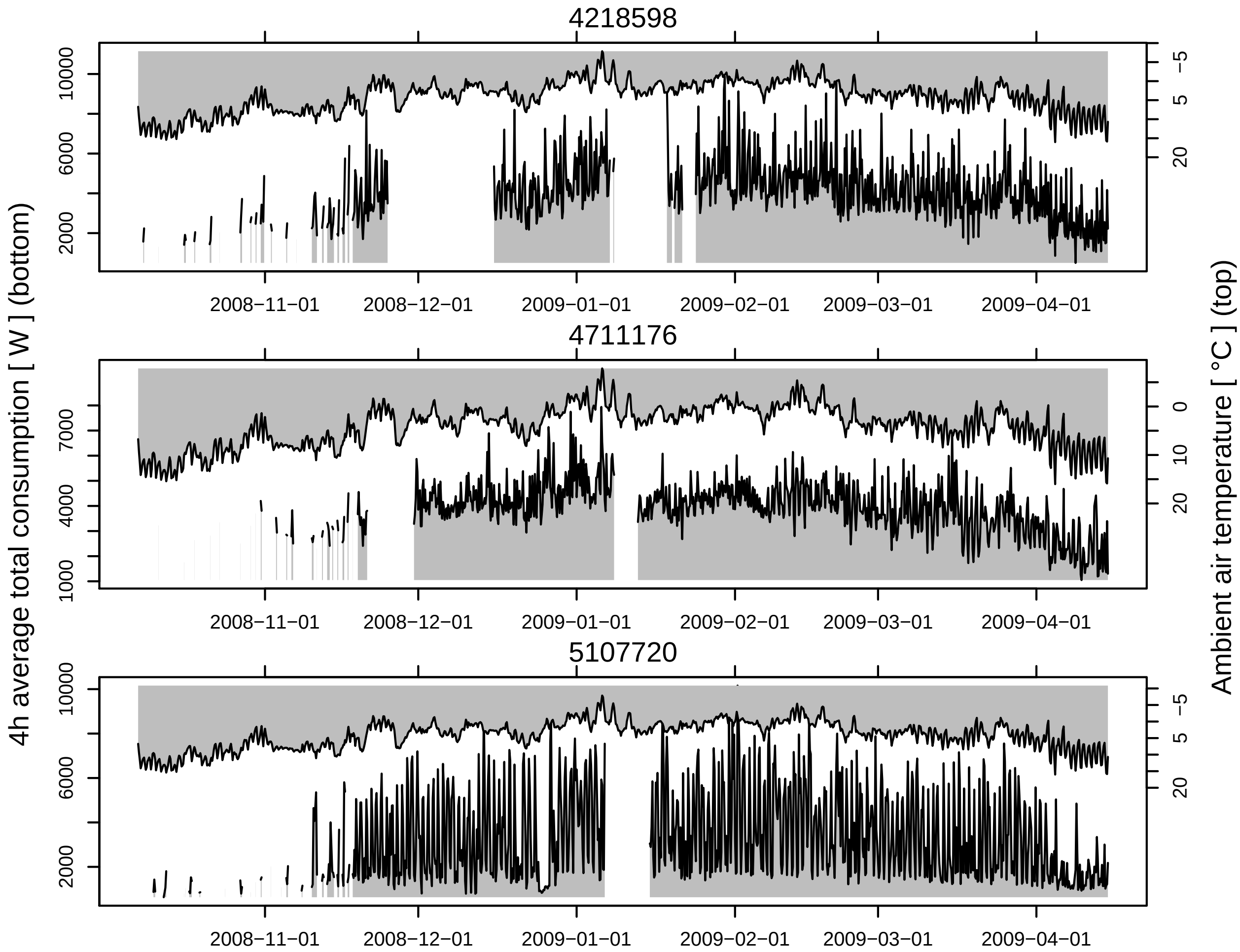

Figure 26: Measurements of total (heat and electricity) consumption and ambient air temperature for three selected consumers.

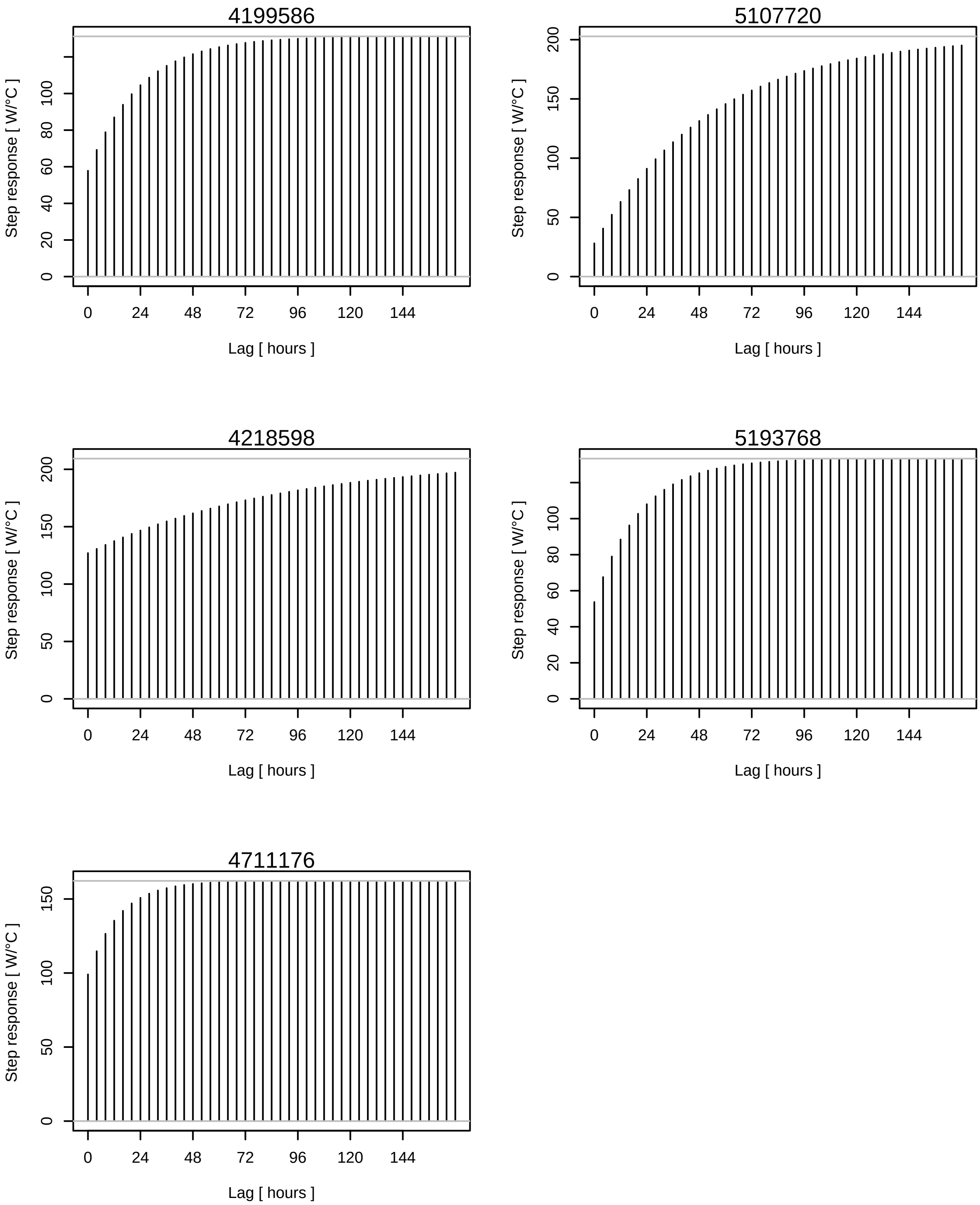

Figure 27: Estimated step response, based on measurements of heat consumption, for model no. 4 (except for 5107720 where model no. 2 is use), when a varying indoor temperature is allowed for. The horizontal lines are placed at zero and at the estimated UA-value.

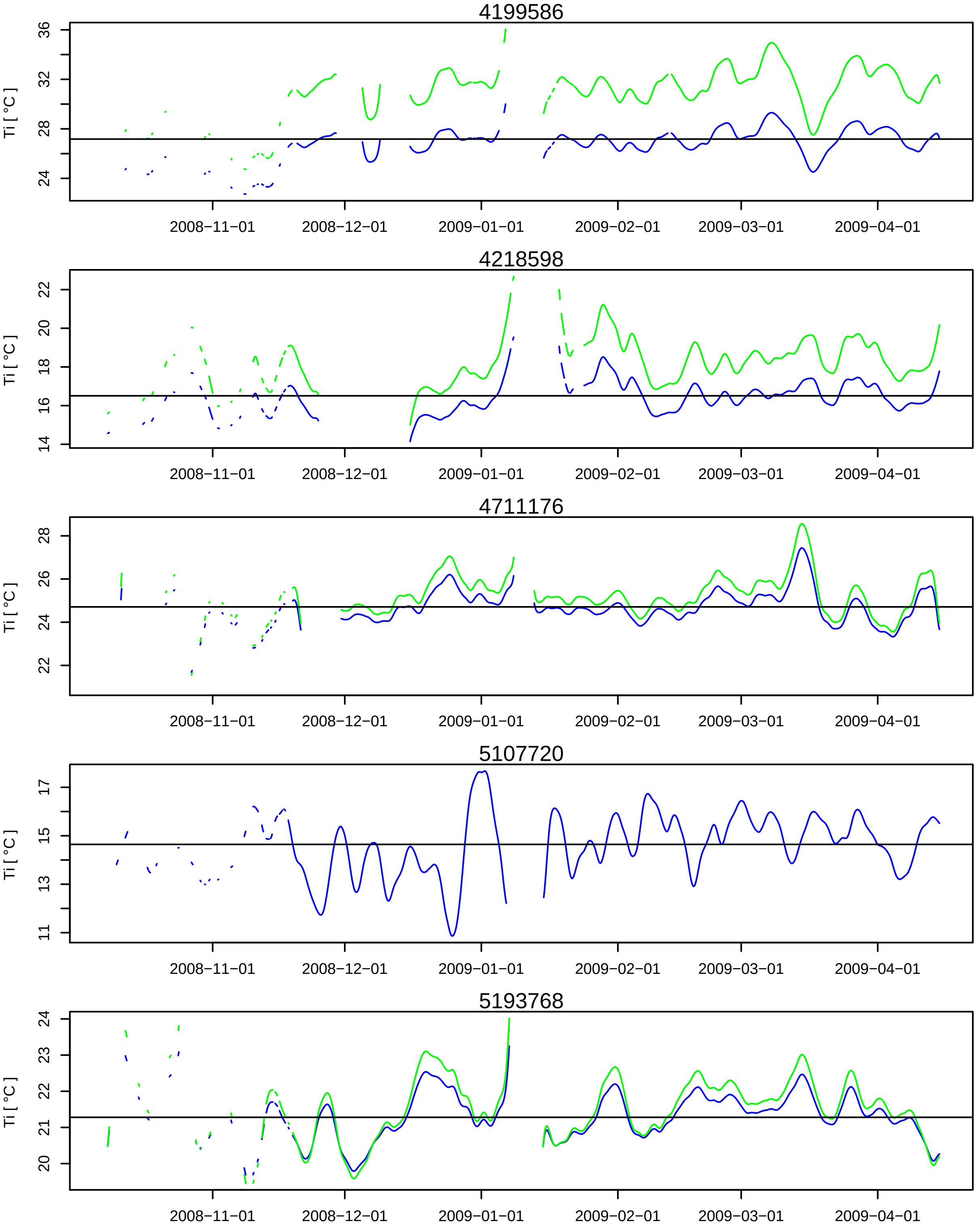

Figure 28: Estimated indoor temperature, based on measurements of heat consumption, for model no. 4 (except for 5107720 where model no. 2 is use). The horisontal line indicates the constant value in the 4h model, the blue line is the value of $T_{i,t}$ *and the green line is the value of* $\hat{a} + \hat{b}T_{i,t}$*, cf. Appendix F, Eq.* (4).

References

[1] ENFOR/03EKS0001A003-A. Estimation of UA-values for single-family houses, 2008.

[2] ENFOR/03EKS0007A001-A. Detailed measurements of heat and electricity consumption in Sønderborg households during the herating season 08/09, 2010.

[3] Øivin Holter, Finn Ingerbretsen, and Hugo Parr. *Fysik og energi ressurser*. Universistetsforlaget, Olso, Norway, 1979.

[4] Henrik Aalborg Nielsen. LFLM version 1.0, an S-PLUS / R library for locally weighted fitting of linear models. Technical Report 22, Department of Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark, 1997.