Modeling and Prediction of Load in District Heating

Final Project for MASM17

André Nüßlein, Anton Palets

Supervisors: Andreas Jakobsson
Department of Mathematics
Division: Mathematical Statistics
January 2021

Abbreviations

ARMAX	Autoregressive–Moving-Average with exogenous input
BJ	Box-Jenkins

Contents

1. Introduction

District heating is a system which distributes centrally produced heat to homes and businesses. This is done by means of a power plant and a distribution network of insulated pipes as sketched in figure 1.1. To model the plant load, we can leverage exogenous inputs such as supply water and ambient air temperatures. The latter will intuitively influence the demand, as people will require more heating when it is cold outside. Similarly, more power will be necessary to provide a warmer water supply.

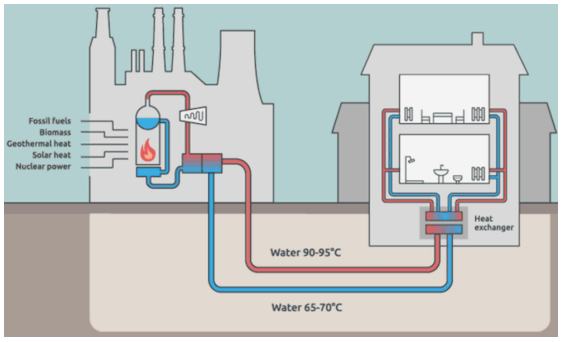


Figure 1.1: A sketch showing how district heating works [1].

As will be discussed and motivated below, air temperature will influence the power load more than water temperature, so we start with a rational transfer function model using the Box-Jenkins methodology, and then use the same approach with both air and water temperatures as inputs. We then continue on to form an ARMAX reformulation of the two-input model and use Kalman filtering to recursively estimate the parameters throughout the 1989 part of the data. Finally, predictions are also done using Facebook's forecasting tool Prophet.

The available 1989 data set from a Danish town was examined for suitable sections of data for modeling, validation, and testing. The power used by its district heating plant in 1989 can be seen in figure 1.2. Eight weeks were chosen as the modelling data set (1989/07/10 to 1989/09/03), with the subsequent two weeks as validation data, and a week following that the first test data. Another test week was chosen in the winter to evaluate model performance in an 'unfamiliar' season (from 1989/11/27 onward).

In the process of creating the models, outlier detection and replacement were considered. Any sample more than 1.5 interquartile ranges over the upper quartile or under the lower quartile was treated as an outlier and cubic interpolation was used to fill its place. The few samples that were found to be outliers did not end up having a meaningful effect on the models, so it was decided to create the models using unprocessed data.

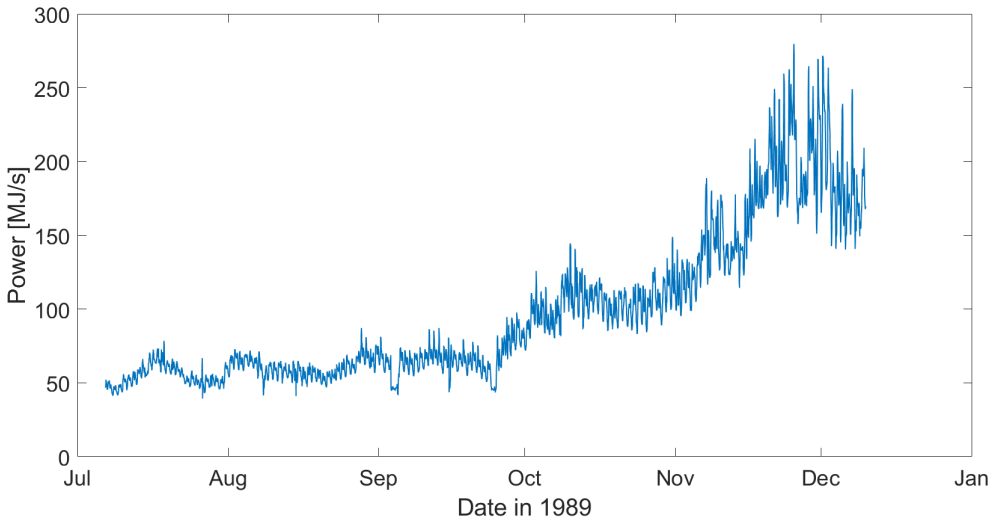


Figure 1.2: The load in the district heating system in Esberg, Denmark, in 1989.

2. Air temperature as exogenous input

To begin with, the load is assumed to be a result of the Box-Jenkins (BJ) model with air temperature as an exogenous input. This model is a slight generalization of the ARMAX model, having the form

$$y_t = \underbrace{\frac{B_1(z)z^{-d_1}}{A_{21}(z)}}_{H(z)} u_t + \underbrace{\frac{C_1(z)}{A_1(z)}}_{\tilde{e}_t} e_t, \tag{2.1}$$

where y_t is the output signal, here the load, u_t is the input signal, here the air temperature, d_1 is the time delay between input and output, and e_t is a white noise. The subscripts are chosen with a second exogenous signal in mind.

All the estimation is done on zero mean data. For prediction, the previously subtracted mean is simply added again in the end.

Since u_t is not a white noise, pre-whitening needs to be performed. This means that a model for the input needs to be found such that the input is driven by white noise. An ARMA model does the job:

$$A_3(z)u_t = C_3(z)u_t^{pw}. \tag{2.2}$$

We find the coefficients to be

$$\begin{aligned} A_3(z) = 1 &- 1.437(\pm 0.01982)z^{-1} + 0.4859(\pm 0.01978)z^{-2} \\ &- 0.4746(\pm 0.02642)z^{-23} + 0.4312(\pm 0.02811)z^{-24} \end{aligned}$$

$$C_3(z) = 1 + 0.07176(\pm 0.02514)z^{-22} - 0.3819(\pm 0.03841)z^{-23} - 0.2133(\pm 0.02717)z^{-24}.$$

The ACF, PACF, and the normal probability plot for the residual from these coefficients are shown in figure 2.1. While the residual is not perfectly white, it is hard to improve on the above mentioned coefficients.

These coefficients can be plugged into equation 2.1 and if this in turn is multiplied by $A_3(z)/C_3(z)$, one gets

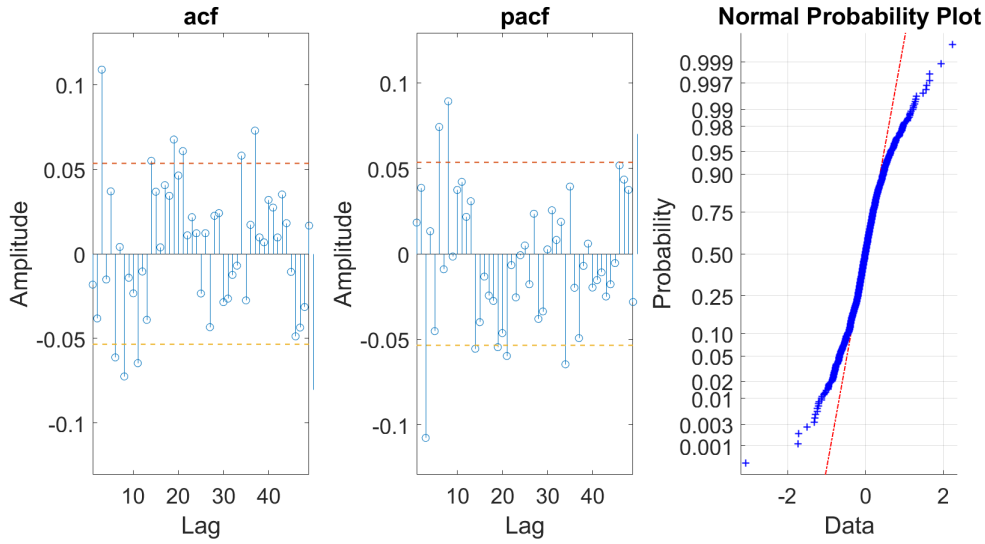


Figure 2.1: ACF, PACF, and normality of the air temperature residual.

$$\frac{A_3(z)}{C_3(z)}y_t = \frac{B_1(z)z^{-d_1}}{A_{21}(z)}u_t^{pw} + \frac{A_3(z)}{C_3(z)}\frac{C_1(z)}{A_1(z)}e_t. \tag{2.3}$$

We call the LHS of equation 2.3 y_t^{pw} and treat the noise summand as to be uncorrelated with u_t. Then the CCF between u^{pw} and y^{pw}, figure 2.2, gives an estimate of the impulse response $H(z)$ (see equation 2.1). Using Table 4.7 in the textbook, we find $(d, r, s) = (0, 0, 0)$ such that initial guesses are $A_{21}(z) = 1$ and $B_1(z) = 1$. Using PEM, we then estimate $\hat{B}_1(z) = -0.9834$ and $\hat{A}_{21}(z) = 1$.

We then find the ARMA model for the residual $\tilde{e}_t$ which gives us the proper model orders of the polynomials $A_1(z)$ and $C_1(z)$. After having obtained all relevant polynomial orders, we reestimate the polynomials jointly for our final model:

$$\begin{aligned}
B_1(z) &= -0.6003(\pm 0.1159) \\
A_{21}(z) &= 1 \\
C_1(z) &= 1 \\
A_1(z) &= 1 - 1.011(\pm 0.02717)z^{-1} + 0.1482(\pm 0.02714)z^{-2} - 0.1968(\pm 0.02629)z^{-23} \\
&\quad - 0.09132(\pm 0.04119)z^{-24} + 0.04829(\pm 0.04139)z^{-25} + 0.1451(\pm 0.02806)z^{-26}
\end{aligned}$$

In figure 2.3 the model fitting residual can be seen to be white noise, as statistically enough correlations are inside the confidence interval and the normality plot shows the residual to be normal enough to interpret these confidence bounds as correct.

This model is then evaluated on the validation set by seeing whether the one-step prediction residual is white noise (figure 2.4) and whether the 8-step prediction residual resembles an

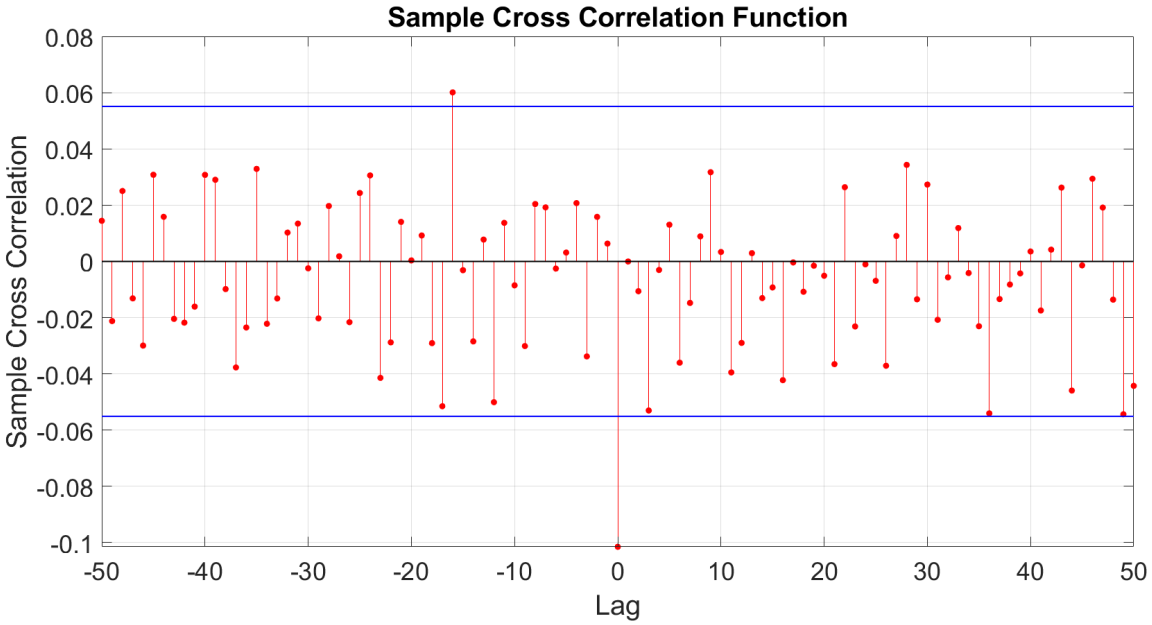


Figure 2.2: The CCF between u^{pw} and y^{pw} giving and estimate of the impulse response.

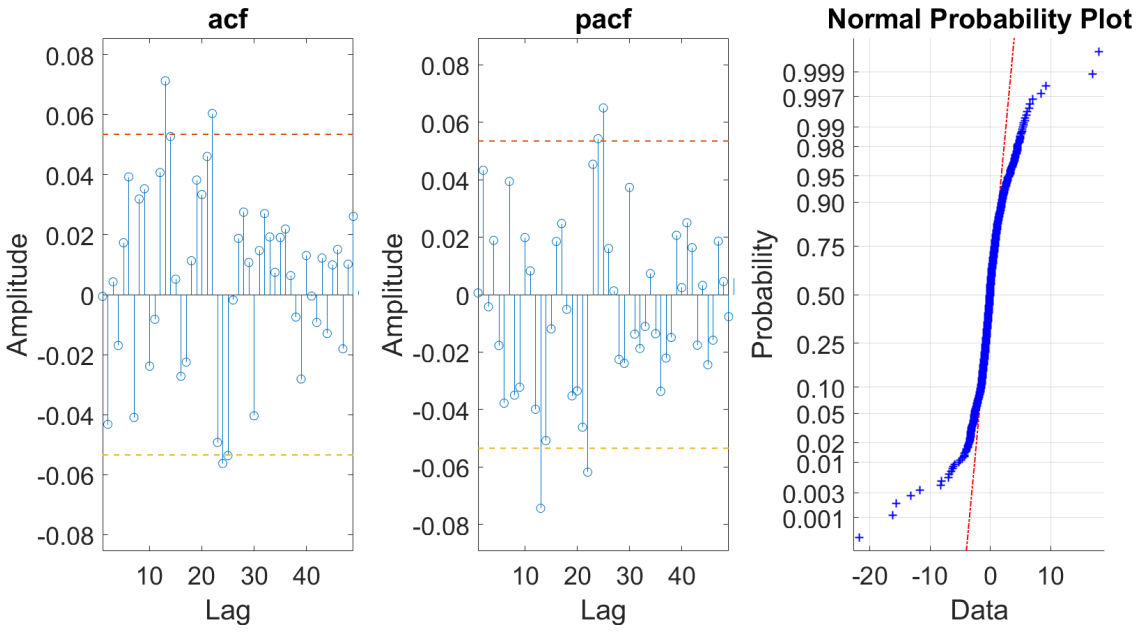
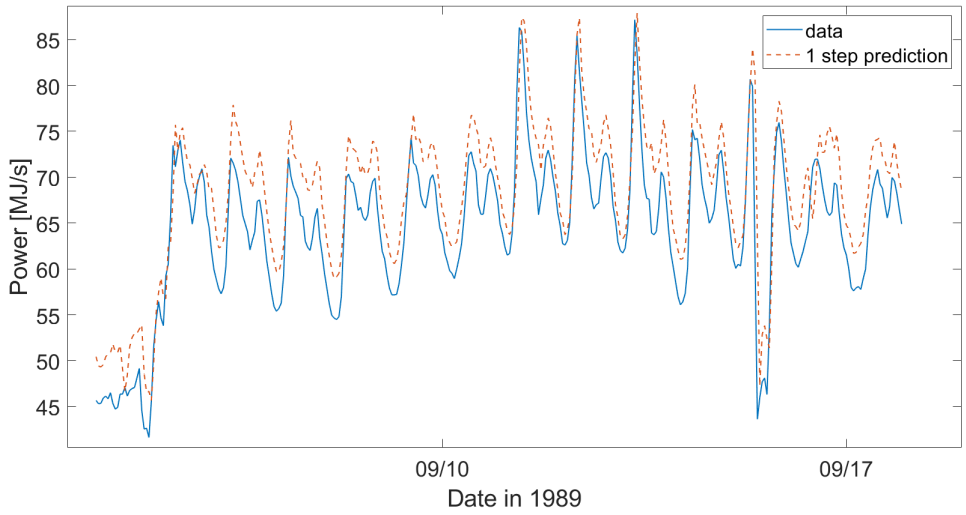


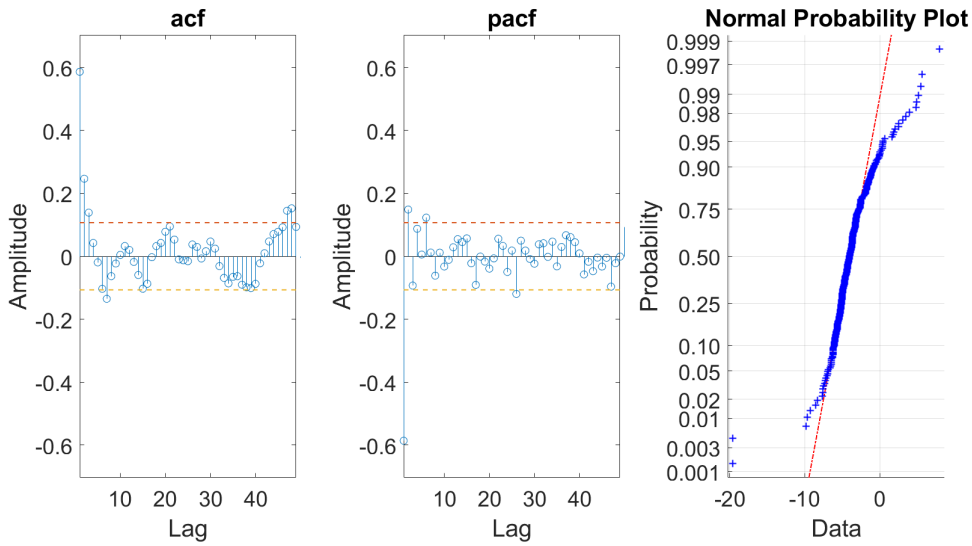
Figure 2.3: ACF, PACF, and normality of single input model fitting residual.

MA(7) process (figure 2.5). The final performance of the model is established based on the test data performance. The predictions and residuals for the summer and winter test data can be seen in figures 2.6 and 2.7 respectively. The 6-step error variances were found to be $\operatorname{Var} \epsilon^{\text{summer}}_{k+6|k} = 27.5881$ and $\operatorname{Var} \epsilon^{\text{winter}}_{k+6|k} = 534.7279$. The performance on the winter data set is significantly worse because the model was not made for that season and the variance of the data is significantly higher during the colder parts of the year.

With such high variance, the predictions cannot be trusted for winter. The performance in the season the model was trained on is not ideal, but it can be used.

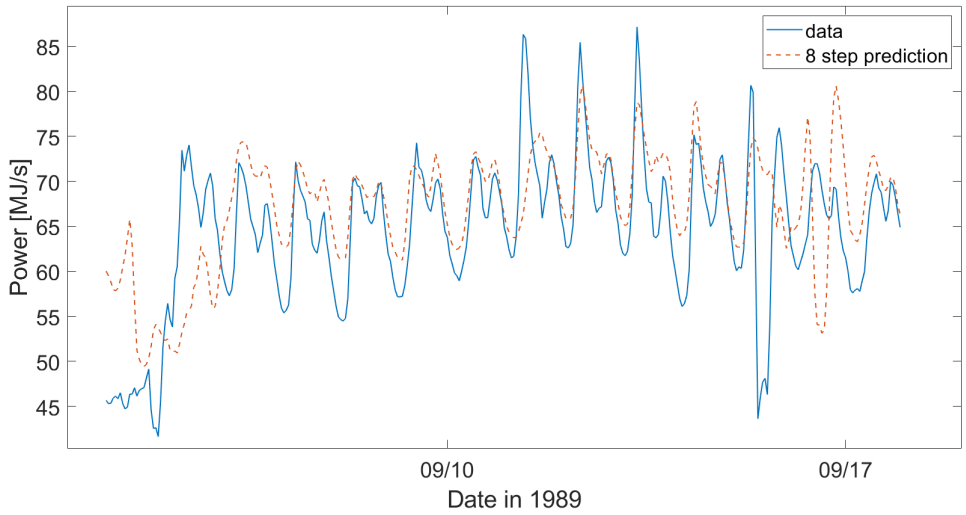


(a) Prediction as compared to the the validation data set.

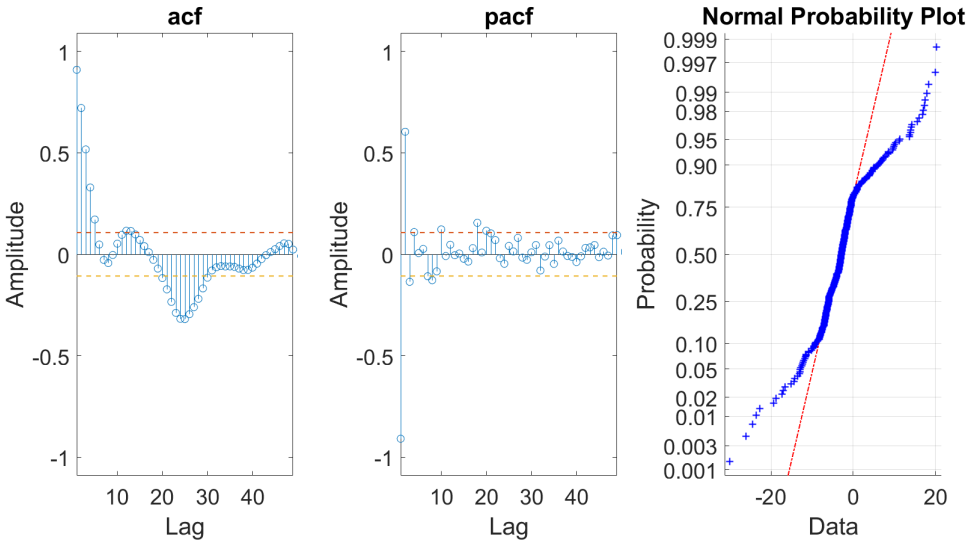


(b) ACF, PACF, and normality of the residual.

Figure 2.4: Single input model 1-step prediction on validation data.

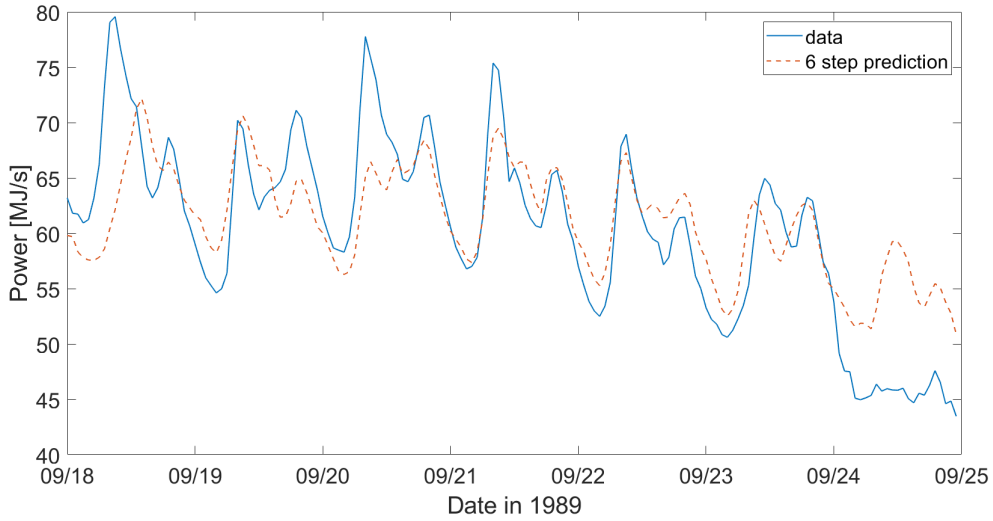


(a) Prediction as compared to the the validation data set.

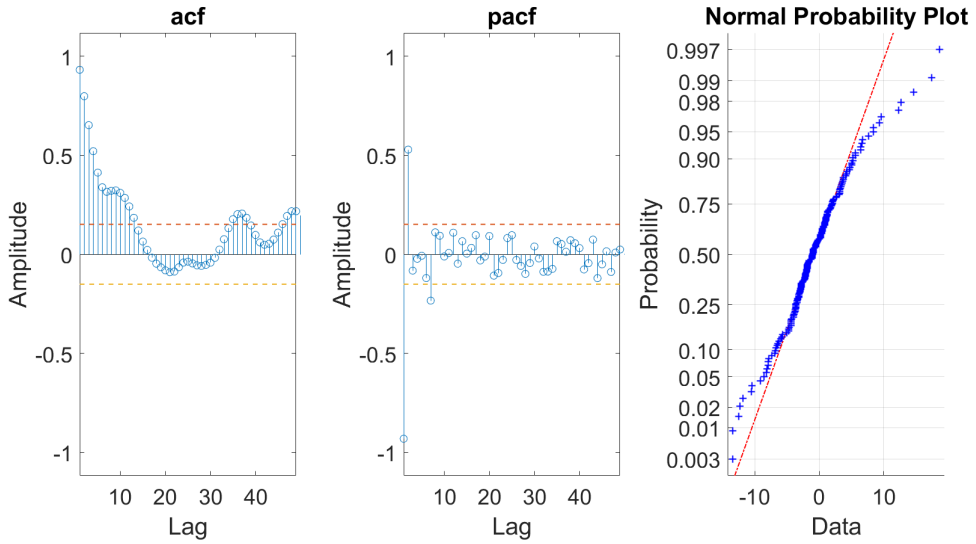


(b) ACF, PACF, and normality of the residual.

Figure 2.5: Single input model 8-step prediction on validation data.

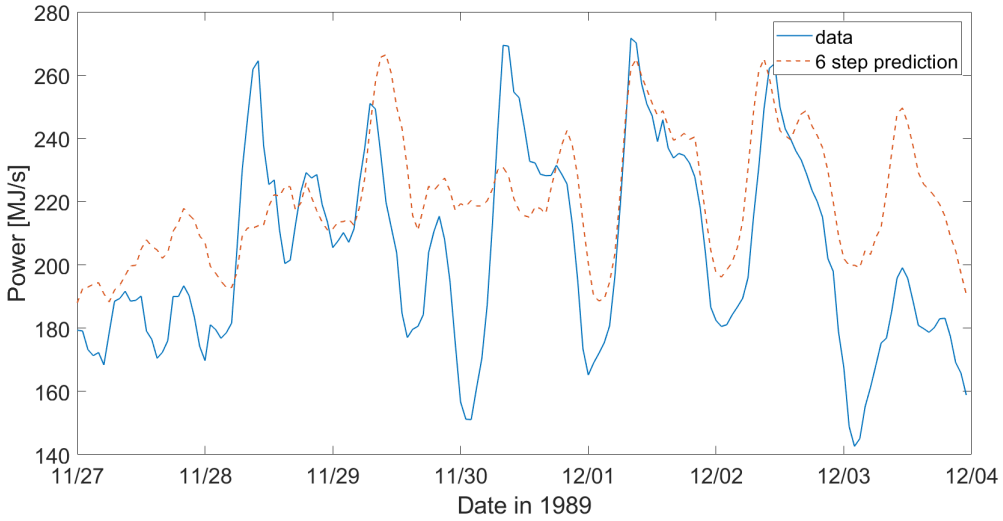


(a) Prediction as compared to the the summer test data set.

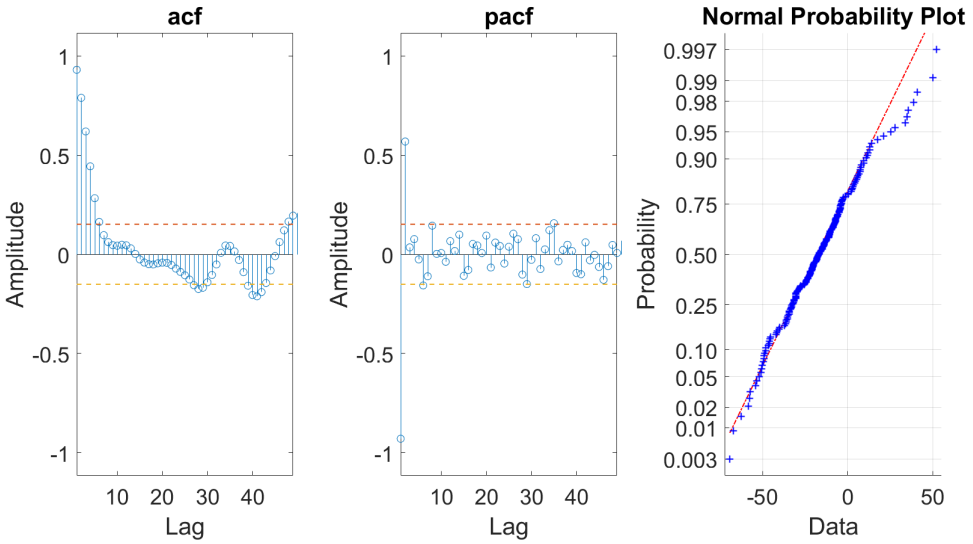


(b) ACF, PACF, and normality of the residual.

Figure 2.6: Single input model 6-step prediction summer test data.



(a) Prediction as compared to the the winter test data set.



(b) ACF, PACF, and normality of the residual.

Figure 2.7: Single input model 6-step prediction on winter test data.

3. Air and water temperatures as inputs

To model the load using air temperature and water temperature as inputs, we create a dual input single output model. We approach the task in a similar manner as in the single input case discussed above, such that the generalized ARMAX model becomes

$$y_t = \frac{B_1(z)z^{-d_1}}{A_{21}(z)}u_t + \frac{B_2(z)z^{-d_2}}{A_{22}(z)}w_t + \frac{C_1(z)}{A_1(z)}e_t, \tag{3.1}$$

where w_t is the water temperature. Now we first model the load with the most correlated of the two inputs and then treat the residual as a single input system. After having obtained all model orders, we reestimate all parameters. From figure 3.1 we clearly see that the outside temperature is more strongly correlated to the load, which is convenient because its model order was already estimated. As a sanity check, we confirm that the outside temperature is negatively correlated and the water temperature positively.

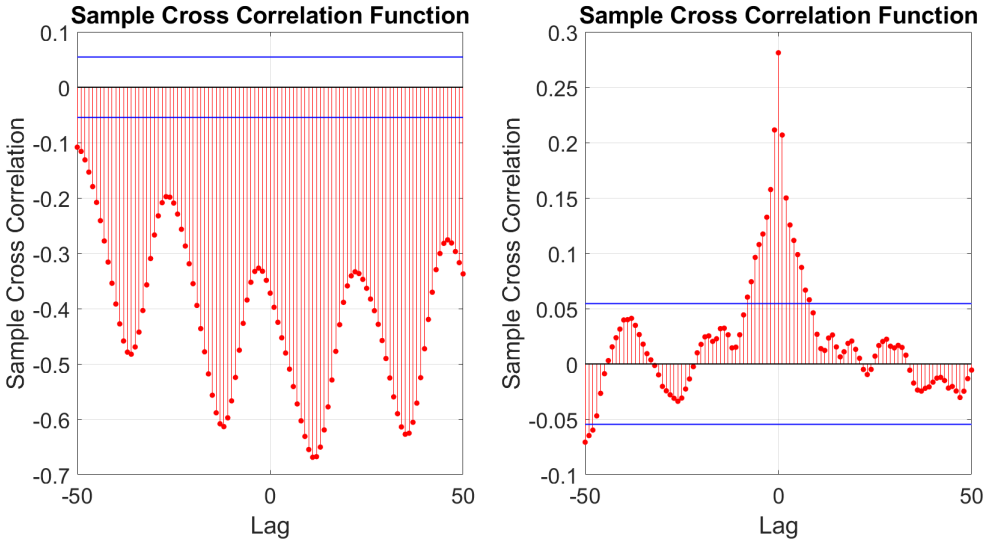


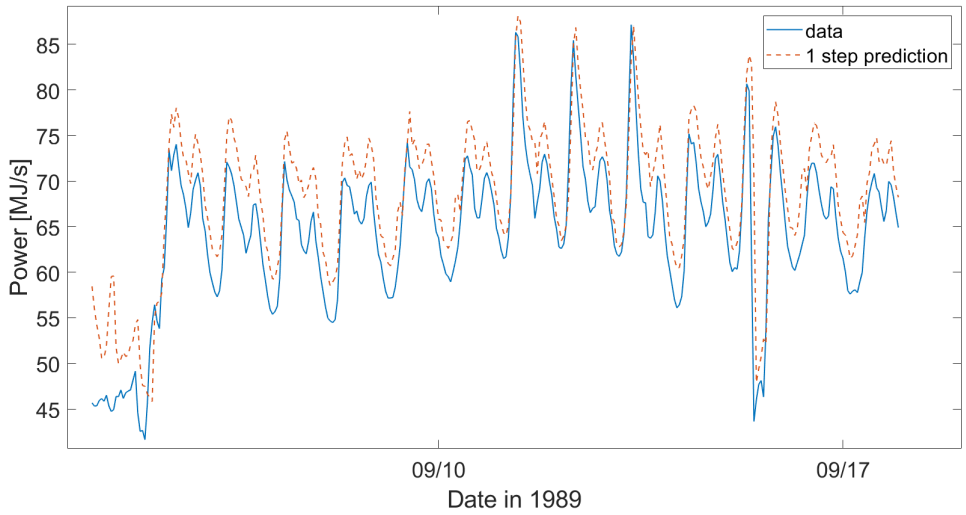
Figure 3.1: *Left:* The cross correlation of ambient air temperature and power load. *Right:* The cross correlation of supply water temperature and power load.

Pre-whitening is carried out as before and model orders are found as described before. The final model is

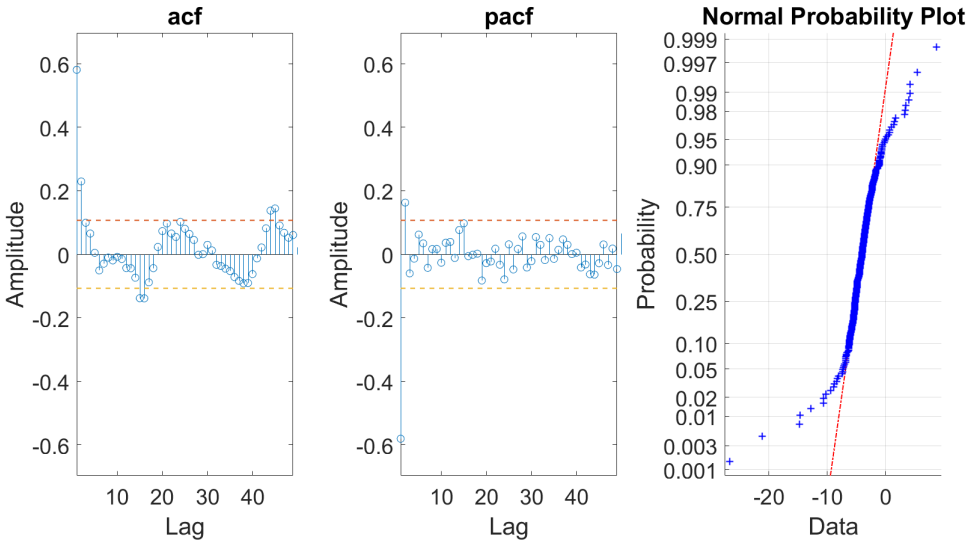
$$
\begin{aligned}
B_1(z) &= -0.1058(\pm 0.08124) \\
A_{21}(z) &= 1 \\
B_2(z) &= 1.928(\pm 0.04144) \\
A_{22}(z) &= 1 \\
C_1(z) &= 1 - 0.8825(\pm 0.0183)z^{-24} \\
A_1(z) &= 1 - 1.166(\pm 0.02799)z^{-1} + 0.1997(\pm 0.02797)z^{-2} - 0.9952(\pm 0.007487)z^{-24} \\
&\quad - 1.151(\pm 0.03076)z^{-25} + -0.1856(\pm 0.02886)z^{-26}
\end{aligned}
$$

The performance of this model can be seen in the figures below. Figures 3.2 and 3.3 show the 1- and 8- step predictions on the validation data. Figures 3.4 and 3.5 show the 6-step predictions on the summer and winter test data sets respectively. The prediction residual variances of the latter two were found to be $\operatorname{Var} \epsilon^{\text{summer}}_{k+6|k} = 23.8518$ and $\operatorname{Var} \epsilon^{\text{winter}}_{k+6|k} = 608.2940$ respectively.

As can be seen in the variance of the residuals, the dual input model performs better in the season its parameters were estimated on than the single input model. Interestingly, the dual input model performs worse in winter than the single input model, which suggests that the single input model was more generalized, albeit worse performing in the season it was estimated on.

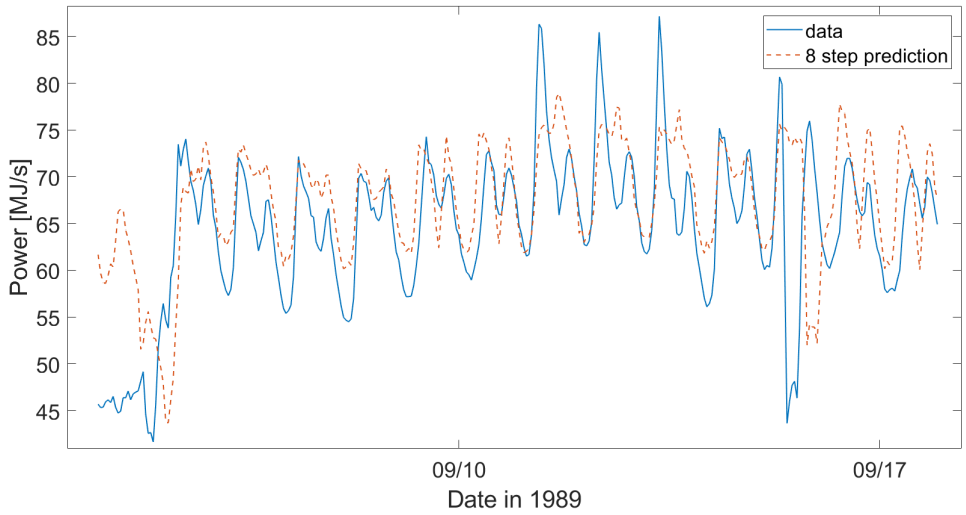


(a) Prediction as compared to the the validation data set.

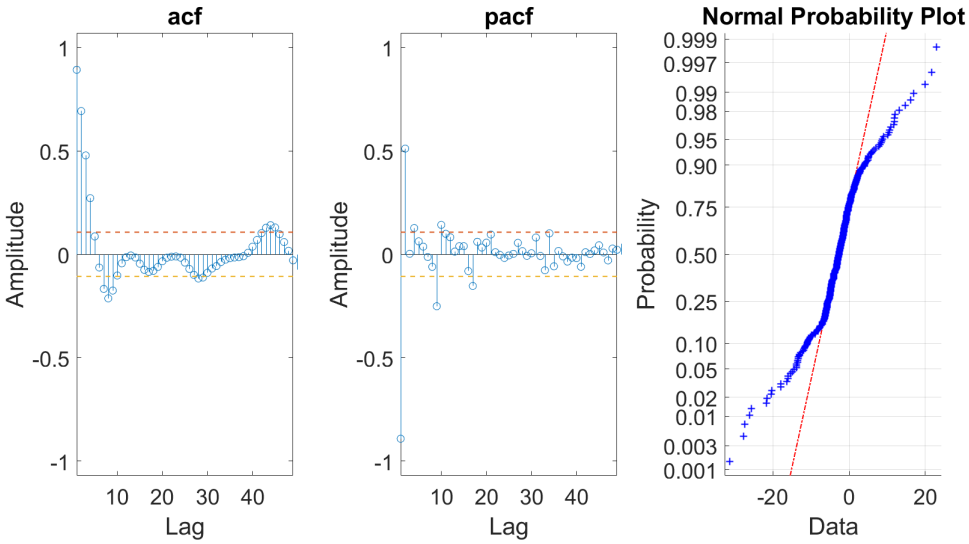


(b) ACF, PACF, and normality of the residual.

Figure 3.2: Dual input model 1-step prediction on validation data.

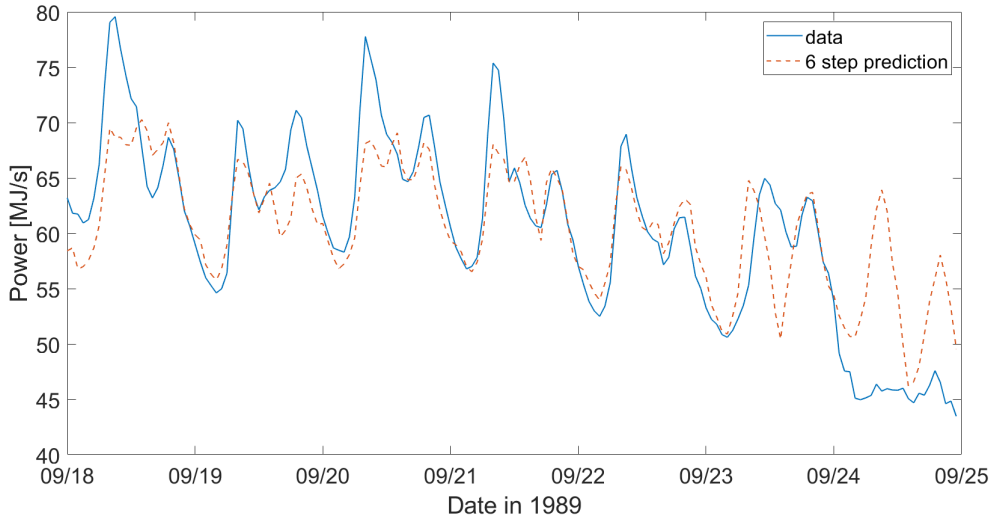


(a) Prediction as compared to the the validation data set.

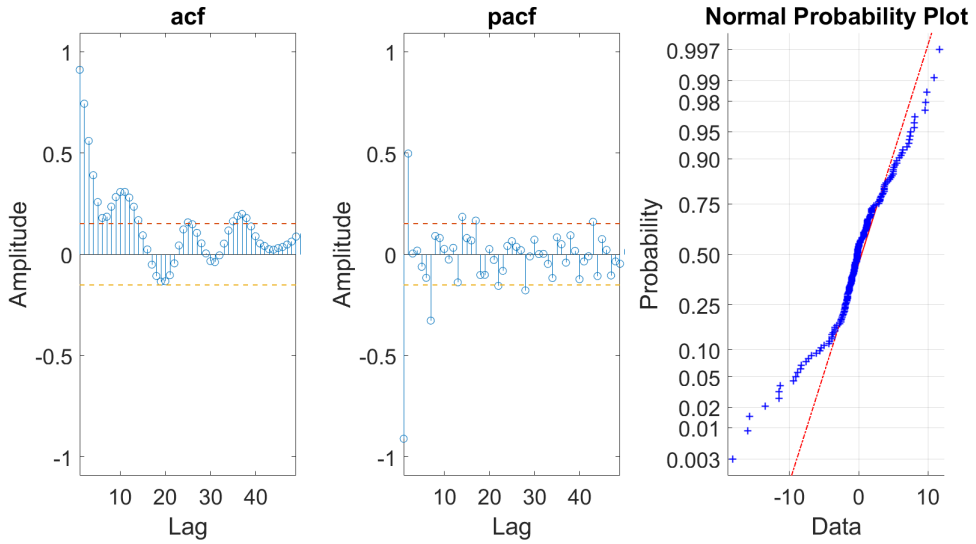


(b) ACF, PACF, and normality of the residual.

Figure 3.3: Dual input model 8-step prediction on validation data.

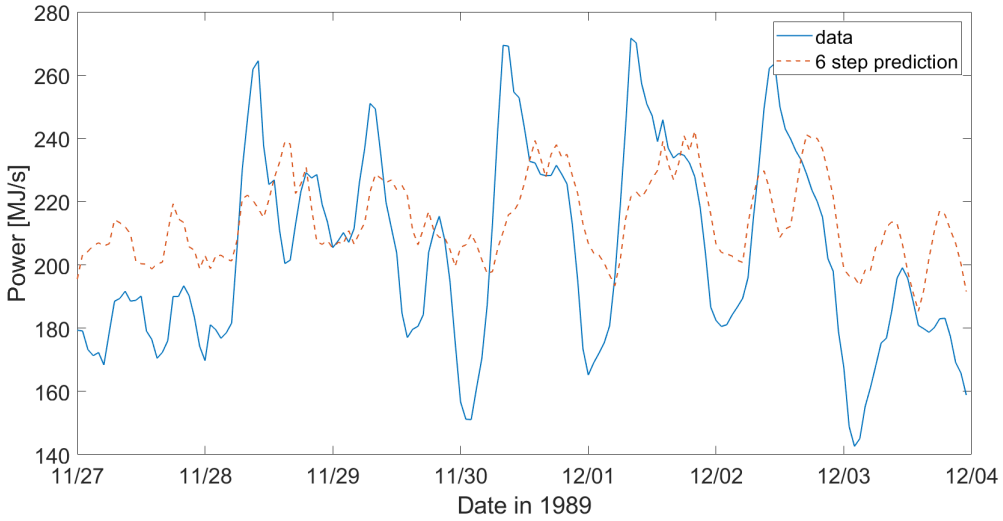


(a) Prediction as compared to the the summer test data set.

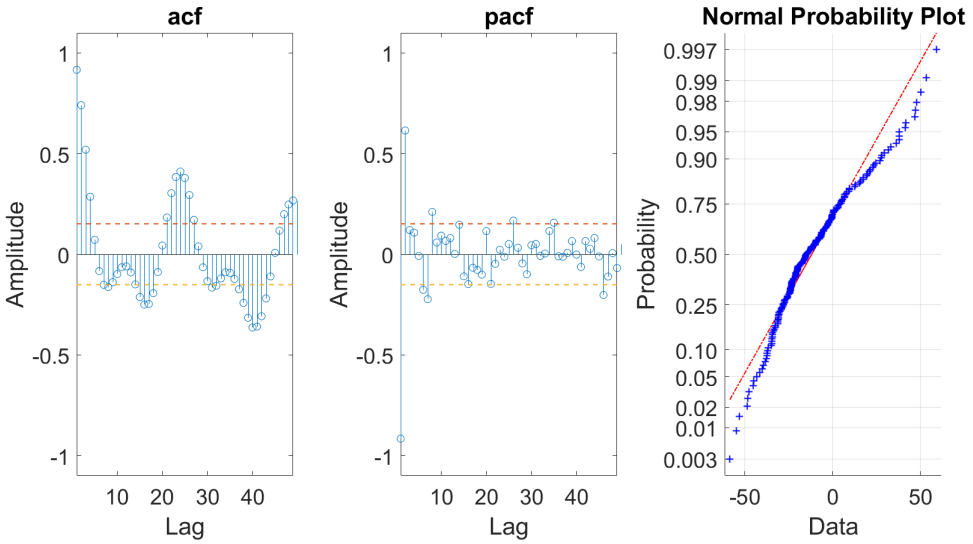


(b) ACF, PACF, and normality of the residual.

Figure 3.4: Dual input model 6-step prediction summer test data.



(a) Prediction as compared to the the winter test data set.



(b) ACF, PACF, and normality of the residual.

Figure 3.5: Dual input model 6-step prediction on winter test data.

4. Recursive parameter estimation

So far, the parameters have been estimated in summer and it was seen that they resulted in poor winter predictions. Using the same model orders as found before, but estimating the parameters in a modelling data set that starts on the second of October 1989 and ends where the winter test data begins, we find $\operatorname{Var} \epsilon^{\text{winter}}_{k+6|k} = 488.3988$ for the 1-input Box-Jenkins model. This is a slight improvement, but it is still high because October and November have notoriously varying data. In the two-input model, we find a worsening to $\operatorname{Var} \epsilon^{\text{winter}}_{k+6|k} = 735.7310$. To further develop on the fact that the parameters are expected to vary over time, we turn to estimating them recursively.

In order to use a Kalman filter, we rewrite equation 3.1:

$$A_1(z)y_t = \underbrace{A_1(z)B_1(z)}_{B_u(z)} u_t + \underbrace{A_1(z)B_2(z)}_{H_w(z)} w_t + C_1(z)e_t, \tag{4.1}$$

where we have used that both $A_{21}(z)$ and $A_{22}(z)$ were previously estimated to be identically 1. Since $A_1(z)$ was previously estimated to have as free parameters a_1, a_2, a_{24}, a_{25} and a_{26}, and that both $B_1(z)$ and $B_2(z)$ are constant, we now have

$$A_1(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + a_{24} z^{-24} + a_{25} z^{-25} + a_{26} z^{-26} \tag{4.2}$$
$$B_u(z) = b_0 + b_1 z^{-1} + b_2 z^{-2} + b_{24} z^{-24} + b_{25} z^{-25} + b_{26} z^{-26} \tag{4.3}$$
$$H_w(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + h_{24} z^{-24} + h_{25} z^{-25} + h_{26} z^{-26} \tag{4.4}$$
$$C_1(z) = 1 + c_{24} z^{-24} \tag{4.5}$$

Given this formulation of the model, we can easily write down the state space representation of the system

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{e}_t \tag{4.6}$$
$$y_t = \mathbf{C}_t \mathbf{x_t} + w_t \tag{4.7}$$

with the hidden state

$$\begin{aligned} \mathbf{x}_t = \big(&b_0 \quad b_1 \quad b_2 \quad b_{24} \quad b_{25} \quad b_{26} \quad h_0 \quad h_1 \quad h_2 \quad h_{24} \\ &h_{25} \quad h_{26} \quad c_{24} \quad a_1 \quad a_2 \quad a_{24} \quad a_{25} \quad a_{26}\big)^T, \end{aligned} \tag{4.8}$$

$\mathbf{A} = \mathbf{I}_{18}$ the identity matrix of appropriate dimension, and

$$\mathbf{C}_t = \big(u_t \quad u_{t-1} \quad u_{t-2} \quad u_{t-24} \quad u_{t-25} \quad u_{t-26} \quad w_t \quad w_{t-1} \quad w_{t-2} \quad w_{t-24} \quad w_{t-25} \quad w_{t-26} \quad \hat{e}_{t-24} \quad -y_{t-1} \quad -y_{t-2} \quad -y_{t-24} \quad -y_{t-25} \quad -y_{t-26}\big), \tag{4.9}$$

with noise values estimated from previous 1-step prediction residuals

$$\hat{e}_{t-k} = \epsilon_{t-k|t-k-1} = y_{t-k} - \mathbf{C}_{t-k|t-k-1}\hat{\mathbf{x}}_{t-k|t-k-1} \tag{4.10}$$

In the implementation of the Kalman filter, we put $\mathbf{R}_w = 8.6415$ and $\mathbf{R}_\mathbf{e} = 10^{-5}\,\mathbf{I}_{18}$. The value of $\mathbf{R}_w$ is the variance of the 1-step prediction error on the validation set using the Box-Jenkins 2-input model presented in the previous chapter. The initial guesses for the hidden state are also taken from there. The k-step prediction procedure follows example 8.12 on page 300 in the course literature [2]. Since we need to initialize the prediction on some values, we let the first 24 e_t values be $N(0,1)$-distributed, and the following are created as in equation 4.10. The 6-step prediction and residual results on the two test data sets can be seen in the figures 4.2 and 4.3 for summer and winter respectively. The residual variances were calculated to be $\operatorname{Var}\epsilon^{\text{summer}}_{k+6|k} = 20.4721$ and $\operatorname{Var}\epsilon^{\text{winter}}_{k+6|k} = 296.6762$ for summer and winter respectively. As was mentioned in the previous sections, these results are to be expected, as the variance of the data itself is significantly higher in the colder months, making it harder for the Kalman filter to find the right parameters.

Figure 4.1 reveals that the coefficients $b_2, b_{26}, h_2, h_{26}, a_2$, and a_{26} are little compared to all others. We therefore create a second ARMAX model that drops these coefficients and estimate the remaining parameters again in a Kalman filter, where we also predict anew. The result is much worse than before, giving $\operatorname{Var}\epsilon^{\text{summer}}_{k+6|k} = 39.3352$ and $\operatorname{Var}\epsilon^{\text{winter}}_{k+6|k} = 579.2934$. We therefore conclude that, even with the Kalman filter, the previously found model orders describe the process well.

5. Automatic prediction using Prophet

Facebook's Prophet works differently from the ARMAX models described above. In particular, it does not does not do a 1- or 6- step prediction but predicts the whole desired future frame at once, and it does not use external inputs. It thus identifies patterns in the "training" data, creates a model, and then predicts as long in the future as specified.

If trained on all data up to September and then predicting the test week from the previous chapter, prophet results in figure 5.1a and a prediction error with variance 60.387. Significantly worse results were obtained for winter, figure 5.1b yielding a prediction error variance of 766.623.

In both cases, Prophet has no problem identifying the periodicity, but it cannot predict the peak height. The previously deployed ARMAX models do a better job because they rely on data that

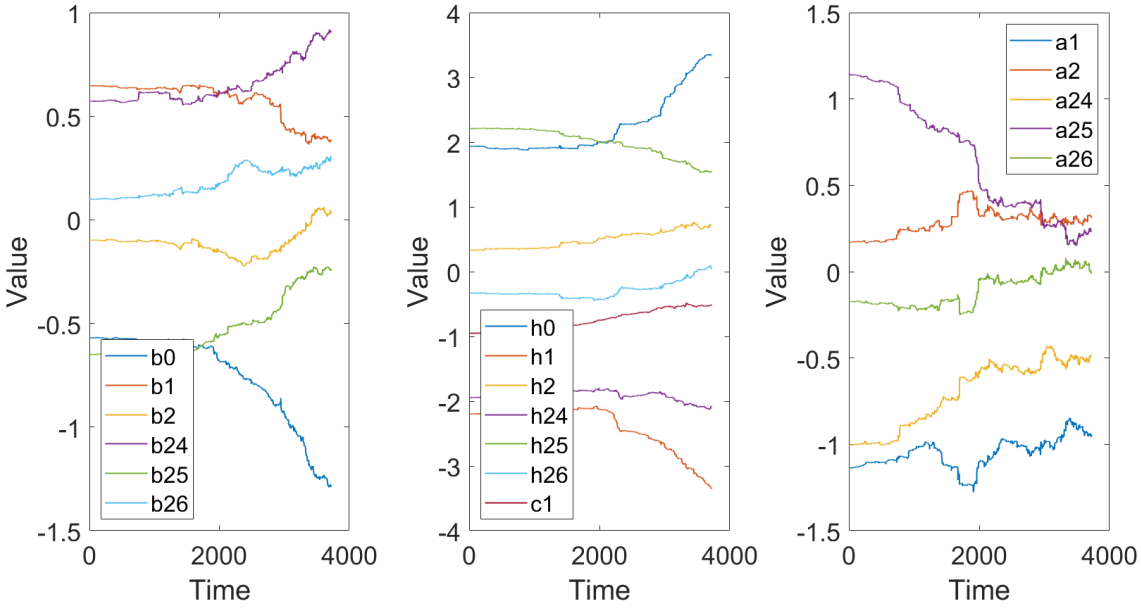
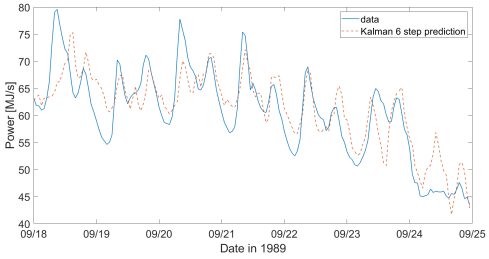
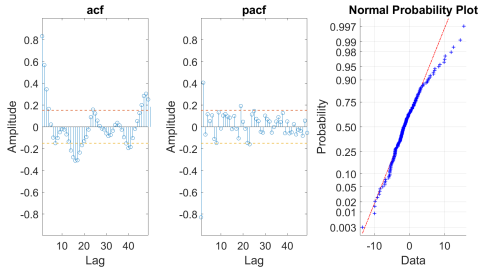


Figure 4.1: Kalman model parameters.

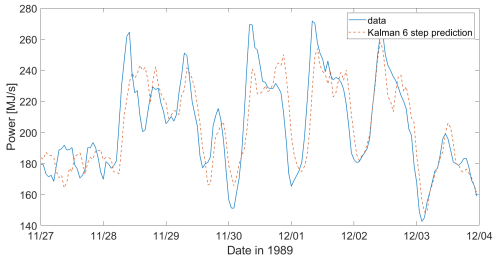


(a) Prediction as compared to the the summer test data set.

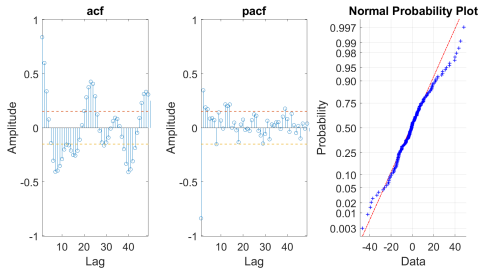


(b) ACF, PACF, and normality of the residual.

Figure 4.2: Kalman filter 6-step prediction on summer test data.



(a) Prediction as compared to the the winter test data set.

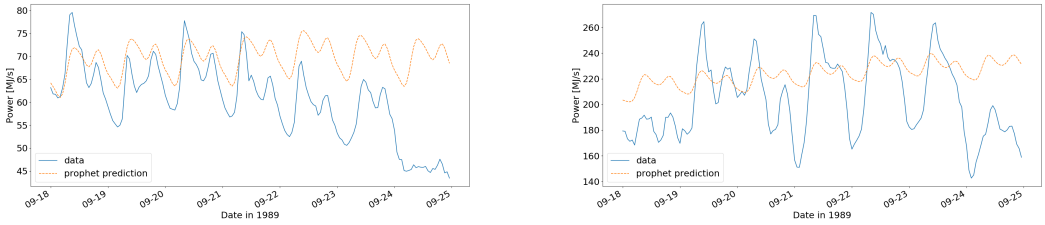


(b) ACF, PACF, and normality of the residual.

Figure 4.3: Kalman filter 6-step prediction on winter test data.

lies only six hours previously and instantaneous input. The fact that Prophet performs worse in winter, just as ARMAX, is to be expected because then the amplitudes vary more greatly.

Prophet's true power would probably be revealed if it were "trained" on more years, but this was outside the project's scope.



(a) prediction in the summer test week.

(b) Prediction in the winter test week.

Figure 5.1: Facebook's Prophet predicting two weeks in two different seasons in 1989.

6. Conclusion

In this project, four models for power load prediction were created. These differed from one another in the information available to them and in the underlying assumptions made. It is reasonable to expect the dual-input Box-Jenkins model to perform better than the single-input model, and the recursive model to be even better in turn, as we expect the parameters to change with the seasons. These expectations are generally reflected in the performance of our models, as can be seen in the 6-step prediction residuals over the two test data sets in table 6.1.

Table 6.1: The variance of the 6-step prediction residuals.

	Summer	Winter
Single input BJ	27.5881	534.7279
Dual input BJ	23.8518	608.2940
Kalman filter	20.4721	296.6762

As can be seen, the exception to this is the winter performance of the dual input model, which hints at the lack of generalization of the more complicated model. Although the winter predictions are generally not great, they at least model the periodicity well. This periodicity was also well modelled by Prophet, but because it predicted so far into the future, it did not reproduce the amplitudes reliably.

In conclusion, the best performing prediction was was found by the Kalman filter, which recursively estimates time-dependent parameters. It did, however, perform well only with the model orders that were estimated in summer. The predictions made by the Kalman filter can be considered good for the summer months, but they deteriorate in winter, presumably due to the higher variance of data.

Bibliography

[1] Laura Toffetti. District heating — Wikipedia, the free encyclopedia, 2016. [Online; accessed 08-January-2021].

[2] A. Jakobsson. *An Introduction to Time Series Modeling*. Studentlitteratur, 2015.