UPPSALA UNIVERSITET

UPTEC STS 20015

Examensarbete 30 hp

Juni 2020

Empirical modeling of the thermal systems in an apartment

A study of the relationship between household electricity consumption and indoor temperature

Jacob Rutfors

Måns Wallentinsson

UPPSALA UNIVERSITET

Teknisk- naturvetenskaplig fakultet
UTH-enheten

Besöksadress:
Ångströmlaboratoriet
Lägerhyddsvägen 1
Hus 4, Plan 0

Postadress:
Box 536
751 21 Uppsala

Telefon:
018 – 471 30 03

Telefax:
018 – 471 30 00

Hemsida:
http://www.teknat.uu.se/student

Abstract

Empirical modeling of the thermal systems in an apartment

Jacob Rutfors, Måns Wallentinsson

In this study, linear and non-linear models were trained on real data to mimic the relationship between household electricity consumption and indoor temperature, in the rooms of an apartment in downtown Stockholm. The aim was to better understand this relationship and to distinguish any divergence between the different rooms. With data from two weeks of measurements, the models proved to perform well when tested on validation data for almost all rooms, only showing performance dips for the middle room. A noticeable correlation between the electricity consumption and the indoor temperature was observed for all rooms except the bedroom. However, the benefits of using this information to predict the indoor temperature are limited and differ between the rooms. The household electricity consumption primarily brought beneficial information to the kitchen models, where most of the heat generating appliances were located. It was found that linear models were sufficient to represent the thermal systems of the rooms, performing equally well and often better than non-linear models.

Handledare: Peter Karlsson
Ämnesgranskare: Per Mattsson
Examinator: Elísabeth Andrésdóttir
ISSN: 1650-8319, UPTEC STS20 015
Tryckt av: Uppsala

Populärvetenskaplig sammanfattning

Automatisk styrning av radiatorer i hushåll är ett väletablerat tillvägagångssätt för att öka bekvämligheten för de boende samtidigt som energianvändningen kan reduceras. Radiatorerna i många av dagens äldre fjärrvärmekopplade byggnader styrs först och främst utifrån utomhustemperaturen, vilket i grova drag betyder att värmeeffekten minskar när utomhustemperaturen ökar, och vice versa. Därigenom kan energi sparas under varma dagar, då radiatorerna inte behöver värma upp hushållet lika mycket. Det finns potential att förbättra den automatiska styrningen för det äldre byggnadsbeståndet och konstruera den på ett sådant sätt att fler variabler tas i beaktning för att reglera hushållets inomhustemperatur. En variabel som potentiellt skulle kunna användas till detta är hushållselförbrukningen, både på grund av att elektriska apparater i hemmet avger värme men också för att en hög förbrukning av el skulle kunna indikera att personer är hemma vilket också är en värmekälla.

Relationen mellan elförbrukningen och inomhustemperaturen har i denna studie grundligt undersökts för en lägenhet i centrala Stockholm. Genom mätningar har det varit möjligt att dels undersöka hur variablerna samvarierar men också huruvida de kan användas för att förutse hur inomhustemperaturen kommer förändras framåt i tiden. För att studera det sistnämnda har modeller skapats för att efterlikna rummens termiska system, det vill säga hur inomhustemperaturen påverkas av faktorer såsom hushållselförbrukningen.

Genom att registrera inomhustemperaturen för kök, sovrum, badrum, vardagsrum och mellanrum separat, och simultant registrera elförbrukningen för hela lägenheten har vi observerat hur hushållselförbrukningen påverkar inomhustemperaturen för respektive nämnt rum. Resultaten visar att en korrelation mellan hushållselförbrukning och inomhustemperatur är noterbar för alla rum, med undantag för sovrummet. I praktiken betyder detta att en ökning i hushållselförbrukningen verkar öka inomhustemperaturen, dock till olika grad och med olika tidsfördröjningar beroende på rum. Variablerna fungerar bra för att förutse framtida förändringar i inomhustemperaturen, dock är inomhustemperaturen för sig själv adekvat för detta ändamål. Det betyder att givet information om en inomhustemperatur vid en viss tidpunkt kan en framtida inomhustemperatur (här 15 minuter framåt) förutses med god noggrannhet. Om även information om hushållselförbrukningen inkluderas blir prediktionerna ofta bättre, men bara marginellt.

Sammanfattningsvis så finns det potential att förbättra radiatorstyrning genom att ta hänsyn till hushållselförbrukningen, allra främst för köket som enskilt rum. Då förbättringarna är begränsade är det svårt att i dagsläget motivera investeringar i att integrera mjukvara i befintliga uppvärmningssystem, men det är möjligt att detta kan vara gynnsamt under vissa förhållanden. Exempelvis är sannolikheten stor att hushållselförbrukningen korrelerar mer med inomhustemperaturen för mindre lägenheter, där exempelvis kök och vardagsrum utgör en gemensam boendeyta och värmeavgivande elektrisk apparatur är mindre utspridd.

Preface

We are grateful to everyone who have made this final thesis possible. Thanks to Uppsala university assistant professor Per Mattsson we have managed to stay on track and manage a complex subject well. A special thanks also goes out to Aktea Energy with Ebba Lindencrona and Peter Karlsson, who have provided crucial knowledge, feedback and equipment throughout this work.

Distribution of work

This work was executed by Jacob Rutfors and Måns Wallentinsson. All sections of the thesis has been written in collaboration but each person was given a set of main responsibilities. Wallentinsson had the main responsibility of the linear modeling, the correlation analysis and the literature study regarding thermal system modeling. Rutfors had the main responsibility of the non-linear modeling, the sensitivity analysis and the collection- and pre-processing of measurement data. For the other parts, the responsibility was divided equally. For example, black-box modeling theory, programming, the choice of method, text revision, the discussion section and the conclusion section was dedicated equal care from both writers.

Each part of the report has in one way or another been edited by both writers, making the distribution of work close to 50/50.

Vocabulary

Estimation data	Data used to estimate a model.
Validation data	Data used to validate a model.
Goodness of fit (GOF)	Expresses how well a 15-step prediction mimics the validation data**,** using the normalized root mean squared error (NRMSE).
SE-fit	A model's ability to predict an output of the validation data, with respect to Sum of squares (SE).
Input-output models	A term describing models using inputs and outputs to predict values of the output.
No-input models	A term describing models without input. Only the previous values of the output are used to predict future outputs.
Tuned one-step predictor	A model designed to predict outputs one step into the future, using inputs and outputs up to $t-1$ to predict the output at time t.
Tuned 15-step predictor	A model designed to predict outputs 15 steps into the future, using inputs and outputs up to $t-15$ to predict the output at time t.
15-step predictions	1. When done by tuned one-step input-output models, inputs up to $t-1$ and outputs up to $t-15$ are used to predict the output at time t. 2. When done by tuned input-output 15-step predictors, inputs and outputs up to $t-15$ are used to predict the output at time t. 3. When done by no-input tuned one-step predictors or the no-input tuned 15-step predictors, outputs up to $t-15$ are used to predict the output at time t.

Number of units	Number of units used in a Sigmoid network.
Household electricity consumption	Refers to household electricity power consumption in this study.

Table of contents

1. Introduction 9
1.1 Background 9
1.2 Purpose 10
1.3 Questions at issue 10
1.4 Delimitations 10
2. Theory 11
2.1 The utility of modeling 11
2.1.1 Different modeling approaches 12
2.2 System analysis 13
2.2.1 Correlation analysis 13
2.2.2 Transient analysis 16
2.2.3 Frequency- and spectral analysis 16
2.3 Black-box modeling in detail 18
2.3.1 Empirical modeling 18
2.3.2 Linear black-box models 18
2.3.3 Non-linear black-box models 22
2.3.4 Objective functions and model validation 24
2.4 Modeling a thermal system 26
2.4.1 Electricity consumption and indoor temperature 26
2.4.2 Thermal zones 27
2.4.3 Sampling time and thermal time constant 28
3. Method 29
3.1 Equipment 29
3.2 Software 29
3.3 Experimental design and data collection 30
3.3.1 Overview of the experimental design 30
3.3.2 Room specifics 30
3.3.3 Evaluation process 33
3.3.4 Data collection 34
3.4 Modeling of the apartment 34
3.4.1 Correlation analysis approach 35
3.4.2 Two-week modeling approach 36
3.4.3 Two-week linear modeling 36
3.4.4 Two-week non-linear modeling 37
3.4.5 Sensitivity analysis 38
4. Data 40

4.1 Data description and pre-processing 40
5. Results 42
5.1 Correlation analysis 42
5.1.1 Kitchen 42
5.1.2 Bathroom 44
5.1.3 Bedroom 45
5.1.4 Livingroom 47
5.1.5 Middle room 48
5.2 Summary- Correlation analysis 50
5.3 Two-week data- Linear modeling 51
5.3.1 Kitchen 51
5.3.2 Bathroom 52
5.3.3 Bedroom 55
5.3.4 Livingroom 57
5.3.5 Middle room 58
5.4 Two-week data- Non-linear modeling 60
5.4.1 NLARX-and NLAR-models 60
5.4.2 15-step LAVA-predictors 61
5.5 Sensitivity analysis 62
5.6 Summary- Two-week study 63
6. Discussion 66
6.1 Correlation analysis 66
6.2 Two-week study 67
7. Conclusion 72
8. Future studies 73
References 75
Appendix A- Three-day data codes 81
A.1. Correlation analysis 81
Appendix B- Two-week data codes 83
B.1. Best linear tuned one-step models with respect to GOF 83
B.2. Linear tuned one-step- and 15-step predictors 85
B.3. NLAR-and NLARX-models 87
B.4. LAVA AR-predictors 88
B.5. LAVA ARX-predictors 89
Appendix C- Sensitivity analysis codes 91
C.1. Kitchen analysis. 91

1. Introduction

This introduction aims to clarify the general motivation behind this work. Firstly, a background is given, specifying the driving forces of this work. Secondly, the purpose is stated, which will be the common thread throughout the report. Thirdly, the questions at issue are established. Finally, we present the set of delimitations which keep the study manageable.

1.1 Background

Energy efficiency will remain a relevant subject for years to come but it will also entail challenges. In Swedish households about 60 percent of the total energy usage can be assigned to heating (Naturskyddsföreningen, 2016). In relation to this, the Swedish government has an expressed goal of achieving a 50 percent more energy effective society between 2020 and 2030 (Regeringskansliet, 2016). In order to achieve this, implementations of new solutions targeted towards the energy sector must be considered and executed.

The strive for energy efficiency has continued to spur innovation even as of 2020. In Sweden a decrease in energy supply for the upcoming 30 years seems plausible given today's forecasts (Energimyndigheten, 2019). This would lead to an increased demand for smart solutions minimizing unnecessary energy waste in households, both in terms of electricity consumption as well as heating, while maintaining a habitable indoor environment. Aktea Energy is a consultancy company in the forefront of the energy sector which aims to supply solutions encouraging reduced energy use and good indoor environments. Aktea Energy has therefore, in consultation with undersigned, put interest in finding new ways to control indoor temperature by considering the possible correlation between electricity usage and indoor temperature in apartments.

Radiator outputs in older district heated households are today mainly controlled with respect to outdoor temperature, not considering possible heat contributions from components such as ovens and tumblers or human body heat. This leaves room for improvements and possible reductions in energy usage as the radiator output may be reduced during heavy use of electricity and/or when people are present. By constructing *empirical models* of an apartment situated in downtown Stockholm we aimed to better understand the relation between household electricity consumption and indoor temperature, and evaluated the benefits of using the former to predict the latter. The models were assessed by their ability to predict the indoor temperature 15 minutes in advance and whether knowledge about electricity usage added predictive performance. The apartment in question consists of six rooms divided on 85 square meters.

The work was executed in collaboration with, and on behalf of, Aktea Energy. It was supervised by Gothenburg based knowledge manager Peter Karlsson and subject reviewed by assistant professor Per Mattsson at the department of Information technology at Uppsala University.

1.2 Purpose

The purpose and goal of this work was to construct thermal system models based on empirical data from the rooms of a district heated apartment in downtown Stockholm to deduce whether household electricity consumption can help predict indoor temperatures.

1.3 Questions at issue

- Can information about household electricity consumption support predictions of indoor temperature in the rooms of the studied apartment?
- Does the possible correlation between household electricity consumption and indoor temperature differ between the rooms of the apartment?
- Are there any improvements in prediction performance using non-linear models compared to linear models to represent the thermal systems of the rooms?
- Are there any improvements in prediction performance by using information about household electricity consumption up to time $t-1$ compared to $t-15$ when predicting indoor temperature at time t, i.e. do valuable information exists in the last 15 minutes before prediction?

1.4 Delimitations

In the realm of empirical modeling there are a multitude of approaches which can be taken. Here, we aim not to find the most complex models possible but rather the simplest which still yield satisfactory results. It is neither in our interest to investigate every single model structure possible, the focus will rather be on producing a few high-quality models and through these evaluate the possible benefits of knowing the household electricity consumption when predicting indoor temperature.

If a software is to be constructed in practice, which utilizes household electricity consumption data, physical hardware must be engineered and attached to the present radiator infrastructure. This is however not in the scope of this work and will not be considered further.

Finally, we do not intend to reinvent the wheel during this six month work. Many software solutions are already available and will be utilized, the aim is instead to understand these solutions, implement, and modify them to suit our needs.

2. Theory

Within this section all necessary theory will be reviewed. Firstly, models and their practical use are explained. Also, the basics of system analysis is presented and a brief explanation of common methods to analyze systems is given. Secondly, a detailed explanation of the main approach for this work is made, namely black-box modeling. Thirdly, the process of model validation is explicated and connected to previously presented theory. Finally, established strategies and important aspects of thermal system modeling are presented.

2.1 The utility of modeling

By definition, a model can be described as a representation of some real world object or system (Merriam-Webster, 2019). This model may be used as a tool to answer questions about the system without conducting experiments on it (Ljung, Glad, 2003). The model can be constructed in several ways but the main structure to be considered here is the empirical model, which is a mathematical system representation based on data. Data is preferably collected from the system it seeks to represent. The model can be used to reflect trends in the data and support predictions (Hernández-Molinar et al., 2016). For instance, the relationship between two variables like household electricity consumption and indoor temperature can be described by an empirical model.

Models do in many cases hold significant value, if applied well and correctly. Predictive models have for example been used for weather forecasting and health outcomes of disease epidemics, making it easier to design warning systems related to these events (Rogers, 2012). It is however important to emphasize that the quality, and thereby utility, of the model is highly dependent on decisions made by the modeler. Modeling is a technical process which relies on formal theory but it also requires common sense. In general, the modeling process consists of three phases (Figure 1) (Ljung, Glad, 2003):

1) **Problem structuring**: This task involves attaining a better grasp of the problem of interest. When dealing with larger systems, a good practice is to divide it into smaller subsystems. For example, if data is to be registered in an apartment, a good idea may be to register data from each separate room and handle each room as a subsystem. It is also important to identify the relevant variables and how these affect one another. This requires some portion of common sense and intuition from the modeler.
2) **Formulation of base equations:** Here the subsystems in question are studied and the relationships between variables are determined. In a physical system, it is relevant to relate proper laws of nature to the behavior of the variables and form the mathematical equations describing these. Most often a portion of idealization is in place to avoid overcomplicating things. Again, the modeler is expected to decide on a reasonable practice.
3) **Model construction:** In this final phase, the aforementioned equations are structured into a model formulation suitable for analysis or simulation, which finally preferably leads to an elegant state description (Ljung, Glad, 2003).

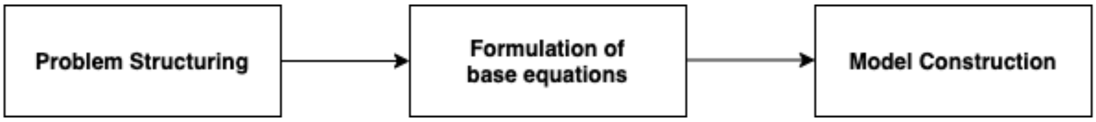

Figure 1. Three-phase method for modeling (Ljung, Glad, 2003).

The three-phase method gives a general structure to assume but may differ depending on the modeling approach chosen by the modeler. In this work, black-box modeling is the method of choice, therefore shorter descriptions of alternative methods are presented but not looked upon in detail.

2.1.1 Different modeling approaches

A researcher is often faced with problems which can be solved in several ways. In the field of modeling there are mainly three approaches necessary to consider: *White-box modeling*, *Grey-box modeling* and *Black-box modeling* (Figure 2). Each approach has its given set of perks and flaws, and one approach will likely yield different results than another. In general, the three modeling approaches mainly differ in terms of accuracy and interpretability, this trade-off is explained further below (Duun-Henriksen et al, 2013).

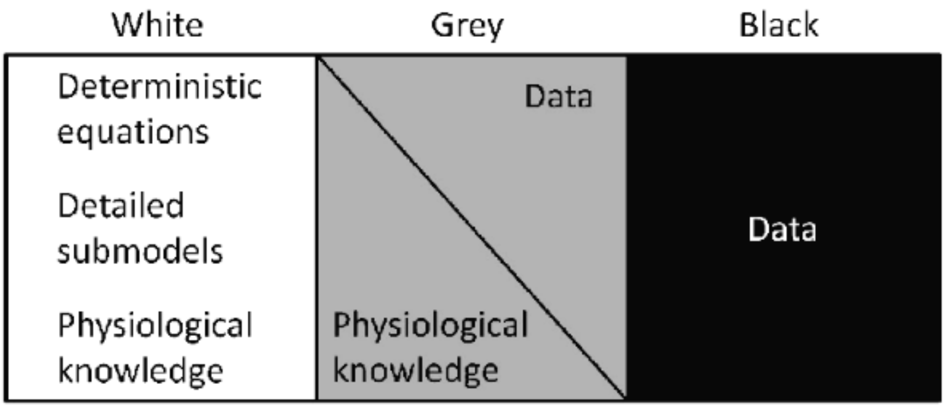

Figure 2. Different modeling approaches (Duun-Henriksen et al, 2013).

A white-box model is primarily based on physiological knowledge about the system it depicts, dealing with deterministic relations and extensive submodels (Duun-Henriksen et al, 2013). A pure white-box model can be interpreted as a copy of reality, this is however not possible in practice.

A grey-box model can be described as an intermediate to a black-box- and a white-box model, i.e. it is based on both data and physiological knowledge about the system (Duun-Henriksen et al, 2013).

A black-box model is more or less entirely based on data. The inner workings of the model can often be hard to interpret but the results can instead prove more accurate than white- or grey-box models (Hulstaert, 2019). An example of a black-box model is the neural network which is given an in depth look in Section 2.3.3. As the goal of this work is to model thermal systems with complex physical properties, black-box modeling is the method of choice.

2.2 System analysis

The primary goal of a system analysis is often to study how a chosen set of variables behaves and covariate. For this reason, it is generally necessary to conduct a *correlation analysis* to find correlated variables and the strength of dependence between them. Therefore, a great part of the theory section will be dedicated to theory about this process. This is followed by a shorter description of the *transient analysis* and the *frequency and spectral analysis* which are methods for finding certain system characteristics but also serves as validation tools.

2.2.1 Correlation analysis

A measure for detecting linear dependency between random variables is the *covariance*. For two random variables (X, Y) the covariance can be expressed in terms of the expected value of the differences between the variables and their respective means μ_X and μ_Y (Alm, Britton, 2008),

$$cov(X, Y) = E((X - \mu_X)(Y - \mu_Y)). \quad (1)$$

It is also possible to express an estimate of covariance from empirical data (Wolfram, 2020). This gives

$$\widehat{cov}(X, Y) = \frac{1}{n}\sum_{i=1}^{n=1}(x_i - \mu_X)(y_i - \mu_Y), \quad (2)$$

where the estimated covariance describes the behavior of the variables and if they imitate each other. If $cov(X, Y) > 0$, the realizations of the random variables tend to be larger or smaller than their means simultaneously. For $cov(X, Y) < 0$ the random variables tend to behave opposite to each other, i.e. when one variable realization is larger than its mean value the other tend to be smaller than its corresponding mean value (Alm, Britton, 2008).

A *correlation* exists between two random variables if a change in one results in a change in the other (Schneider, 2009). Pearson's correlation coefficient p measures the linear dependency between random variables and it can be expressed in terms of the covariance of two variables *(X, Y)* and their respective standard deviations σ_X and σ_Y (Oja et al, 2016), i.e.

$$p(X, Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y}. \quad (3)$$

The correlation coefficient p is dimensionless and can be expressed as a scaled covariance (Rychlik, Rydén, 2006). The value of p ranges from -1 to 1, where $p = 0$ indicates no linear correlation between the variables (Oja et al, 2016). Using the command `corrcoeff` in MATLAB, the correlation coefficient matrix of two random variables can be computed. This matrix contains Pearson's correlation coefficients p of each pairwise combination of variables (Mathworks, 2020a). When using a set of two random variables, *X* and *Y*, the correlation coefficient matrix is of the form $2x2$, presented as

$$corrcoeff(X,Y) = \begin{pmatrix} p(X,X) & p(X,Y) \\ p(Y,X) & p(Y,Y) \end{pmatrix}. \quad (4)$$

Since a variable always show maximum linear dependency when compared to itself, the diagonal of the matrix will be 1 (Mathworks, 2020a).

A realization of a random variable is called an observation and multiple observations forms a time-series. It is possible to calculate an alternative correlation coefficient ρ based on the rankings of two time-series'. Spearman's rank correlation coefficient is calculated by ranking the observations in the arrays $[x_1 \dots x_n]$ and $[y_1 \dots y_n]$ of random variables *X* and *Y* from 1 to n and then summarize over the differences on quadratic form (Alm, Britton, 2008). This gives

$$\rho(X,Y) = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \quad (5)$$

which can be used if no doublets exist. The method measures the level of order association, in contrast to Pearson's correlation coefficient which measures the linear relationship between *X* and *Y*. In equation (5), d_i represents the difference between observations x_i and y_i from the rankings of arrays **x** and **y** (Alm, Britton, 2008).

Using the command `corr` in MATLAB, which utilizes Spearman's method, the rank correlation coefficient is calculated for variables *X* and *Y*. This is done by applying Pearson's linear correlation coefficient to the rankings of *X* and *Y* or, if no doublets exist, by equation (5) (Mathworks, 2020b). The rank correlation coefficient adopts a value between -1 to 1 (Alm, Britton, 2008). A correlation of -1 or 1 represents the maximum degree of negative or positive linear relationship between the ranks of *X* and *Y*.

It is also possible to test the null hypothesis H_0, i.e. whether *X* and *Y* are uncorrelated. For large n, a rule of thumb is to approximate the rank correlation coefficient as normally distributed and p-values can thereby be calculated (Alm, Britton, 2008). The p-value ranges from 0 to 1, where values close to 0 indicate a non-zero correlation. Through the command `corr` it is possible to calculate a p-value which rejects the null hypothesis if it is smaller than 0.05, i.e. smaller than the significance level for a confidence interval of 95 percent (Mathworks, 2020b). The hypothesis depends on the level of confidence, which is selected by the modeler.

It should be noted that there are limitations in only studying the correlation coefficients of Pearson's or Spearman's rank to understand the relationship between two random variables. Schober et.al (2018) explains that a common misconception is that values of correlation coefficients close to 0 entails that no relation exists between the variables. In fact, the correlation describes linear or monotonic association but do not regard other types of relationships. For instance, non-linear relationships can easily be overlooked by these methods. It is therefore important to also adapt other methods to fully understand the behavior of the variables and their impact on each other. One way is to plot the variables as functions of time and observe the visual relationships in the graphs. It is also useful to utilize a scatter plot to analyze the dependency between the variables. In this graph the variables are plotted as functions of each other and the relationship can thereby be visualized. This approach does not consider any time delay between the variables but can show tendencies of non-linear relationships (Alm, Britton, 2008).

Another way to understand the relationship between two variables, while simultaneously regarding potential time delays, is to create and compare prediction models. For example, if information about one variable X improves predictions of another variable Y, compared to predictions based only on knowledge about Y, a correlation between the variables may exist.

A well-established method of deducing the degree of correlation between two time-series' for different time lags (positive delay) and leads (negative delay) is the *cross correlation*. The correlation is specified with a number between -1 and 1 to symbolize either negative or positive correlation between time-series' x and y. The correlation is calculated using different lags or leads d, shifting the series some defined number of steps in positive or negative direction respectively (Bourke, 1996). This is presented mathematically as

$$r(d) = \frac{\sum_{i=1}^{n}[(x_i-\mu_x)(y_{i-d}-\mu_y)]}{\sqrt{\sum_{i=1}^{n}(x_i-\mu_x)^2} * \sqrt{\sum_{i=1}^{n}(y_{i-d}-\mu_y)^2}}, \tag{6}$$

where μ_x and μ_y is the mean of x and y. The lag/lead generating the highest correlation coefficient can be considered the most likely time delay between the two time-series'. However, this is mainly true when the dependency is linear (Mathworks, 2020c). It is also possible to estimate the time delay between two time-series' by constructing models of the input and the output and observe for which time delay n_k they are most similar. The command `delayest` in MATLAB does this by creating ARX-models of the input and output data and compare them for different time delays n_k (Mathworks, 2020d).

A good way to assess whether a model can be adapted better to the real system is to observe the cross-correlation between the inputs and the residuals, i.e. the prediction errors. For predictive models, the prediction error ε can be expressed as the residuals between measurements y and predictions $\hat{y}$ (Svensson, 2018), given as

$$\varepsilon(t) = y(t) - \hat{y}(t). \tag{7}$$

If the correlation is small, the model is likely well-adapted to the data (Ljung, Glad, 2003).

In any time-series analysis it is common to not only investigate how one time-series correlates with another but also how the time-series correlates with itself, i.e. *autocorrelates* (ESH, 2020). This can be expressed as

$$autocorr(k) = \frac{\sum_{i=1}^{n}(x_i-\mu_x)(x_{i+k}-\mu_x)}{\sum_{i=1}^{n}(x_i-\mu_x)^2}, \tag{8}$$

which is a special case of equation (6). It makes intuitively sense for most time-series' that the value of a data point x_i is correlated to previous data points x_{i-k} as well as future data points x_{i+k}, where k corresponds to the *lead* or *lag*. A lag of 0 means that a data point is compared to itself, yielding an autocorrelation of 1. In general, it is important to be aware of autocorrelation when performing a correlation analysis, this to avoid picking up nonsense correlations between data points because of trends randomly matching (Vilela, Danuser, 2013). When validating or comparing prediction models, the autocorrelation of the prediction errors can be studied. This shows whether the residuals are independent of each other and whether noise has been successfully regarded in the design of the model (Ljung, Glad, 2003).

2.2.2 Transient analysis

In order to better understand a system, one must identify the relevant magnitudes and variables describing the behavior of the system, and more importantly how these interact and affect each other. A general approach is to conduct a *transient analysis*. This is done by varying the system input u as a *step* and register the corresponding behavior of the other measurable variables, i.e. as a *step response* (Ljung, Glad, 2003),

$$u(k) = \begin{cases} 1, k \geq 0 \\ 0, k < 0 \end{cases}. \tag{9}$$

With this simple experiment, system information can be deduced from the results, for example: [1] How other variables are affected by the input signal, [2] What time constants the system possesses, [3] The character of the step response and the level of static amplification (Ljung, Glad, 2003).

It is also possible to gain understanding of a system through its *impulse response*. The impulse response for a dynamic time discrete system is the system output when the input is on the form of a unit pulse (Carlsson, Samuelsson, 2017). For the input u, this can be written as

$$u(k) = \begin{cases} 1, k = 0 \\ 0, k \neq 0 \end{cases}. \tag{10}$$

From the impulse response, information can be obtained such as: [1] The time delay, [2] How fast the system is and [3] If the system is unstable (Carlsson, Samuelsson, 2017).

Sometimes it is difficult to perform a transient analysis on the real system. It is however possible to estimate a step- or an impulse response based on a model of the system. From this, good estimations of system behaviors can be achieved depending on the quality of the model. These can, for instance, be used when validating or comparing different models.

2.2.3 Frequency- and spectral analysis

Another important tool to assess the behavior of a linear system graphically is the Bode-diagram, which is a visualization of the frequency function $G(i\omega)$ of some transfer function G. It consists of two separate plots related to different frequencies ω, one showing the amplitude $|G(i\omega)|$ and the other the argument arg $(G(i\omega))$ (Carlsson, Samuelsson, 2017).

The amplitude curve serves several purposes, for example to deduce the number of poles in the system from counting the number of resonance peaks. In general, one resonance peak entails two poles, two resonance peaks entails three poles and so on. Also, if the resonance peak occurs at higher frequencies then the step response will also tend to fluctuate at a higher frequency and vice versa. The height of the resonance peak determines how fluctuating the step response will be (Carlsson, Samuelsson, 2017).

From the argument curve the phase shift can be read, for every 90 degree shift a general rule of thumb is that there should be one more pole than zeros (Carlsson, Samuelsson, 2017).

Even though the frequency analysis is a sufficient analysis tool it can also be necessary to perform a *spectral analysis*. A spectrum, or a spectral density, ϕ_v describes the frequency content of a signal $v(k)$ (Ljung, Glad, 2003). It is defined as the square absolute value of the Fourier transform of the signal. The spectrum has the unit energy per frequency and the integral of the spectrum between two frequencies ω_1 and ω_2 express the energy in this frequency interval (Ljung, Glad, 2003). The spectrum can be estimated for some given time-series of sampled inputs and outputs. To do this, for example on the input u, the time-discrete Fourier-series (TDF) is used (Carlsson, Samuelsson, 2017). Mathematically it is formulated as

$$U_{TDF}(i\omega) = \sum_{k=1}^{n} u(k)e^{-i\omega k}. \tag{11}$$

The spectrum $\hat{\phi}_{U,TDF}(\omega)$ is then approximated and this estimation is called a periodogram, given as

$$\hat{\phi}_{u,TDF}(\omega) = \frac{1}{n}|U_{TDF}(i\omega)|^2, \tag{12}$$

which can be graphically visualized (Ljung, Glad, 2003). Here n is the number of samples of the input u. From the spectral analysis, it is possible to derive valuable information. For example, an approximation of the frequency function $\hat{G}(i\omega)$ can be derived from the TDF-series of the input and output (Carlsson, Samuelsson, 2017), by solving

$$\hat{G}(i\omega) = \frac{Y_{TDF}(i\omega)}{U_{TDF}(i\omega)}. \tag{13}$$

This frequency function can be useful when constructing a model. One way to utilize it is to compare the frequency function from a created model with the function $\hat{G}(i\omega)$ derived from spectral analysis of the given time-series. If the spectral analysis is performed correctly and the frequency functions behaves similarly, the model is likely to have adapted correct system characteristics (Ljung, Glad, 2003). The command `spa` in MATLAB estimates the spectrum and standard deviations from time-series data. This estimation can be complemented by the command `bodeplot` to visualize the frequency function (Mathworks, 2020e). It is however important to regard some of the limitations with this type of signal analysis. Spectral analysis assumes that the system is in fact linear and does not work if the real system operates with feedback during data collection, i.e. spectral analysis assumes that the input u and the disturbance v are uncorrelated (Ljung, Glad, 2003).

The theoretical background for finding system characteristics is indeed important but the process of modeling also spans beyond the mathematics. Modeling is also a process of intuition and common sense.

2.3 Black-box modeling in detail

2.3.1 Empirical modeling

Creating a detailed physical model (white or grey) of a thermal system in a building is both time-consuming and requires extensive information about the building. This information can be hard to obtain, especially if the building is old and has undergone several restorations. To gain understanding of materials and building characteristics through destructive methods can also be challenging since many older buildings are protected (Kramer et al., 2012). Therefore, a simplified model of the building is often used to approximate the real system.

A way of designing a simplified model is to identify the parameter values through *empirical modeling* (Balan et al., 2011). This concept describes a modeling approach where the parameters are determined by matching the model output to real measurement data. This matching uses an optimization algorithm to minimize the objective function, e.g. the root-mean-squared-error between model outputs and the collected validation data. It is possible to use both linear parametric models and non-linear models for empirical modeling. In contrast, when applying white- or grey-box modeling, the method is to go forward from model characteristics to data instead of vice versa (Kramer et al., 2012).

The process of constructing a linear parametric model via empirical modeling can be explained in three main steps, similar to the general three-phase method approach presented earlier in (Figure 1). Firstly, measurements of the real system are made, i.e. inputs and outputs of the real system are collected. Secondly, a model structure is chosen and its parameters fitted to the data using an optimization algorithm. Lastly, model validation is performed with data which was not used in the parameter fitting (Mustafaraj et al., 2010).

2.3.2 Linear black-box models

There are a significant number of linear black-box models for a modeler to consider, all which vary in complexity and purpose. A general model structure can be found within the realm of linear models which utilizes polynomial structures with or without added noise (Ljung, Glad, 2003). This structure is formulated as

$$y(t) = G(q,\theta)\, u(t) + H(q,\theta)\, e(t). \tag{14}$$

The output y is generated via a transfer function G specifying the dynamics of input u affecting output y and another transfer function H describing the effect of noise e on y. e is assumed to be white noise and non-white noise are being modeled using the term $H(q,\theta)\, e(t)$. The transfer functions provide the polynomial structure of the models (Ljung, Glad, 2003). G and H is expressed as

$$G(q,\theta) = \frac{B(q)}{F(q)} = \frac{b_1 q^{-nk} + b_2 q^{-nk-1} + \cdots + b_{nb} q^{-nk-nb+1}}{1 + f_1 q^{-1} + \cdots + f_{nf} q^{-nf}} \tag{15}$$

and

$$H(q,\theta) = \frac{C(q)}{D(q)} = \frac{1 + c_1 q^{-1} + \cdots + c_{nc} q^{-nc}}{1 + d_1 q^{-1} + \cdots + d_{nd} q^{-nd}}. \tag{16}$$

Here, θ represents the parameter values of $[b_1 \ldots a_{nb} \;\; f_1 \ldots f_{nf} \;\; c_1 \ldots c_{cf} \;\; d_1 \ldots d_{nd}\,]^T$ and n_k the time delay between input u and output y.

There are essentially six different linear models worth considering in this family of linear models: the *ARX-model*, the *AR-model*, the *ARMAX-model*, the *ARMA-model*, the *Box-Jenkins-model* and the *OE-model* (Ljung, Glad, 2003).

One of the simplest linear model structures for input and output time-series' is the autoregressive exogenous (**ARX**) model,

$$A(q)\, y(t) = B(q)\, u(t) + e(t). \tag{17}$$

A common practice is to let *D* and *F* coincide as demonstrated in equation (18) and to set *C = 1* in the general model structure, thereby introducing polynomial *A* on the left-hand side in the ARX structure (Ljung, Glad, 2003), forming

$$D(q) = F(q) = A(q) = 1 + a_1 q^{-1} + \cdots + a_{na}\, q^{-na}. \tag{18}$$

The predictor of the ARX-model takes into consideration previous inputs as well as previous outputs to estimate values of the output (Svensson, 2018), formulated as

$$\hat{y}(t;\theta|t-1) = \big(1 - A(q)\big) y(t) + B(q) u(t). \tag{19}$$

θ represents the parameter values of $[a_1 \ldots a_{na} \;\; b_1 \ldots b_{nb}]^T$.

By allowing $B(q)$ to be zero in the ARX-structure, i.e. excluding the input signal, the model takes on the form of an autoregressive (**AR**) model. The AR-model is a common structure used for representing a stochastic time series (Carlsson, Samuelsson, 2017). In comparison to the general model, no input exists and there is no modeling of the noise *e*. The AR-model can be written as

$$A(q) y(t) = e(t). \tag{20}$$

The AR-predictor, given as

$$\hat{y}(t;\theta|t-1) = \big(1 - A(q)\big) y(t), \tag{21}$$

is only dependent on previous values of the time series itself (Carlsson, Samuelsson, 2017). θ represents the parameter values of $[a_1 \ldots a_{na}]^T$.

A slightly more advanced model compared to ARX is the exogenous autoregressive moving average (**ARMAX**) model which allows for modeling of the noise (Ljung, Glad, 2003). This model can be formulated as

$$A(q)y(t) = B(q)\, u(t) + C(q)\, e(t). \tag{22}$$

The predictor of ARMAX, expressed as

$$\hat{y}(t;\theta|t-1) = \left(1 - \frac{A(q)}{C(q)}\right) y(t) + \frac{B(q)}{C(q)} u(t), \tag{23}$$

use both inputs and outputs to predict future output values (Svensson, 2018). θ represents the parameter values of $[a_1 \ldots a_{na}\; b_1 \ldots n_{nb}\; c_a \ldots\; c_{nc}]^T$.

By allowing $B(q)$ to be zero in the ARMAX-structure, only considering a single time series without input, the model takes on the form of an autoregressive moving average (**ARMA**) model. The ARMA-model is also an enlargement of the AR-model (Carlsson, Samuelsson, 2017), formulated as

$$A(q)y(t) = C(q)e(t). \tag{24}$$

Similarly to the AR-model, the noise is considered an unmeasurable input to the model. By setting C to one in the ARMA-structure, the AR-model is received. The ARMA-predictor, expressed as

$$\hat{y}(t;\theta|t-1) = \left(1 - \left(\frac{A(q)}{C(q)}\right)\right) y(t), \tag{25}$$

regards previous outputs y when determining the predictions (Carlsson, Samuelsson, 2017). θ represents the parameter values of $[a_1 \ldots a_{na}\; c_1 \ldots c_{nc}]^T$.

The most complete linear model structure is Box-Jenkins (**BJ**) and utilizes all available polynomials in the general model, see equation (14) (Ljung, Glad, 2003). The model is structured as

$$y(t) = \frac{B(q)}{F(q)}\, u(t) + \frac{C(q)}{D(q)}\, e(t). \tag{26}$$

The BJ-predictor, formulated as

$$\hat{y}(t;\theta|t-1) = \left(1 - \frac{D(q)}{C(q)}\right) y(t) + \frac{D(q)B(q)}{C(q)F(q)} u(t), \tag{27}$$

use information about previous inputs and outputs to estimate the output (Svensson, 2018). θ represents the parameter values of $[b_1 \ldots b_{nb}\; c_1 \ldots c_{nc}\; f_1 \ldots f_{nf}\; d_1 \ldots d_{nd}]^T$.

A final special case well worth to consider is the Output-Error (**OE**) model which excludes modeling of the noise (Ljung, Glad, 2003), given as

$$y(t) = \frac{B(q)}{F(q)}\, u(t) + e(t). \tag{28}$$

The OE predictor is formulated as

$$\hat{y}(t;\theta|t-1) = G(q,\theta)u(t) = \frac{B(q)}{F(q)} u(t), \tag{29}$$

i.e. the OE predictor only depends on the input u (Ljung, Glad, 2003). θ represents the parameter values of $[b_1 \dots b_{nb}\ f_1 \dots f_{nf}]$. The other predictors differ a bit from the OE-predictor as they depend, completely or partly, on earlier values of the output (Svensson, 2018).

As earlier mentioned predictive models are estimated as one-step predictors, tuned to predict indoor temperature at time t, using inputs and outputs up to $t-1$, but later to be validated by inputs up to $t-1$ and outputs up to $t-k$ to predict the output at time t, it is well-worth to also consider estimating k-step predictors analytically. This can be done with linear regression by solving *normal equations* with the *least squares method* using a prediction horizon of *k* steps for both inputs and outputs (Svensson, 2018). This is possible for both AR- and ARX-models. In general, the normal equation is structured as

$$\hat{\theta} = [\sum_{k=1}^{N} \varphi(t)\varphi(t)^T]^{-1} \sum_{k=1}^{N} \varphi(t)y(t) = [\phi^T\phi]^{-1}\phi^T Y. \quad (30)$$

Here, ϕ can be expressed as $[\varphi(n_0+1)\ \varphi(n_0+2) \dots\ \varphi(N)]^T$ where n_0 is selected as the maximum between n_a and $n_k + n_b + 1$. In the estimation process, N is the amount of data points used for estimation and φ corresponds to the measurement data of $[-y(t-n_k) - y(t-n_k-1) \dots - y(t-n_k-n_a+1)\ u(t-n_k)\ u(t-n_k-1) \dots u(t-n_k-n_b+1)]^T$ for the ARX-model (Svensson, 2018). To estimate a k-step predictor, n_k is selected to be the desirable prediction horizon *k* steps ahead, i.e. 15-steps in this study. The solution of $\hat{\theta}$ entails a minimization of the cost function *V,* given as

$$V(\theta) = \sum_{k=1}^{N} (y(k) - \varphi(k)^T\theta)^2, \quad (31)$$

which is the *sum of squares*. This corresponds to finding the model parameters $\hat{\theta}$ which best fit the data $\varphi(t)$ and $y(t)$ (Carlsson, Lindholm, 2019). The resulting predicted outputs on the estimation data can be computed as

$$\hat{y}(t;\theta|t-1) = \varphi(t)^T\theta. \quad (32)$$

If the aim is to receive outputs from the predictor on validation data, this can be done by constructing ϕ from the validation data and then compute the predictions as in (32) (Svensson, 2018). The predictor uses values of the inputs and outputs up to $t-k$ to predict outputs at time instance *t*. The purpose of applying this additional method is simply to assess the one-step predictors estimated through the MATLAB *System Identification Toolbox* (SITB) and to evaluate whether valuable information exists in knowing the household electricity consumption 15 minutes before each prediction instance *t.*

To summarize, the estimated predictors are validated in two different ways:

1) Tuned one-step predictors uses inputs up to $t-1$ and outputs up to $t-15$ to predict the output at time t.
2) Tuned 15-step predictors uses inputs and outputs up to $t-15$ to predict the output at time t.

Linear models have some advantages compared to non-linear models. Mustafaraj et al. (2010) highlights the simplicity of the models. Often linear models have a low number of estimated parameters and are likely to be compatible with physical models of the studied system. In contrast, a non-linear model like a neural network cannot be related to a physical model. It is also easier to use linear models in control schemes. However, linear models are not always sufficient when modeling reality as many real-world processes tend to be non-linear. Therefore, non-linear models must often be considered.

2.3.3 Non-linear black-box models

Non-linear black-box models are, as the name suggests, used to describe non-linear dynamics of a system. One such model structure is the *neural network* (Sjöberg et al., 1995). It consists of multiple layers of connected nodes, where the connections are called weights. In general, a neural network will be structured around one input layer, one or more hidden layers and one output layer (Schmidhuber, 2014). See (Figure 3) for an example of a neural network with five input nodes, two hidden layers with two and five nodes respectively, and five output nodes.

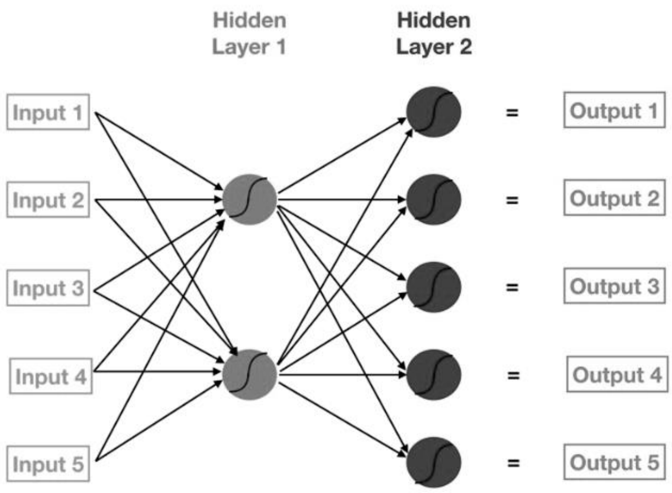

Figure 3. Neural Network (Yiu, 2019).

Like any model, the neural network can be used for predictions of the output given some input. Each input u_i connects to each node in the first hidden layer and is given a weight w_{ij}, the product sum of these plus a *bias* are then passed into an activation function f associated with the hidden layer (Schmidhuber, 2014). The first hidden node output is expressed as

$$f(u_1\, w_{11} + u_2\, w_{12} + \cdots + u_k * w_{1k} + bias). \tag{33}$$

The output values of the first hidden layer's nodes are passed to each node in the next layer (with a new weight) until the final output is reached, which represents the prediction. A common choice of activation function is the *Sigmoid function* (Schmidhuber, 2014). The choices of activation function, initial weights and the dimensions of hidden layers are made by the modeler.

For the purpose of this work, a static neural network model will not be sufficient to describe the relation between electricity and temperature. This because a static neural network will try to fit each instance of electricity to the corresponding instance of indoor temperature at each time step, thereby not considering the dynamics of the system. The indoor temperature does not only depend on the household electricity consumption at time t, but also on earlier values of this variable as well as earlier values of the indoor temperature. Therefore, the MATLAB model framework `nlarx` is utilized which is simply a non-linear ARX-structure that combines the linear ARX-model with a non-linear neural network, yielding a flexible model for prediction (Mathworks, 2020f). The NLARX-model works similarly to equation (32), but instead of $\hat{y}$ being estimated by $\varphi(t)^T\theta$, $\varphi(t)^T$ is inserted into some non-linear function f yielding $\hat{y} = f(\varphi(t)^T;\ \theta)$, in this case a neural network. If $n_b = 0$, a non-linear AR-model is created (NLAR).

NLARX- and NLAR-models requires a neural network structure as an input. While it is possible to construct a neural network manually, we instead opted for MATLAB's pre-defined `sigmoidnet` which only requires the parameter `number of units` to be defined. This represents the number of nonlinearity terms in the sigmoid expansion (Mathworks, 2020g).

Neural networks is one of the most used model structure's when the goal is to capture non-linear dynamics. However, an alternative approach has been developed by Mattsson et al. (2018). In their study, their modeling framework LAVA outperformed the neural network in terms of fit to data and therefore this model structure will also be considered in this work.

LAVA is a system modeling framework developed to learn non-linear models with multiple inputs and outputs by Mattsson et al. (2018). LAVA is itself supported by complex modeling theory but it is not in the scope of this work to dive into the mathematical details of it.

LAVA assumes a model structure with a nominal part $\Theta\varphi(t)$, a latent part $\mathrm{Z}\gamma(t)$ and a white noise process $v(t)$ forming the model presented as

$$y(t) = \Theta\varphi(t) + \mathrm{Z}\gamma(t) + v(t). \tag{34}$$

The idea is then to estimate the parameter matrices Θ and Z to form the final model. The parameters are linear but the model is in fact input-output non-linear, i.e. the relation between the input and the output is non-linear. Here, $\gamma(t)$ is as a non-linear function of $\varphi(t)$. If $\mathrm{Z} = 0$, the prediction errors are solely generated by white noise, allowing the nominal part to capture the system dynamics by itself (Mattsson et al., 2018).

In contrast to the neural network models, the LAVA models are estimated as 15-step predictors. This means that for the non-linear models in this study, the NLARX- and NLAR-models represents tuned one-step predictors while LAVA is tuned 15-step predictors. When predicting the output at time t, the NLARX-models uses inputs up to $t - 1$ and outputs up to $t - 15$, whereas the input-output LAVA-predictors use both inputs and outputs up to $t - 15$.

The purpose of modeling both linear models and non-linear models, as well as tuning one-step predictors and 15-step predictors, is to assess which structure works best for prediction given the studied thermal systems. This will be determined by the process of *model validation*.

2.3.4 Objective functions and model validation

To validate whether a model can be used to describe a real system it is important to analyze the performance of the model. This is often done by cross-validation which is a method to evaluate the prediction errors. The main issue with cross-validation is that not all data is used as estimation data, some has to be earmarked as validation data (Ljung, Glad, 2003). This means that the model will be unable to utilize all available data when estimated. However, from a positive viewpoint, the validation data allows for reliable model testing.

The results of a one-step prediction can often be good even for low-performance models and it is therefore recommended to analyze the prediction errors further, e.g. by the autocorrelation of the residuals, or to use a greater prediction horizon, like the 15 minutes used in this study. For a model structure with noise, the autocorrelation of the residuals should show that the residuals are independent. This independency is true if the autocorrelation is close to zero. It is also ideal that the residuals should be independent of the inputs, otherwise some system dynamics has not been modeled properly (Ljung, Glad, 2003). Therefore should the cross-correlation between the residuals and the inputs be close to zero.

When constructing a model it is crucial to measure its accuracy, partly upon construction but also post construction. This is done with *objective functions*, which outputs the error of the prediction, similarly to equation (7). For example, if one wishes to optimize a neural network, the approach would be to optimize the model parameters so that the objective function output error is minimized (Kenton, 2019). The objective function used during the estimation process helps to estimate the parameters of the model so that the best fit to estimation data is received with respect to the objective function error (Hernándes-Molinar et al., 2016). If instead one wishes to test the performance of the estimated model, the approach would be to compare the model output to validation target outputs, also with an objective function.

A general objective function is the *sum of squares* (SE) of the residuals between the prediction response data $\hat{y}_i$ and the validation data y_i. It is formulated as

$$SE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2. \tag{35}$$

Here, *n* is the number of data samples (Kenton, 2019). In this study, the *SE-fit* is used during the correlation analysis to evaluate how well the one-step predictors mimics the validation data, before selecting the linear model structures. The sum of squares is also used as an objective function to estimate the parameters of the tuned 15-step predictors, see 2.3.2.

Another common objective function is the *Mean squared error* (MSE) which can be described as an extension of the SE now divided with the number of data points n, formulated as

$$MSE = \sum_{i=1}^{n}\frac{(y_i - \hat{y}_i)^2}{n}. \tag{36}$$

MSE simply compares the predicted value $\hat{y_i}$ to the target value y_i, sums the squares of the error and finally averages the error to the number of observations, generating a mean squared error. However, MSE has been criticized for not always being a reliable error measuring tool. LeCun et al. (1990) showed in a study of neural networks that a factor 2 decrease in the number of network parameters yielded an increase in estimation data MSE by a factor 10 while simultaneously reducing the MSE on the validation data. Thus, implying MSE is not always a suitable tool for performance measures.

Another objective function is the *Normalized root mean square error* (NRMSE) function, which in comparison to MSE outputs the normalized root of MSE. It can be expressed as

$$NRMSE = \frac{\sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y_i})^2}{n}}}{\bar{y}}, \tag{37}$$

where $\bar{y}$ is the mean of the measured output y. The NRMSE is used to validate the constructed models in this work. To evaluate how well a model predicts validation data the *goodness of fit* (GOF) can be studied by withdrawing the NRMSE from 1 and multiply it by 100 to receive a percentage value. This is formulated as

$$FIT_{GOF} = 100 * (1 - NRMSE). \tag{38}$$

The GOF between the predictor output and validation data is calculated through the command `compare` in MATLAB (or the command `goodnessOfFit`)[1] (Mathworks, 2018).

Moreover, when validating a model, it is important to consider its stability and its order (Ljung, Glad, 2003). The stability of the model can be analyzed in several ways. For a linear time-invariant discrete system, one approach is to observe the poles of the system. If they are strictly placed within the unit circle, the system is input-output stable. Another way of analyzing the stability is to observe its impulse response. If the output does not converge to a stationary value, the system is not stable (Carlsson, Samuelsson, 2017). This could be expressed in terms of the weight function $h(k)$ of the system and the criteria

$$\sum_{n=0}^{\infty} |h(n)| < \infty. \tag{39}$$

Regarding orders of models there are drawbacks in using too high orders. This can be explained by the concept of overfitting, where the parameters of the model adapts to noise characteristics from the estimation data. To analyze if the order of the model can be reduced, the positions of poles and zeros should be investigated. If some zeros or poles are overlapping, or are located close to each other, it may be suitable to reduce the order (Ljung, Glad, 2003). Another method is to study the parameters of the model. If some of the parameters are estimated close to zero, they may be redundant.

1 For the MATLAB version R2018b the command `goodnessOfFit` calculates the fit in a manner identical to `compare`, when using cost function NRMSE (Mathworks, 2018). For later versions, such as MATLAB R2020a, this do not hold.

While the details of black-box modeling are certainly important, some attention should also be dedicated to understanding the specific characteristics of the system to be depicted. The next section gives a detailed description of how thermal systems has been modelled in earlier works.

2.4 Modeling a thermal system

Modelling the behavior of the indoor temperature in an apartment is challenging, as it is a time-varying system affected by several factors (Huang et al., 2012). A changing outdoor temperature, the number of persons in the apartment, wind conditions and solar insolation are some examples. The usage of household electricity and heat flow between adjacent climate zones in the apartment also affects the indoor temperature. Several studies have tried to model these types of systems. For example, Huang et al. (2012) has presented a model for predicting the temperature between diverse thermal zones inside Terminal one of the Adelaide airport using a black-box modeling approach with neural networks. By measuring both controllable and uncontrollable variables and applying empirical modeling to the data set, two days of accurate predictions were achieved.

Another researcher, Mustafaraj (2010), compared linear parametric models for predicting the indoor temperature in an office using different valuation criteria. GOF, coefficient of determination, mean absolute error and MSE were analyzed between the output of the model and the real data and the results showed that the Box-Jenkins-model performed better than the ARMAX- and ARX-models. This result depends first and foremost on the Box-Jenkins noise handling, which proved more accurate than the other models.

2.4.1 Electricity consumption and indoor temperature

A significant amount of previous work has been dedicated to estimating heat output of common objects present in an apartment. Zavattoni et al. (2014) has listed the wasted thermal energy of some of the most common household electrical appliances during duty cycles. Electrical ovens were estimated to generated approximately 245 Wh of thermal energy for each cycle. Washing machines contributed to 550 Wh and dishwashers to 230 Wh thermal energy during one of the cycles (Zavattoni et al., 2014). In a similar work, Suszanowicz (2017) determined average values for the heat emission coefficient of different light sources. A halogen light bulb had an average value of 0.82 W/W and a led bulb 0.08 W/W.

When electrical appliances are used, the indoor temperature will be affected. This must however not only depend on the actual heat emissions from the electrical components. An increase of household electricity could also indicate that people are present inside the building which can be seen as an additional thermal energy source. Akiful et al. explains this: “So human body [*sic*] has an interior core which acts like a heat generation source where the heat generation depends on the rate of metabolic activities” (2017, p.1).

When a human sleeps, the body generates about 75 W of thermal energy and 1000 W for more demanding aerobic activities (Akiful et al., 2017). There is a possibility that the usage of electricity can indicate that people are present and that they are not in rest. During the night for example, the household electricity consumption goes down as electrical appliances are not used in the same scale. During these times, the power consumption first and foremost consists of cycling appliances, e.g. fridge and freezers, or electrical components that are on standby while people are at rest (Firth et al., 2008). Also, the usage of hot water tends to decrease during the hours when people normally sleep or are away from home (Svensson, 1973). Usage of hot water, through for example showers, will affect the indoor temperature and it is possible that this could be detected by observing the trends of household electricity consumption.

2.4.2 Thermal zones

Heat, ventilation and air-conditioning (HVAC) are strongly related to the energy consumption of a building (Afroz et al., 2018). In this work, the focus is on modeling the temperature dynamics (the thermal systems) of the rooms in the studied apartment, all of which are highly affected by the HVAC system. The aim is not to understand the dynamics of the thermal system, but to construct a model able to predict future temperature changes. Therefore, the relation between the HVAC and the thermal system is not analyzed deeply but an understanding of usual assumptions made about the HVAC system helps delimit the range of studied variables of the thermal system during the tests. Afroz et al. (2018) assume upon modeling a HVAC zone: [1] normal temperature distribution for each zone, [2] that the effect of opposing walls on the zone temperature is equal, [3] that the floor does not affect the zone temperature and [4] that there are no pressure losses across a zone or in the mixed regions (which could lead to increased airflows between them).

For larger buildings, e.g. offices or industrial buildings, the thermal system is often divided into multiple thermal zones (Huang et al., 2012). Each zone is considered a subsystem and is controlled by an individual air handling unit. Huang et al. (2012) highlights the fact that a global model of the thermal system cannot predict the outcome of every thermal zone and that individual models therefore has to be used. However, a relation between the thermal zones does exist through heat transfers between adjacent zones, which should be considered in the model. Also, for smaller buildings, the usage of multiple thermal zones is commonly adapted. For instance, Voll et al. (2016) defines each room as a thermal zone upon construction a simulation model of a nearly-zero energy building. The reference building in that study was connected to district heating and used radiators as its energy source, similarly to our apartment of study.

Souza and Alsaadani (2012) also investigates the impact of different zoning strategies through thermal simulations of an office area. A five-zone model and a single-zone model were simulated and then compared to a target model. The target model was divided into several zones and simultaneously regarded the specific heat generating activities for each room, making it the most accurate. Results showed that the five-zone model estimated the annual heating demand of the office area well, but more importantly that an accurate target model can be constructed based on multiple zones regarding each room’s characteristic.

Fielsch et.al. (2017) shows that the usage of Model Predictive Control can improve energy consumption and human comfort also in small-scale heating systems, compared to the more conventional usage of Model Predictive Control in office buildings. When constructing the model of the building, every room was considered one thermal node which is then connected to the adjacent rooms, the outdoor temperature and the ground temperature. Thus, when constructing a model of a thermal system of a building, choosing each room to represent a thermal zone is a well-established approach.

2.4.3 Sampling time and thermal time constant

For the dynamic variable of indoor temperature, a change in heating or cooling will not lead to a direct variable change, which can be explained by the building's thermal time constant. This constant describes the thermal inertia of the building, i.e. how long it takes for the indoor temperature to change to the outdoor temperature when the heating- or cooling system is turned off (Karlsson, 2012). Large buildings often have time constants that exceeds 100 hours (Hietaharju et al., 2018). Therefore, when collecting measurement data for indoor temperature, the sample time is often set to hours or days. Hietaharju et al. (2018) uses a sample time of one hour to collect data for indoor temperature, outdoor temperature and heating power, aiming to construct a predictive model of the indoor temperature via grey-box modeling. This sample scale is however not generalizable for all studies in this area. Mustafaraj et al. (2010) instead uses a sample time of five minutes to study the potential of linear parametric models to predict indoor temperature and humidity.

Comparing the variable household electricity consumption with the indoor temperature, the electricity is of a more fast-changing kind. As soon as electrical appliances are turned on or off, this variable will change instantly. Some of the electrical activities, only performed during limited times, can potentially indicate a change in human activity and/or heat generation from electrical components, and potentially (with a time delay) a changing indoor temperature. It is therefore important to choose a sampling time that is not too big. Chujai et al. (2013) uses a sample time of one minute to measure the electricity consumption of a household to find a model to predict this consumption. This sample time is adapted in our study.

3. Method

This section aims to clarify how the study was designed and executed. Firstly, the necessary equipment and software are described. Then the process of data collection and the validation strategy is dealt with, as well as a layout description of the apartment. After this, a detailed description of the modeling process follows.

3.1 Equipment

In order to collect data from the apartment certain measuring tools were necessary. Indoor temperature and electricity consumption are the primary data points assessed. Therefore, two different measuring tools were utilized, *Tinytag Ultra 2* (Figure 4) and *Logger 2020* (Figure 5). A total of 8 Tinytags were placed inside the apartment, measuring indoor temperature, surface temperature from radiators and relative humidity. Logger 2020 registers the electricity consumption online, giving close to instant access to the electricity consumption at any time.

Figure 4. Tinytag (Intab, 2019).

Figure 5. Logger 2020 (Energibutiken, 2019).

3.2 Software

MATLAB is the primary software of this project. Given its rich library of tools for empirical modeling and technical calculations, it was a suitable software to work with. For example, the *System Identification Toolbox* (SITB) and the neural network construction library *Deep Learning Toolbox* both provide tools which serve to simplify the modeling process. We have chosen to execute the data pre-processing with Python. Also, we had have the opportunity to use already existing code sequences that simplifies the construction of LAVA-models, brought from assistant professor Per Mattsson at Uppsala university.

3.3 Experimental design and data collection

3.3.1 Overview of the experimental design

In order to construct an accurate model of the thermal system it is important that the experiments and measurements of the system are carried out carefully. To evaluate the potential of predicting indoor temperature in an apartment with knowledge of household electricity consumption, an understanding of the relation is clearly of essence. Beyond this, to understand the attributes of the heating system, e.g. its inertia or its sensibility towards disturbances, measurements had to be made so that disturbance signals can be described by their typical characteristics (Ljung, Glad, 2003). This was of importance to build an understanding of the behavior of the indoor temperature.

In every room of interest, i.e. kitchen, bedroom, bathroom, living room and middle room, an internal temperature sensor was placed. This sensor measured the indoor temperature and it was positioned with a minimum of one meter from the floor and outer wall to prevent it from being affected by the surface temperature of these areas (Karlsson, 2019). The sensors placed in the kitchen, bathroom and bedroom also measured the relative humidity. In the kitchen, bedroom and living room, external temperature sensors were attached to the radiators, measuring the surface temperature. Every sensor used a sample time of one minute.

3.3.2 Room specifics

As the disposition of the apartment significantly impacts the measurements from the different rooms, an overview of the apartment will now be given. For reference, see Figure 6.

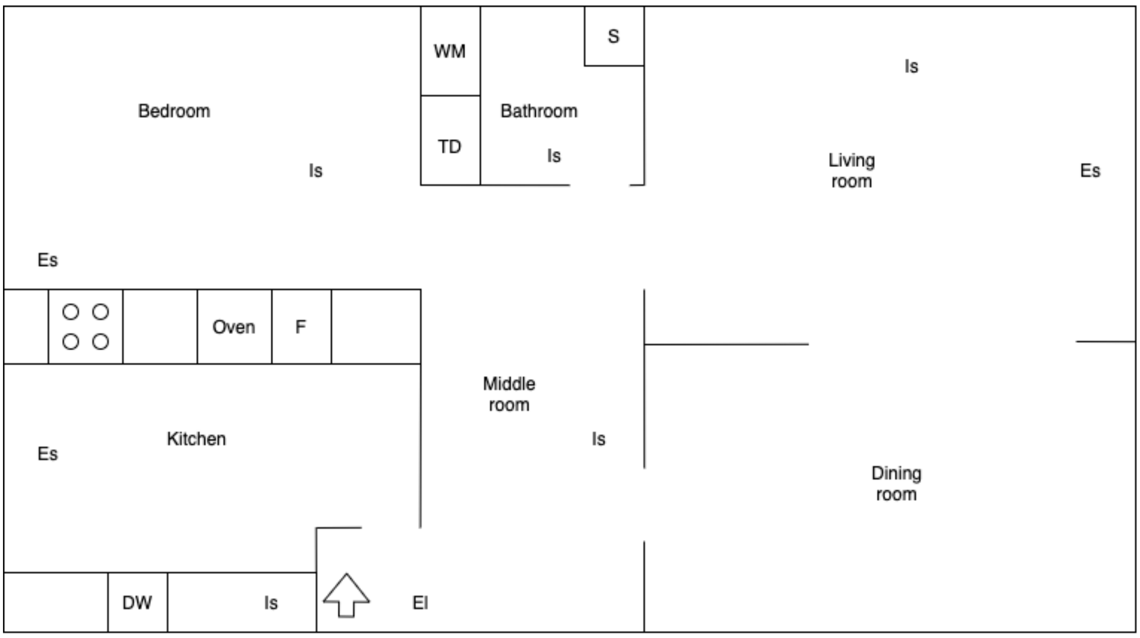

Figure 6. Apartment overview and the location of internal sensors (Is) and external sensors (Es).

The kitchen faces south and is equipped with standard kitchen appliances. It has an oven, an induction stove, a fridge (**F**), a freezer (**F**) and a kitchen fan above the stove. Except from this, a coffee maker and several light sources are installed. Attached to the radiator, an external Tinytag sensor (**Es**) was placed measuring the radiator surface temperature (Figure 7). This radiator is placed below a window facing south. An internal Tinytag sensor (**Is**) was placed on the kitchen sink (Figure 8) at the east interior wall measuring indoor temperature and the relative humidity.

The middle room connects to all other rooms, it is an entrance lounge. The switchboard (**El**) is placed to the right of the entrance and a Logger 2020 was attached to this. Above the entrance, a router is located and activated. The middle room has two light sources. At the interior wall, in connection to the bathroom, a smaller radiator is located, but it was not measured as it was non-functioning. By the wall between the doorways of the living room and the dinner room an internal Tinytag was placed measuring the indoor temperature (Figure 9).

The bedroom faces south and links directly to the middle room. The room is adjacent to the kitchen and the bathroom, with interior walls separating them. The bedroom is equipped with a TV, two bed lights and a main light source at the center of the room. Adjoint to this room, at the south outer wall, double doors opens to a balcony. At the direction towards the middle room, an internal Tinytag was placed measuring the indoor temperature (Figure 10). In the southeast corner of the room, at the interior wall towards the kitchen, a radiator is located. On this, an external Tinytag was attached (Figure 11).

The bathroom is equipped with a washing machine (**WM**) and a tumble dryer (**TD**). Except from this, the bathroom has a toilet, a shower and a sink. By the door, an internal Tinytag was placed measuring the temperature and the relative humidity of the room (Figure 12).

The living room faces north and has a big window in this direction. The main electrical appliances are a TV, a PlayStation, a sound system and a light source in the middle of the room. Adjacent is the middle room and the dinner room. Beneath the window at the outer wall facing north, a radiator operates. An external Tinytag was attached to this (Figure 13) and by the northwest interior wall an internal Tinytag was placed measuring the indoor temperature of the room (Figure 14).

The dinner room is located next to the living room and is facing north with a large window. The main electrical consumer is a central light source. Below a window, a radiator operates but it is not measured. This room is considered the least important for this study and is therefore of lower priority. Mainly because of its lack of electrical appliances and human activities.

Figure 7. Es: Kitchen *Figure 8. Is: Kitchen* *Figure 9. Is: Middle room*

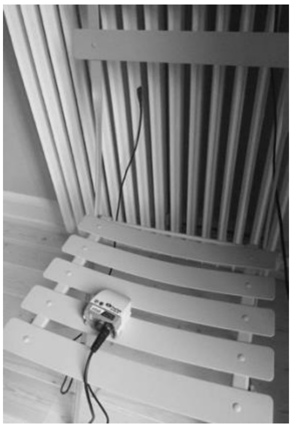

Figure 10. Is: Bedroom *Figure 11. Es: Bedroom* *Figure 12. Is: Bathroom*

Figure 13. Es: Livingroom *Figure 14. Is: Livingroom*

3.3.3 Evaluation process

The primary validation method used was focused on answering whether benefits exist in knowing the household electricity consumption when predicting future indoor temperatures. To gain understanding of this, different types of models were constructed and compared for every room. Firstly, for each room and model type (linear and non-linear), input-output models were modelled. These models handled the household electricity consumption as an input signal and used information about it, as well as information about previous indoor temperatures, to estimate future values of the latter (Figure 15). Secondly, no-input predictive models were constructed. These models only consider information about previous indoor temperatures to estimate future values of the temperature (Figure 16). The comparison between the different predictive models was based on their ability to mimic the validation data with respect to the NRMSE, i.e. GOF.

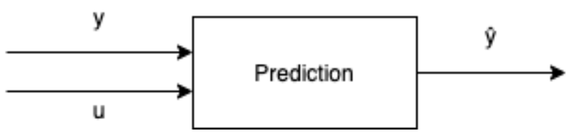

Figure 15. Predictor that both considers the indoor temperature y and the household electricity consumption u when estimating future indoor temperatures $\hat{y}$.

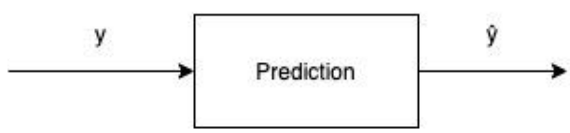

Figure 16. Predictor that only considers the indoor temperature y when estimating future indoor temperatures $\hat{y}$.

This evaluation method was used for all tested models. At first, linear models were constructed as one-step predictors for every room and then evaluated and compared to the predictive performance of tuned 15-step linear predictors. Secondly, non-linear NLARX- and NLAR-models were constructed for every room and they were compared to tuned 15-step LAVA-predictors.

On validation, the tuned one-step predictors uses more input data than the tuned 15-step predictors while predicting the output at time t, i.e. input data up to $t-1$ are used in comparison to the tuned 15-step predictors which uses inputs up until $t-15$.

All models were evaluated on a prediction horizon of 15 minutes, using more or less data points from household electricity. However, only the tuned 15-step linear predictors and the tuned 15-step LAVA-predictors were tuned to estimate the indoor temperature 15-steps in advance. The idea behind this is to determine the significance of the household electricity consumption. If tuned one-step predictors, which uses more input data, performs better predictions on a horizon of 15 steps than tuned 15-step predictors, the advantage is likely to exist in the input signal itself. In that case, up to one minute pre prediction, the household electricity consumption brings valuable information to predict the indoor temperature. Otherwise, if the input is redundant, it is reasonable to believe that tuned 15-step predictors would be better at predicting the indoor temperature at the very prediction horizon it is tuned for.

To evaluate the benefits of tuning predictors on the prediction horizon of interest, no-input tuned one-step predictors are compared to their corresponding no-input tuned 15-step predictors in terms of their GOF.

3.3.4 Data collection

The measuring stage of the project was initiated with a three-day pilot study. Between the 11th and 14th of December, three days' worth of data was collected, primarily focusing on indoor temperature and electricity consumption. During this period, extreme tests were performed. For one hour multiple electrical devices were turned on, and then switched off for three hours. This procedure was repeated for multiple cycles during day-time. The reasoning behind this method was that if electrical appliances has a clear impact on the indoor temperature it should be visible from these tests. The three hours of "rest" made it possible for the apartment to return towards a steady-state in terms of indoor temperature. The purpose of this shorter study was to deduce whether an actual correlation exists between indoor temperature and electricity consumption. It also serves the purpose of assessing the equipment, making sure it is working correctly. Finally, simpler models were constructed to provide some information on how well a model can fit to the data.

In a similar manner to the pilot study, a longer measuring period was executed between the 10th and 24th of February, collecting two weeks' worth of data. This period was chosen as it is generally cold in Sweden in February, meaning that the radiators are operating on full effect periodically. At this stage, all different data points were collected, i.e. indoor temperature, outdoor temperature, radiator heat, electricity consumption and relative humidity. Instead of executing extreme tests on the apartment, two persons lived in it and used it as a normal household. The purpose of this study was to investigate the prediction performance of different model types, representing the thermal systems of the rooms in the apartment, and through that evaluate if knowledge about household electricity consumption could improve the model performance.

3.4 Modeling of the apartment

The modeling of the apartment's thermal systems is the core of this study and it was executed in three instances of modeling. A correlation analysis on the three-day data was the first step. For this, a prediction horizon of 15 minutes was deemed sufficient, meaning that it is considered a small enough time-span to compare the predictive performance of tuned one-step predictors with tuned 15-step predictors. It is also considered a large enough time span to evaluate the importance of using more information about the household electricity consumption, i.e. up to $t-1$ for the tuned one-step predictors. The second step was the modeling of both linear- and non-linear models from two weeks of data, again using a prediction horizon of 15 minutes, to evaluate which model structures best represents each room of the apartment. The third step was a sensitivity analysis which was made for the purpose of testing the impact of different input signals (electricity consumption, radiator temperature and relative humidity) on the indoor temperature.

3.4.1 Correlation analysis approach

The modeling process was initiated with a correlation analysis to deduce if a relationship exists between the variables of interest, i.e. household electricity consumption and indoor temperature. This was done for all the rooms, generating an overview of the differences between the thermal zones of the apartment. The correlation analysis was executed in three main steps, with focus on studying the visual relationship, the mathematical correlation and finally the benefits of knowing the household electricity consumption to predict the indoor temperature. Lastly, comparisons between the 15-step predictions from constructed ARX- and AR-models (tuned one-step predictors) and tuned 15-step ARX-and AR-predictors of the same polynomial order were made.

Firstly, a simple graph of the three-day data was plotted, visualizing the relationship between temperature and electricity. From this, a visual correlation was analyzed. A scatter plot of the temperature as a function of electricity was also made, with $C°$ on the y-axis and kW on the x-axis.

Secondly, the correlation was tested mathematically. We used two methods for this task, Spearman's correlation method and Pearson's correlation method, both of which compares the correlation between two arrays of values. Through Spearman's method, p-values to test the null hypothesis were also calculated. However, these methods are not optimal to find non-linear correlations and impacts of time delays, meaning alternative methods were necessary. Therefore, we also investigated how much additional information the household electricity data brought in terms of its ability to help a model predict the indoor temperature. If a simple AR-model, only modelled to predict indoor temperature based on previous indoor temperatures, performs equally well or better than an ARX-model with information of <u>both</u> indoor temperature and electricity, then the household electricity consumption can be considered redundant as a model parameter. To evaluate this, both AR- and ARX-models were constructed.

Thirdly, the time delay n_k between indoor temperature and household electricity consumption were estimated for each room. The ARX-models used to estimate the time delay had a A-and B-polynomial of the fourth order, this was decided to be sufficient, minimizing the number of parameters and avoiding parameters near zero. $SITB$ was used to find the ARX-model with the estimated time delay n_k which best predicted the data one step into the future, i.e. had the best SE-fit of the one-step predictions. It was found trough observations that higher orders of the A-and B-polynomial tended to result in small parameters with none, or small, improvements in performance. Therefore, only polynomials up to the fourth order were investigated. Lastly, the ARX-model was compared with an AR-model with respect to their predictive performance 15-steps into the future. The order of the A-polynomial of the AR-model was chosen to be the same as the A-polynomial of the ARX-model for the corresponding room. The models were tuned on the first half of the collected data (estimation data) and validated on the second half of the data (validation data).

Lastly, 15-step predictors were constructed from the estimation data of each room. The previously estimated time delays and orders of the constructed ARX- and AR-models were used to decide the structure of the 15-step predictors. Also, the same estimation data was used to tune both types of models. The prediction performance of the validation data was then calculated as GOF and compared to the performances of the ARX- and AR-models of the same room.

3.4.2 Two-week modeling approach

The aim of the two-week data modeling was to compare the predictive performance of input-output models and no-input models, as well as tuned one-step predictors and tuned 15-step predictors, for both linear and non-linear models of the rooms in the apartment (Figure 17). Unlike the correlation analysis, the process of selecting a representative model was extensive and included deeper analysis, primarily regarding the linear models. For these models, focus was not only fixed on finding the best predictive models in terms of GOF. Consideration was also taken to other model characteristics, e.g. stability, placements of poles and zeros and the behavior of the residuals.

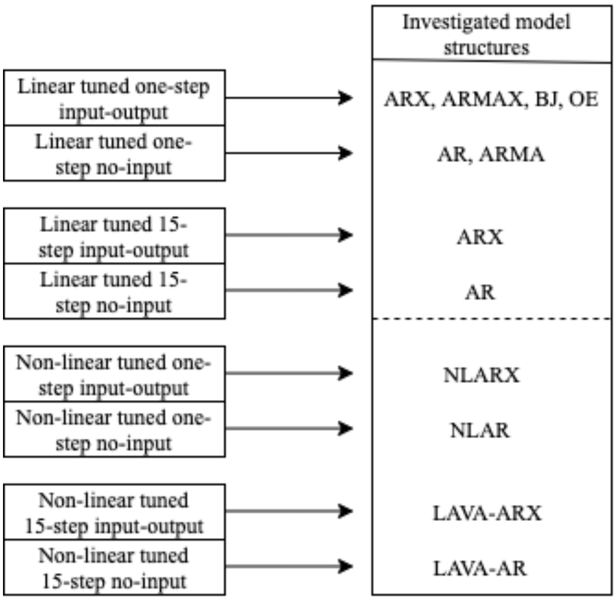

Figure 17. Model types investigated during the two-week data modeling.

Firstly, linear models were constructed and compared by their 15-step prediction performance to AR- and ARMA models. Also, 15-step ARX-and AR-predictors were compared to these models and to each other in terms of GOF. Secondly, NLARX-models were also constructed and compared in terms of their 15-step prediction performance to NLAR-models, and the previously constructed linear models. After this, 15-step LAVA-predictors were constructed and compared with the previous models regarding their GOF. The models were tuned on the first half of the two-week data (estimation data) and validated on the second half (validation data).

3.4.3 Two-week linear modeling

The process of finding the most suitable linear models was executed in the same way for every room. To begin with, the time delay between the household electricity consumption and the indoor temperature was obtained through ARX-estimations of the inputs and outputs. To ensure that a good polynomial order was used, all orders for the

A-and B polynomial were tested up to the fourth order and the most frequently attained time delay was selected. After obtaining the time delay, the best models with respect to GOF of the 15-step predictions were generated. This was done for the linear models ARX, ARMAX, BJ and OE and the estimated models were then evaluated. The order limit was chosen as a result of the observations received during the correlation analysis, which indicated that higher polynomial orders tended to result in small values for the estimated parameters without yielding any significant improvements in model performance. Every model was then analyzed by their residual characteristics, placement of poles and zeros, step- and frequency response and estimated parameters using a checklist. A satisfying behavior for a model in any of these areas was awarded with a check (√) and an unsatisfying behavior with a cross (X).

The residuals were analyzed in terms of the autocorrelation between the prediction errors, and the cross-correlation between the input and the residuals. A confidence interval of 99 percent was used to evaluate if the correlations could be interpreted as white noise only, which is preferable. The poles and zeros of the system was studied primarily in terms of stability for the model, i.e. if poles were placed inside the unit circle. Their placement was also analyzed in relation to each other, so that no poles risked being cancelled by other poles or zeros.

The step responses were examined for every model and compared to the behavior of the other models, to clarify deviations in for example convergence. This was also true for the frequency responses, which also compared the behavior of the various models with the estimated frequency response derived from spectral analysis. These frequency responses were visualized in Bode-diagrams. For every model, the estimated parameters were examined by their significance for the model in question. Thus, estimated parameters smaller than 0.001 resulted in order reduction of the corresponding polynomial. This did, first and foremost, concern the B-polynomials. However, the minimum B-polynomial order was set to one for the input-output models.

When the linear input-output models were analyzed, the most suitable models for every room were selected and compared to the best performing AR- and ARMA-models with respect to GOF. Here, AR- and ARMA-models up to the fourth order were tested.

Lastly, the tuned one-step predictors were also compared to tuned 15-step predictors of ARX- and AR-structures with the same orders on the A-polynomial (and B for ARX) as the previous best performing input-output model.

3.4.4 Two-week non-linear modeling

For the neural network models, the aim was to try to find better 15-step prediction performances compared to the linear models while also investigating the benefits of using information about household electricity consumption to predict indoor temperatures. The focus was not to find the absolute best performing models, but simply to see if the prediction performances could improve using neural networks. Therefore, the same estimation- and validation data used to construct and evaluate the linear models were used to train and evaluate the neural network.

Firstly, NLARX-models where constructed for every room with similar n_a, n_b and n_k as the selected linear model of the corresponding room. For a specific room, different layer structures were tested and the neural network was trained on the estimation data using a sigmoid network nonlinearity estimator. The number of units tested ranged from 1 to 20 and the GOF was calculated, using information about inputs up to $t-1$ and outputs up to $t-15$ to predict the output at time t. The best performing NLARX-model for every room was then compared with a NLAR-model of that specific room. The structures of the NLAR-models, i.e. the value of n_a, were selected to be the same as for the corresponding best performing linear no-input model regarding GOF. The number of units tested for the NLAR-models ranged from 1 to 20.

The LAVA-models were the final model structures tested in this work and they were tuned in as 15-step predictors. The purpose of this was to investigate whether any non-linear dynamics were lost when 15-step predictors were tuned in on the estimation data and only used inputs up to $t-15$ to predict the output at time t. Also, comparisons between input-output LAVA-models and no-input LAVA-models were made.

The LAVA-models were constructed based on the best linear AR- and ARX-model structures earlier found for each room. The data was also divided into estimation data and validation data in the same manner as for the previously tested types of models. Then, a value for the resolution of the *Laplace basis M* and a value of number of iterations per sample in cyclic minimization *L* was chosen. We opted for three variations on [M, L] which was [4, 5], [3, 5] and [4, 6] resulting in three results for each room and each model (AR and ARX). After constructing the models, the best one's were selected based on the GOF and they were then compared to the performance of the previously constructed models.

3.4.5 Sensitivity analysis

For the kitchen, data has been collected for both the relative humidity and the radiator temperature during the two-week period. Intuitively, the radiator temperature should help predict the indoor temperature as its surface temperature directly affects the indoor temperature. This means that the heat of one radiator through thermal radiation affects the indoor temperature in the surrounding area. This is not true for the household electricity consumption as it was measured for the whole apartment, meaning that its impact on indoor temperature is spread throughout the rooms. Apart from this, the household electricity consumption has to be converted to heat before affecting the indoor temperature in the apartment, e.g. by heat losses in electrical appliances or by heat generating human activities.

Hot water usage and human activities in a room may affect the humidity as well as the indoor temperature. To study the difference in prediction performance when using household electricity consumption, relative humidity and radiator temperature as input signals, a sensitivity analysis was performed.

OE-models were constructed with one input of each kind. Polynomial orders up to the fourth order were tested and the best performing structures with regard to GOF were selected. The same procedure was executed with ARX-models. Their predictive performance was compared to each other and the OE-models. At last, a *multiple input-single output (MISO)* ARX-model were constructed which used all inputs to predict indoor temperature. The structure of this model corresponded to the single input ARX-models previously constructed and its predictive performance was studied in relation to the other models.

As the models with the input relative humidity proved to achieve the worst GOF, a MISO ARX-model with household electricity consumption and radiator temperature as inputs was also constructed.

4. Data

A crucial part of this project is the collection and pre-processing of data. This section aims to clarify the data types being used and how the pre-processing takes place.

4.1 Data description and pre-processing

The data collected is divided into six different variables, *date*, *time*, *indoor temperature, radiator temperature*, *relative humidity* and *electricity.* While time and date only contribute with information in terms of indexing, the other variables are measurements of the systems. Essentially, two datasets are utilized, one set for the pilot study (three-day data) and one set for the empirical modeling phase (two-week data). The data has the following types and forms (Table 1).

Table 1. Data for each measured variable, their format type and form.

Variable	*Type*	*Form*
date	String	YY-MM-DD
time	String	HH-MM-SS
indoor temperature	Float	xx.yy
radiator temperature	Float	xx.yy
relative humidity	Float	xx.yy
electricity power	Float	xx.yy

Tinytags were used to collect indoor temperature data. The Tinytag-software automatically constructs a downloadable .txt-file containing indoor temperature and relative humidity, indexed by date and time. Five rooms had internal temperature sensors, meaning five different files were made and subsequently five pre-processing iterations were conducted. As the data was collected in one continuous time interval, no appending of data files was necessary. However, the beginning and the end of the files had a smaller number of outliers, these outliers were removed in all files. Finally, trends and means were subtracted.

The electricity power data is collected with the Logger 2020 which utilizes an online software to save electricity power data, indexed by date and time. From each room a different number of usable data points were collected. The final file format can be seen in Figure 17.

Temp	kWh
22.70	0.130
22.60	0.128
22.50	0.128
22.45	0.130
22.30	0.136
22.30	0.124
22.20	0.077
22.20	0.455

Figure 17. Part of three-day data.

The pre-processing of the two-week data is the same as for the three-day data with two exceptions. The first difference being that the data was not collected in one continuous time interval but two, meaning two files had to be appended for each room. This created some issues at the points of attachment which created temporary peaks in the data, this was resolved by simply removing data points to create a smooth interpolation. The second difference was that the data for the kitchen and bedroom (both rooms oriented to the south) contained outliers which created unnatural data peaks, probably due to sunlight hitting the sensors. These peak values were cut to not exceed a fixed temperature limit. These limits were selected by observing how the temperature had behaved before the peaks. After this, the data was ready to be modelled.

5. Results

This section is dedicated to the collected results. Firstly, the results from the correlation analysis are described. Secondly, the results from the two-week study follows. Finally, the sensitivity analysis is presented.

5.1 Correlation analysis

5.1.1 Kitchen

From the three-day tests, a visual correlation can be seen between the household electricity consumption and the indoor temperature in the kitchen (Figure 18). Each local peak of electricity usage corresponds to a delayed local peak of the indoor temperature. The indoor temperature possesses more thermal inertia compared to the electrical consumption, as the changes of the first signal are slower and softer than the changes in the latter. The smaller periodical electricity peaks found during the times when the electricity consumption is low corresponds to cycling appliances such as the fridge and freezer.

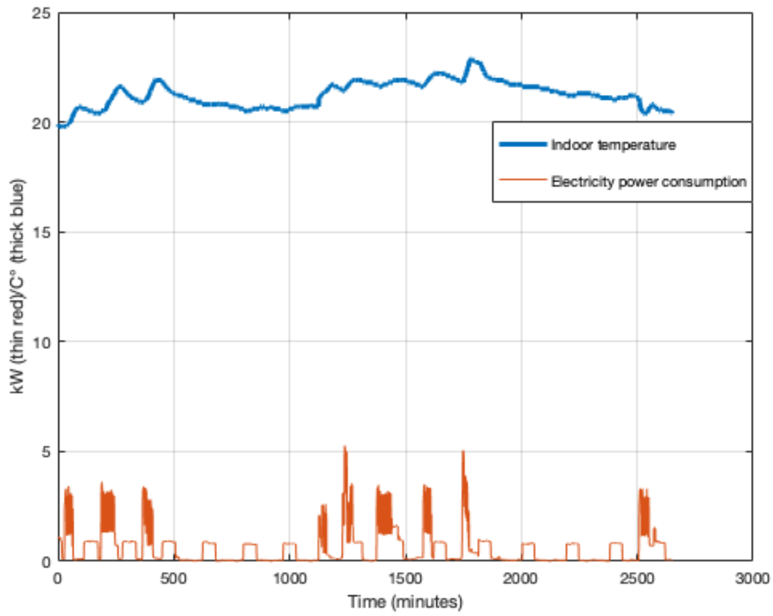

Figure 18. The indoor temperature and the household electricity consumption in the kitchen during three-day tests.

Calculations of Pearson’s correlation coefficients and Spearman’s rank correlation coefficients show no clear indications of linear or order relationship association (Table 2). However, the null hypothesis can be rejected within the confidence interval of 95 percent, meaning that we cannot with certainty conclude that there is no correlation.

Table 2. For kitchen calculated values of Pearson's correlation coefficient p, Spearman's rank correlation coefficient ρ and its p-value.

p	0.0156
ρ	0.1169
p-value	1.5794e-09

The scatter plot of the indoor temperature and the household electricity consumption shows no specific tendencies of correlation (Figure 19). This scatter plot analysis was conducted for each room but all show the same pattern, therefore these results are excluded further.

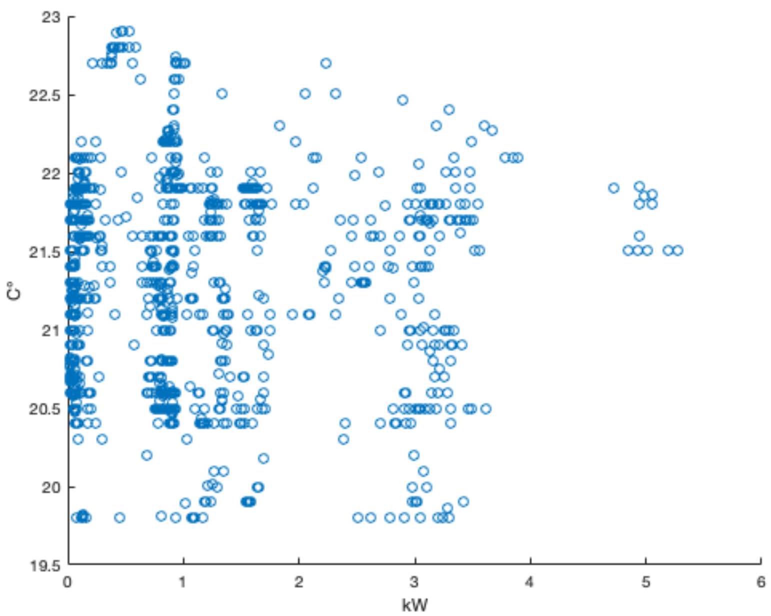

Figure 19. Scatter plot of the indoor temperature in the kitchen as a function of the household electricity power consumption during three-day tests.

By estimating the time delay between the household electricity consumption and the indoor temperature, a time delay of 15 minutes was found. The ARX-model with best SE-fit was a model with an A- and B-polynomial of order 2 and 1. The model has a GOF of 80.51 percent (Table 3).

The constructed AR-model has an A-polynomial of order 2. Its 15-step predictions performed worse than the ARX-model with a GOF of 78.64 percent (Table 3). Regarding the constructed 15-step predictors, the ARX-predictor performed slightly better than the AR-predictor. The tuned 15-step predictors performed worse than the tuned one-step AR-and ARX-models (Table 3).

Table 3. For the kitchen, performance of the tuned 15-step AR-and ARX-predictors and the tuned one-step AR- and ARX-models.

FIT_{GOF} *(ARX)*	80.51%
FIT_{GOF} *(AR)*	78.64%
FIT_{GOF} *(15-step ARX)*	77.58%
FIT_{GOF} *(15-step AR)*	76.35%

5.1.2 Bathroom

The bathroom results only shows weak tendencies of correlation (Figure 20). Small local peaks of the indoor temperature seem to occur slightly after the peaks of electricity consumption but it can be argued inconclusive. Longer rest periods in household electricity consumption seems to result in periods of less oscillating character for the indoor temperature, i.e. between minute 700 and 1200 and between minute 2000 and 2500. However, the visual relationship cannot be determined with certainty.

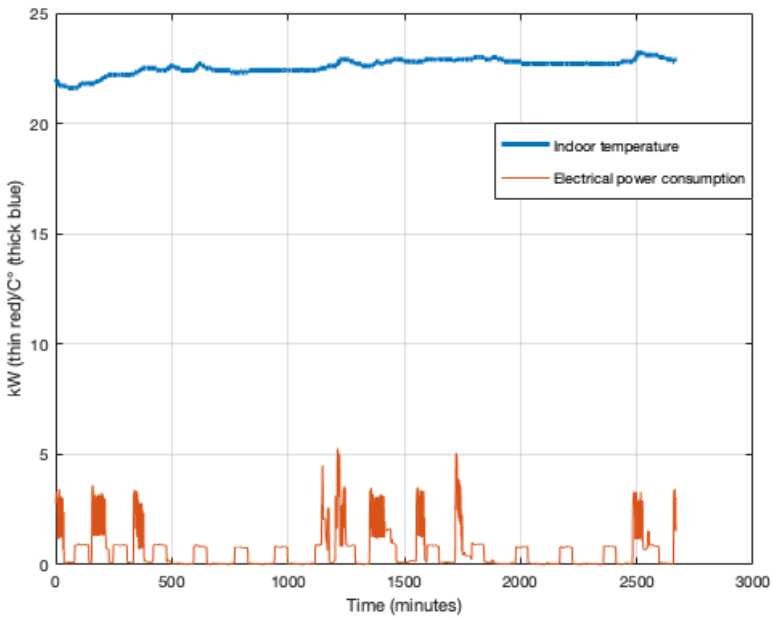

Figure 20. The indoor temperature and the household electricity consumption in the bathroom during three-day tests.

Pearson’s correlation coefficient and Spearman’s rank correlation coefficient do not indicate any strong linear correlation or order association (Table 4). In fact, the coefficients have opposite signs which can indicate that the correlation is weak or zero. However, the null hypothesis can be rejected, thereby a correlation cannot be rejected.

Table 4. For bathroom calculated values of Pearson's correlation coefficient p, Spearman's rank correlation coefficient ρ and its p-value.

p	-0.0158
ρ	0.0809
p-value	2.8624e-05

The time delay was estimated to 2 minutes. The best ARX-model performed a GOF of 69.61 percent (Table 5). The A-polynomial has an order of 4 and the B-polynomial an order of 2.

The AR-model performed a GOF of 66.99 percent making it slightly worse than the corresponding ARX-model. It has an A-polynomial of order 4. The performance of the tuned 15-step predictors was worse than for the ARX- and AR-models (Table 5). The ARX-predictor performed better than the AR-predictor and the tuned one-step AR-model.

Table 5. For the bathroom, performance of the tuned 15-step AR- and ARX-predictors and the tuned one-step AR-and ARX-models.

FIT_{GOF} *(ARX)*	69.61%
FIT_{GOF} *(AR)*	66.99%
FIT_{GOF} *(15-step ARX)*	67.10%
FIT_{GOF} *(15-step AR)*	64.18%

5.1.3 Bedroom

The temperature in the bedroom consistently rose during the measuring period (Figure 21). This room may have required a longer period from that of occupying the apartment until reaching a steady state indoor temperature. For the bedroom, no visual correlation is found between the household electricity consumption and the indoor temperature.

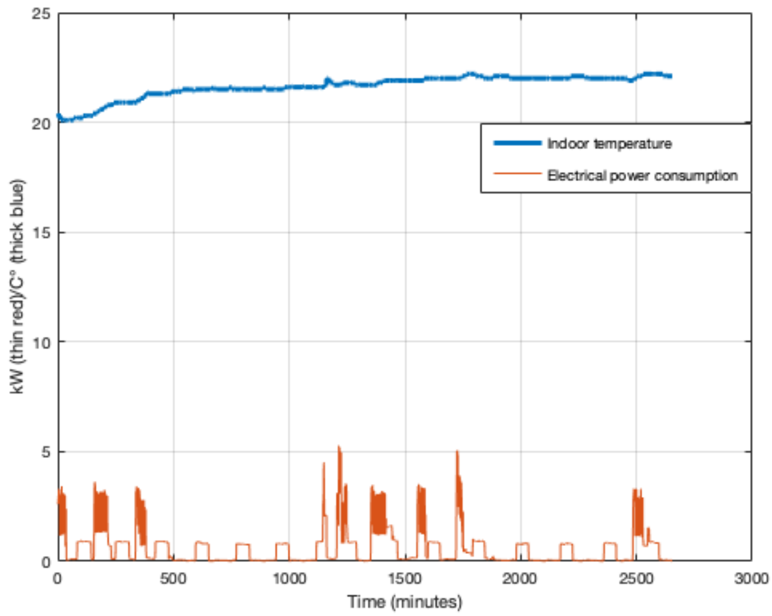

Figure 21. The indoor temperature and the household electricity consumption in the bedroom during three-day tests.

Pearson's correlation coefficient and Spearman's rank correlation coefficient show similar results. The null hypothesis can be rejected and therefore a correlation cannot be rejected (Table 6).

Table 6. For bedroom, calculated values of Pearson's correlation coefficient p, Spearman's rank correlation coefficient ρ and its p-value.

p	-0.1513
ρ	-0.1212
p-value	3.8117e-10

The time delay was estimated to 14 minutes. The selected ARX-model performs a GOF of 78.21 percent (Table 7). The A-polynomial and B-polynomial has an order of 4.

The AR-model of order 4 yielded a GOF of 75.94 percent (Table 7). The bedroom is the room with the most similar performance between the AR- and the ARX-model. The tuned 15-step predictors performed worse than the corresponding constructed ARX- and AR-models (Table 7). The 15-step ARX-predictor performed better than the AR-predictor and the tuned one-step AR-model regarding GOF.

Table 7. For the bedroom, performance of the tuned 15-step AR-and ARX-predictors and the tuned one-step AR-and ARX-models.

FIT_{GOF} (ARX)	78.21%
FIT_{GOF} (AR)	75.94%
FIT_{GOF} (15-step ARX)	76.87%
FIT_{GOF} (15-step AR)	74.00%

5.1.4 Livingroom

A visual correlation between the household electricity consumption and the indoor temperature of the living room can be observed (Figure 22). As the electrical power consumption increases so does the indoor temperature, with a time delay.

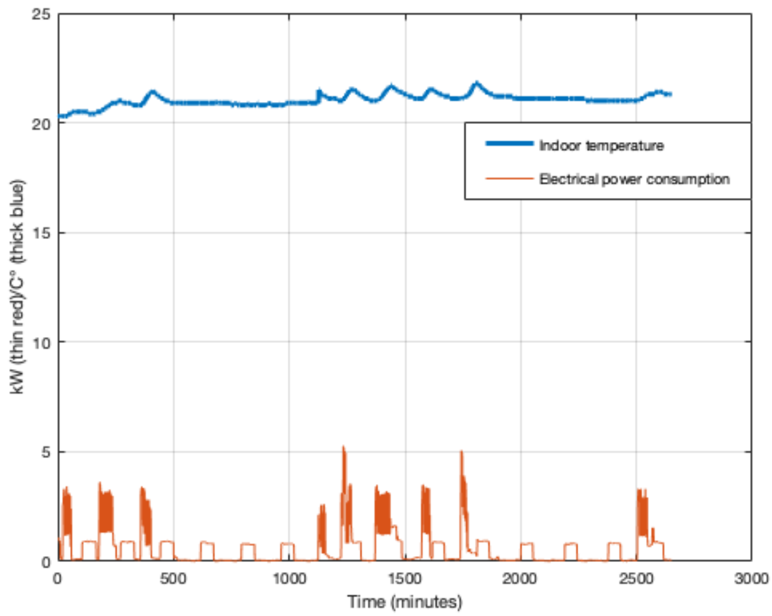

Figure 22. The indoor temperature and the household electricity consumption in the living room during three-day tests.

Pearson's correlation coefficient and Spearman's rank correlation coefficient were both positive (Table 8). The p-value differs from zero, i.e. the null hypothesis can be rejected.

Table 8. For living room calculated values of Pearson's correlation coefficient p, Spearman's rank correlation coefficient ρ and its p-value.

p	0.0542
ρ	0.2049
p-value	1.6059e-26

The best ARX-model with regards to the SE-fit was a model of order 2 for the A- and B-polynomial. This achieved a GOF of 66.52 percent (Table 9). The estimated time delay is 9 minutes.

The AR-model of order 2 performed about 4 percentage units worse than the ARX-model. The GOF of the AR-model was 62.27 percent (Table 9). For the living room, the tuned 15-step predictors performed worse than the corresponding constructed ARX- and AR-models (Table 9). Still, the tuned 15-step ARX predictor performed a better GOF than the tuned one-step AR-model. As for the other rooms, the 15-step ARX-predictor performs better than the AR-predictor.

Table 9. For the living room, performance of the tuned 15-step AR-and ARX-predictors and the tuned one-step AR-and ARX-models.

FIT_{GOF} *(ARX)*	66.52%
FIT_{GOF} *(AR)*	62.27%
FIT_{GOF} *(15-step ARX)*	62.75%
FIT_{GOF} *(15-step AR)*	56.84%

5.1.5 Middle room

The middle room data consists of about 1000 measurements less than the other rooms. This was due to a configuration failure for the first session of the three-day tests. At first sight, no obvious visual correlation seems to exist (Figure 23). Still, the number of peaks for the indoor temperature seems to correspond to the number of peaks for the electricity consumption. Also, during the longer period of rest for the electricity consumption, i.e. between minute 750 and 1300, the indoor temperature seems to slowly decrease. However, none of these observations can be considered conclusive.

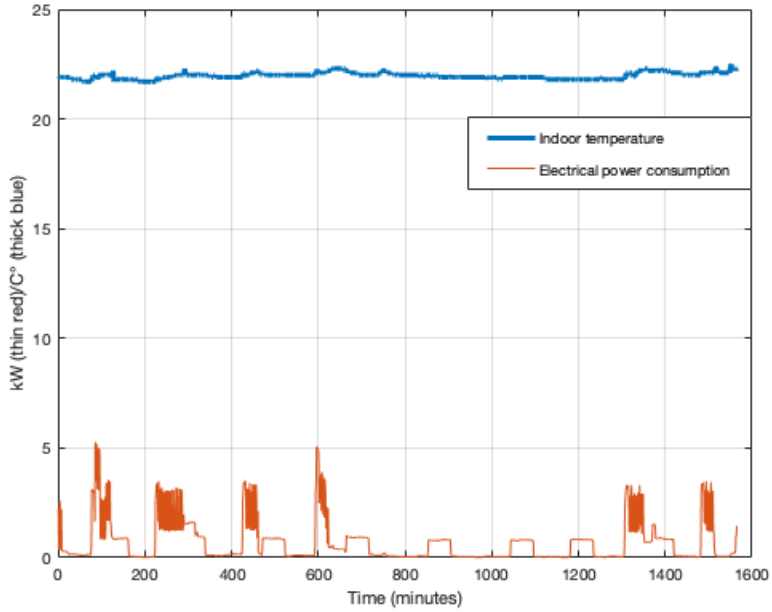

Figure 23. The indoor temperature and the household electricity consumption in the middle room during three-day tests.

Pearson's correlation coefficient and Spearman's rank correlation coefficient took on positive values (Table 10). The correlation coefficients has relatively large absolute values, compared to the rest of the rooms. The p-value lies within the significance level and therefore the null hypothesis can be rejected.

Table 10. For middle room calculated values of Pearson's correlation coefficient p, Spearman's rank correlation coefficient ρ and its p-value.

p	0.2378
ρ	0.3973
p-value	2.4141e-60

The best ARX-model with respect to the SE-fit has an order of 4 of the A- and B-polynomial. The time delay was estimated to 0 and the ARX-model achieved a GOF of 46.73 percent (Table 11).

The AR-model of order 4 also performs poorly. The GOF is only 41.24 percent, which is the worst result of all the rooms (Table 11). The lack of measurement data can the reason for this bad performance. The tuned 15-step predictors performed worse than the constructed ARX- and AR-models (Table 11).

Table 11. For the middle room, performance of the tuned 15-step AR-and ARX-predictors and the tuned one-step AR-and ARX-models.

FIT_{GOF} *(ARX)*	46.73%
FIT_{GOF} *(AR)*	41.24%
FIT_{GOF} *(15-step ARX)*	23.92%
FIT_{GOF} *(15-step AR)*	21.86%

5.2 Summary- Correlation analysis

For every room, the null hypothesis can be rejected. This means that a correlation cannot be disproven conclusively between the household electricity consumption and the indoor temperature. This does not mean that a strong correlation exists or even a positive one. Still, it means that a hypothesis of zero correlation is wrong within the significance level of 0.05 and that a relationship, direct or indirect, may be present. A visual correlation is observed for all rooms, except for the bedroom. It seems to be strongest for the kitchen, which is the room with the most heavy heat generating appliances.

For every room, the 15-step predictions from tuned one-step linear models improves when they consider the household electricity consumption, i.e. the ARX-models. The same is true for the tuned 15-step predictors. However, for the tuned one-step predictors, the estimated parameters of the B-polynomial are small and only affects the outcome of the predictions to a limited extent, which coincides with the fact that the input-output models only perform slightly better than the corresponding no-input models. For every room, the tuned one-step predictors, which use information about electricity consumption up to $t - 1$ to predict the indoor temperature at time t, outperforms the tuned 15-step predictors of the same structures.

The correlation analysis justifies some pre-assumed characteristics. The correlation seems to be time delayed and non-linear, meaning that no instant temperature changes occur in parallel with the change in electricity consumption. The absolute Pearson's and Spearman's correlation coefficients are for every room low and the scatter plots shows no visual relationship. Still the performances of the ARX-predictors are throughout the experiments better than the AR-predictors, but the differences are marginal. The signals also show tendencies of a delayed visual correlation for most of the rooms, except for the bedroom.

Regarding the performance of the models, it is important to have enough estimation data. Looking at the middle room models, the results diverge compared to the other rooms in terms of performance, likely because this room had about 1000 data points less to utilize in the estimation process than the other rooms.

The purpose of the correlation analysis was to study whether a correlation exists between the household electricity consumption and the indoor temperature, the results show that this is likely the case but to a limited extent. Therefore, the modeling of the two-week data builds on the assumption that a correlation cannot be rejected conclusively.

5.3 Two-week data- Linear modeling

5.3.1 Kitchen

The best performing model to represent the kitchen system is an ARX-model with a time delay of 8 minutes and the order of 3 and 1 for the A- and B-polynomial. The model GOF was only 0.01 percent worse than the best higher order model and performed equally good, or in some cases better, regarding the characteristics of the residuals, the poles and zeros and the step-and frequency responses. The ARX-model performed better than the best ARMAX-, BJ-and OE-models (Table 12).

Table 12. Checklist from the analyzation of the most suitable model of the kitchen. The best models with respect to GOF are shown together with the finally selected model ARX318.

Model	FIT_{GOF}	Residuals	Poles/zeros	Step and frequency response	Model parameters
ARX318	80.01%	X	√	√	√
ARMAX1228	79.25%	X	√	√	√
BJ31218	79.99%	X	√	√	X
OE248	10.19%	X	-	-	-

The behavior of the residuals was unsatisfying for all tested ARX-models (Figure 24). The autocorrelation of the residuals shows that up until the lag of |20|, correlation seem to exist. This means that there are still dependencies between the residuals and noise is not successfully regarded by the model. However, the cross-correlation between the residuals and the household electricity consumption is likely white noise, which is preferable.

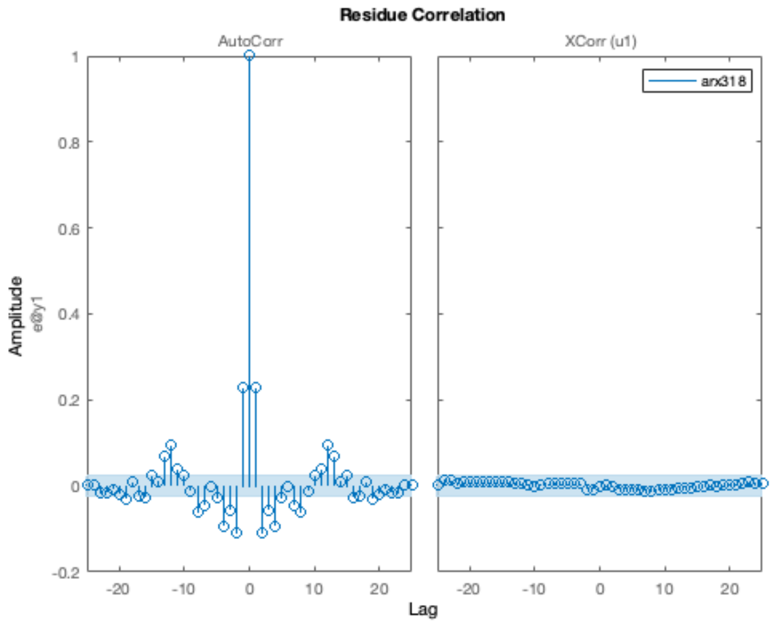

Figure 24. The autocorrelation of the residuals and the cross-correlation between the residuals and the household electricity consumption for the ARX-model of the kitchen.

Comparing the ARX-model with the best performing AR-and ARMA models, the performance of the 15-step predictions are slightly better for the ARX-model. The best performing AR-model coincide with validation data up to a degree of 79,80 percent regarding GOF (Table 13). The corresponding ARMA-model performed similar 15-step predictions (79.37 percent).

The tuned 15-step AR- and ARX predictors performed a worse GOF than the corresponding AR-and ARX models (Table 13).

Table 13. For the kitchen, the GOF performance of the selected input-output model, the best performing no-input model and the tuned 15-step ARX- and AR-predictors.

FIT_{GOF} *(ARX)*	80.01%
FIT_{GOF} *(AR)*	79.80%
FIT_{GOF} *(15-step ARX)*	78.24%
FIT_{GOF} *(15 step AR)*	77.79%

5.3.2 Bathroom

The best linear model to represent the bathroom is an ARX-model with $n_a = 4$, $n_b = 1$ and $n_k = 1$. It performed a GOF of 85.16 percent, which becomes better for higher order models but not significantly. The best model regarding GOF was a BJ-model but this model resulted in small parameters of the B-polynomial.

All models, except for the OE-model, performed well regarding GOF but performed worse regarding other characteristics (Table 14). For example, the autocorrelation of the residuals and the cross-correlation between the residuals and the input were not satisfying for neither model.

Table 14. Checklist from the analyzation of the most suitable model of the bathroom. The best models with respect to GOF are shown together with the finally selected model ARX411.

Model	FIT_{GOF}	Residuals	Poles/zeros	Step and frequency response	Model parameters
ARX411	85.16%	X	√	√	√
ARMAX4321	85.05%	X	√	√	*X*
BJ32441	85.19%	X	√	√	X
OE111	20.38%	X	-	-	-

As the estimated parameters of the B-polynomial of the BJ-model were small, a reduced model was tried. The only B-polynomial order resulting in acceptable parameters was a BJ-model with $n_b = 1$. However, this model resulted in a poorer performance of the 15-step predictions than the ARX-model. This was also the case for a reduced ARMAX-model, which performed similarly to the ARX-model for all criteria but the size of the parameters. The ARMAX-model of $n_a = 4$, $n_b = 1$, $n_c = 2$ and $n_k = 1$ also deviated from the rest of the models in its frequency response (Figure 25). The rest of the models behaved quite similarly for larger frequencies and followed the response estimated from the spectral analysis fairly well, but the reduced ARMAX-model did differ. Even though the behavior at low frequencies are more interesting for us, this divergent behavior did not exist for the original ARMAX-model and is not preferable.

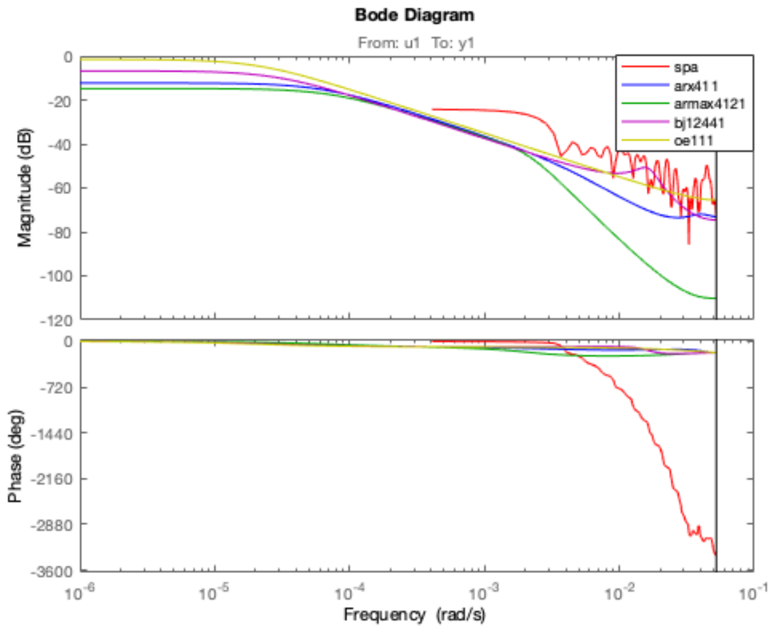

Figure 25. For the bathroom, a bode-diagram of the linear models and the frequency response estimated through spectral analysis (spa). The reduced BJ-and ARMAX-models are shown.

Comparing the ARX-model to AR-and ARMA-models, their performance on the 15-step predictions was slightly worse than for the ARX-model. The best performing AR-model, with a A-polynomial of order 4, managed to achieve a GOF of 85.08 percent (Table 15). For the best ARMA-model, this value was 84.83 percent.

Regarding the tuned 15-step predictors, their performances were worse than the performance of the ARX-and AR-models (Table 15). The 15-step ARX- and the AR-predictor received almost the same GOF, with an advantage for the ARX-predictor on the second decimal.

Table 15. For the bathroom, the GOF performance of the selected input-output model, the best performing no-input model and the tuned 15-step ARX- and AR-predictors.

FIT_{GOF} *(ARX)*	85.16%
FIT_{GOF} *(AR)*	85.08%
FIT_{GOF} *(15-step ARX)*	84.09%
FIT_{GOF} *(15-step AR)*	84.04%

5.3.3 Bedroom

The linear model best representing the bedroom is an ARX-model with $n_a = 3$, $n_b = 1$, and $n_k = 0$. Its GOF was 78.03 percent, which was beaten by a BJ-model of higher order but the BJ-model was not satisfying on other criteria.

The best linear models of each kind, with respect to GOF, had a similar performance of their 15-step predictions except for the OE-model. However, the behavior regarding the residuals, the location of the poles and zeros and the behavior of the step-and frequency responses did differ (Table 16). For example, the BJ-model had poles and zeros placed close to each other, its step-and frequency response was unsatisfying and the estimated parameters of the B-polynomial was smaller than 0.001 for all parameters except one.

Table 16. Checklist from the analyzation of the most suitable model of the bedroom. The best models with respect to GOF are shown together with the final selected model ARX310.

Model	FIT_{GOF}	Residuals	Poles/zeros	Step and frequency response	Model parameters
ARX310	78.03%	X	√	√	√
ARMAX1120	77.84%	X	√	√	√
BJ44140	78.10%	X	X	X	X
OE430	5.661%	X	-	-	-

A reduced BJ-model was also tested and performed worse results than the ARX-model, therefore it was also rejected. The step response from the BJ-model became oscillating (Figure 26), which is not likely the real case for the indoor temperature when the household electricity consumption is turned on to a constant value from zero. A reason to this behavior could be that the poles of the model lies close to the borders of the unit circle.

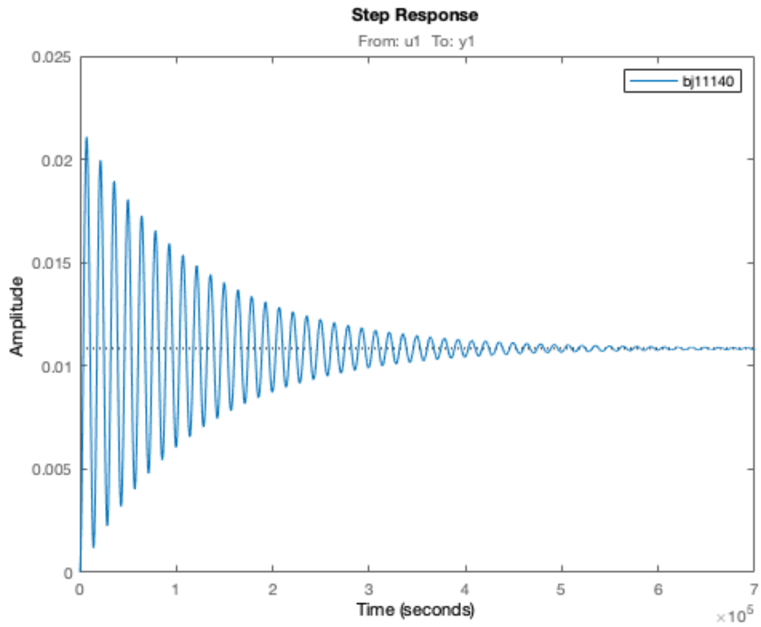

Figure 26. The step-response for the BJ-model of the bedroom.

The 15-step predictions of the best AR-and ARMA-models performed similar results as the ARX-model. In fact, the AR-model had a GOF of 78.04 percent which is slightly better than the corresponding ARX-model result (Table 17). This AR-model had the order of 3 of the A-polynomial. The 15-step predictions of the ARMA-model, with an order of 1 and 4 on the A-and C-polynomial, had a GOF of 77.97 percent.

The GOF received from the 15-step predictors were slightly worse than for the 15-step predictions of the ARX-and AR-models (Table 17). However, the 15-step ARX-predictor performed marginally better than the corresponding AR-predictor which is the other way around when comparing the performance of the AR-and the ARX-model.

Table 17. For the bedroom, the GOF performance of the selected input-output model, the best performing no-input model and the tuned 15-step ARX- and AR-predictors.

FIT_{GOF} *(ARX)*	78.03%
FIT_{GOF} *(AR)*	78.04%
FIT_{GOF} *(15-step ARX)*	76.75%
FIT_{GOF} *(15-step AR)*	76.71%

5.3.4 Livingroom

An ARMAX-model of $n_a = 3, n_b = 1, n_c = 1$ and $n_k = 23$ was selected to represent the system of the living room. In fact, this was the model that performed best with respect to the GOF. Lower orders of ARX-and BJ-models were tested but resulted in a worse GOF than the ARMAX-model. The residuals of all models had unsatisfying behaviors (Table 18).

Table 18. Checklist from the analyzation of the most suitable model of the living room. The best models with respect to GOF are shown together with the final selected model ARMAX31123.

Model	FIT_{GOF}	Residuals	Poles/zeros	Step and frequency response	Model parameters
ARX4423	89.92%	*X*	√	√	*X*
ARMAX31123	91.54%	*X*	√	√	√
BJ314423	91.47%	X	X	√	X
OE3423	15.74%	X	-	-	-

Comparing the ARMAX-structure with AR- and ARMA-models, the GOF was better for ARMAX (Table 19). The AR-model achieved a GOF of 89.61 percent and the ARMA-model achieved a GOF of 91.44 percent. The tuned 15-step predictors where slightly worse and performed a GOF of around 90 percent for both the ARX-and the AR-predictor (Table 19). However, the 15-step ARX-predictor generated marginally better results compared to the 15-step AR-predictor.

Table 19. For the living room, the GOF performance of the selected input-output model, the best performing no-input model and the tuned 15-step ARX- and AR-predictors.

FIT_{GOF} *(ARMAX)*	91.54%
FIT_{GOF} *(ARMA)*	91.44%
FIT_{GOF} *(15-step ARX)*	89.86%
FIT_{GOF} *(15-step AR)*	89.64%

5.3.5 Middle room

The best model to represent the middle room was an ARMAX-model of $n_a = 4$, $n_b = 1$, $n_c = 2$ and $n_k = 8$. An ARMAX-model of order 2 of the B-polynomial achieved a negligibly better GOF but this model resulted in small parameters of the B-polynomial and was therefore rejected. The selected ARMAX-model was the first model to behave satisfying on each analyzed area and was the first room to generate models with desirable residual characteristics (Table 20). However, the GOF was poor for each model.

As for the rest of the rooms, the OE-model performed poorly both regarding GOF and regarding the residual characteristics. Therefore, this model was not further analyzed.

Table 20. Checklist from the analyzation of the most suitable model of the middle room. The best models with respect to GOF are shown together with the final selected model ARMAX4128.

Model	FIT_{GOF}	Residuals	Poles/zeros	Step and frequency response	Model parameters
ARX428	67.64%	*X*	√	√	*X*
ARMAX4128	68.45%	√	√	√	√
BJ31228	68.47%	√	√	√	*X*
OE138	11.74%	X	-	-	-

The best BJ-model regarding GOF resulted in better 15-step predictions than the selected ARMAX-model, this was also true for another reduced BJ-model. However, the best BJ-model had small B-polynomial parameters and the reduced BJ-model resulted in a somehow questionable behavior of its step response, see Figure 27 and 28. The value the output converges towards do change between different orders of the BJ-model, from around 1 for the original model to 0.8 for the reduced one. This behavior does not apply for the ARMAX-model were the value is 0.6 for both the original and the reduced model. Due to this, and the more complex character of a BJ-model, the ARMAX-model was selected to represent the system of the middle room.

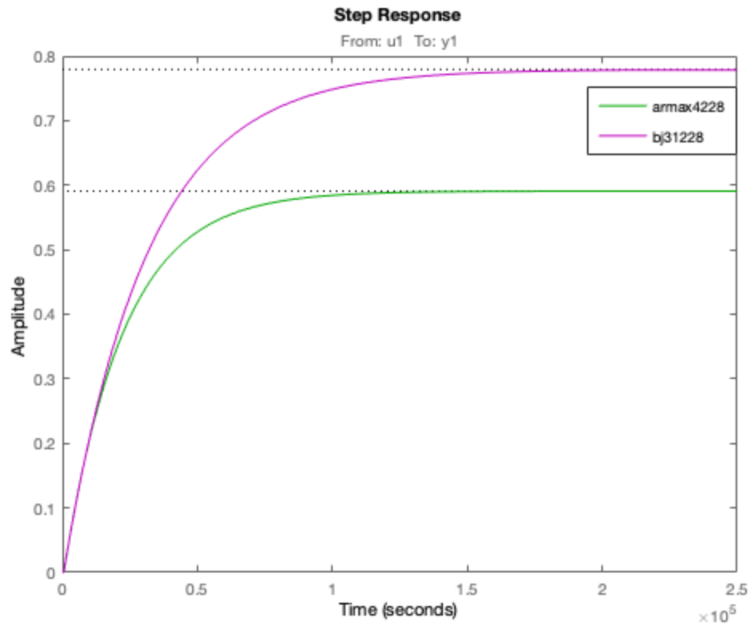

Figure 27. The step-response for the ARMAX-and BJ-models with best performance regarding GOF, i.e. the original models.

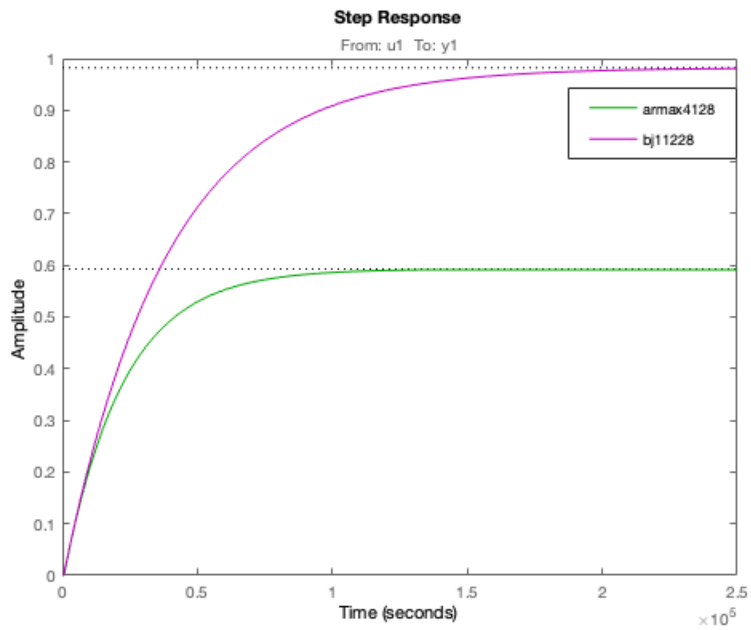

Figure 28. The step-response for the reduced ARMAX-and BJ-models.

The AR-and ARMA models performed a worse GOF than the ARMAX-model. The AR-model performed a GOF of 66.99 percent and the ARMA-model a GOF of 67.75 percent. The tuned 15-step predictors generated similar GOF-values, where the ARX-predictor performed best (Table 21).

Table 21. For the middle room, the GOF performance of the selected input-output model, the best performing no-input model and the tuned 15-step ARX- and AR-predictors.

FIT_{GOF} *(ARMAX)*	68.45%
FIT_{GOF} *(ARMA)*	67.75%
FIT_{GOF} *(15-step ARX)*	66.06%
FIT_{GOF} *(15-step AR)*	65.02%

5.4 Two-week data- Non-linear modeling

5.4.1 NLARX-and NLAR-models

For most of the rooms, the NLARX- and NLAR-models performed similarly to the best linear models when estimated and validated on the same sets of data (Table 22). For the kitchen and the bedroom, the best performing tuned one-step predictor was non-linear. For the rest of the rooms, the model with the best performance was linear. However, the difference in performance between the linear and non-linear models are small.

For the bathroom, bedroom and the living room, the NLAR models performed slightly better than the NLARX models. The perks of knowing the previous and present household electricity consumption to predict future indoor temperatures is therefore questionable.

When compared to the previous tuned linear 15-step predictors, all NLARX-and NLAR-models perform worse regarding GOF, except for the bedroom.

Table 22. For every room, the layer shape of the sigmoid network together with the structure and the GOF for the selected NLARX and NLAR models.

Room	Nr.Units (Sigmoidnet, NLARX)	Structure (NLARX) $[n_a n_b n_k]$	FIT_{GOF} (NLARX)	Nr.Units (Sigmoidnet, NLAR)	Structure (NLAR) $[n_a]$	FIT_{GOF} (NLAR)
Kitchen	[7]	[3 1 8]	82.59%	[5]	[3]	80.81%
Bathroom	[11]	[4 1 1]	84.67%	[3]	[4]	84.96%
Bedroom	[3]	[3 1 0]	78.73%	[10]	[3]	79.02%
Living room	[2]	[3 1 23]	89.48%	[4]	[4]	89.59%
Middle room	[1]	[4 1 8]	68.36%	[17]	[4]	67.56%

5.4.2 15-step LAVA-predictors

The tuned 15-step LAVA-predictors generates for all rooms generally less accurate predictions than the linear models and the neural network models (Table 23). The LAVA- ARX- and AR-structures generates results with similar GOF:s for each room. Small performance improvements are found for the ARX-models compared to the AR-models regarding kitchen, living room and middle room but the differences are small. The bedroom and bathroom AR-models are better than their ARX counterparts. The electricity consumption seems to provide little additional information for the estimated LAVA-models.

Table 23. The structures of LAVA- AR- and ARX-models for every room and their 15-step predicted response performance.

Room	[M, L] (ARX)	Structure (ARX) $[n_a n_b]$	FIT_{GOF} (ARX)	[M, L] (AR)	Structure (AR) $[n_a]$	FIT_{GOF} (AR)
Kitchen	[3, 5]	[3 1]	76.80%	[3, 5]	[3]	76.39%
Bathroom	[3, 5]	[4 1]	75.57%	[3, 5]	[4]	76.62%
Bedroom	[4, 6]	[3 1]	76.27%	[4, 6]	[3]	76.29%
Living room	[3, 5]	[3 1]	86.55%	[3, 5]	[3]	86.32%
Middle room	[3, 5]	[4 1]	64.36%	[3, 5]	[4]	64.35%

5.5 Sensitivity analysis

When comparing the impact of household electricity consumption on the indoor temperature of the kitchen with the impact from relative humidity and radiator temperature, interesting results were found. The OE-models of the signals varied significantly in their ability to predict the indoor temperature. The constructed OE-model with the radiator temperature as input achieved a GOF of 42.98 percent, which was the best received result. The corresponding GOF from the OE-model with input electricity achieved a 10.19 percent GOF and the OE-model with input relative humidity performed a GOF of 8.37 percent (Table 24).

The single-input ARX-models performed similar GOF:s compared to each other (Table 24). However, the ARX-model with household electricity as an input performed best. Second best performance was achieved by using radiator temperature as an input, closely followed by the ARX-model which used relative humidity.

The MISO ARX-model modelled by all three signals achieved a GOF of 80.28 percent, which is slightly better than for the original ARX-model of the kitchen which only regarded the household electricity consumption as an input. A MISO ARX-model with the inputs household electricity and radiator temperature performs equally well (Table 24).

Table 24. GOF:s from OE, ARX-and MISO ARX-models with inputs household electricity (el), relative humidity (hum) and radiator temperature (rad).

Model type	*Inputs*	FIT_{GOF}
OE	el	10.19%
OE	hum	8.37%
OE	rad	42.98%
ARX	el	80.02%
ARX	hum	79.90%
ARX	rad	79.98%
ARX (MISO)	el, hum, rad	80.29%
ARX (MISO)	el, rad	80.29%

5.6 Summary- Two-week study

The results of this study shows that the tuned one-step predictors, which uses household electricity consumption up to $t-1$ to predict outputs of indoor temperature at time t, outperforms the tuned 15-step predictors for every room except the living room. For the living room, the tuned linear 15-step predictors perform a better GOF than the tuned one-step NLARX-and NLAR models. However, the best performing models of the living room are still tuned linear one-step predictors.

For the kitchen and the bedroom, the best performing models are non-linear. This is not true for the rest of the rooms and may be explained by the characteristics of the rooms. For the bathroom, the difference between the best performing model and the worst is bigger than for the rest of the rooms, i.e. 9.59 percentage units. This difference is smallest for the bedroom where the best performing model achieves a GOF of 2.75 percentage units better than the worst performing one. For the bedroom, the best performing model is a no-input model, in contrast to the best performing models for the rest of the rooms.

The OE-models performs poor GOF results for all rooms, which is to be expected as they lack access to earlier temperature data points upon prediction. The household electricity consumption is, by itself, thereby a weak temperature prediction parameter.

For each room, the overall performance of the models are similar with two exceptions. When comparing each ranking with the corresponding GOF for each room the models of the middle room stands out as having more trouble mimicking the validation data (Table 25). In contrast to this, the living room stands out by having a better performance than the rest of the rooms on every ranking. Commonly for all room is the marginal difference between the performance of an input-output model of a specific kind and the corresponding no-input model of the same kind. The LAVA-models performs the worst GOF.

Table 25. The best performing model types for each room, ranked by GOF in descending order.

Ranking	*Kitchen* $[\boldsymbol{FIT_{GOF}}]$	*Bathroom* $[\boldsymbol{FIT_{GOF}}]$	*Bedroom* $[\boldsymbol{FIT_{GOF}}]$	*Living room* $[\boldsymbol{FIT_{GOF}}]$	*Middle room* $[\boldsymbol{FIT_{GOF}}]$
1	NLARX **[82.59%]**	ARX **[85.16%]**	NLAR **[79.02%]**	ARMAX **[91.54%]**	ARMAX **[68.45%]**
2	NLAR **[80.81%]**	AR **[85.08%]**	NLARX **[78.73%]**	ARMA **[91.44%]**	NLARX **[68.36%]**
3	ARX **[80.01%]**	NLAR **[84.96%]**	AR **[78.04%]**	15-step ARX **[89.86%]**	ARMA **[67.75%]**
4	AR **[79.80%]**	NLARX **[84.67%]**	ARX **[78.03%]**	15-step AR **[89.64%]**	NLAR **[67.56%]**
5	15-step ARX **[78.24%]**	15-step ARX **[84.09%]**	15-step ARX **[76.75%]**	NLAR **[89.59%]**	15-step ARX **[66.06%]**
6	15-step AR **[77.79%]**	15-step AR **[84.04%]**	15-step AR **[76.71%]**	NLARX **[89.48%]**	15-step AR **[65.02%]**
7	15-step LAVA-ARX **[76.80%]**	15-step LAVA-AR **[76.62%]**	15-step LAVA-AR **[76.29%]**	15-step LAVA-ARX **[86.55%]**	15-step LAVA-ARX **[64.36%]**
8	15-step LAVA-AR **[76.39%]**	15-step LAVA-ARX **[75.57%]**	15-step LAVA-ARX **[76.27%]**	15-step LAVA AR **[86.32%]**	15-step LAVA-AR **[64.35%]**

In the sensitivity analysis, the best performing OE-model used radiator temperature as the input signal. This model was significantly better than the other OE-models. However, the best performing single input ARX-model used household electricity as input, even if all ARX-models performed similar results. The models with the relative humidity as input signal did, for both the OE- and ARX-models, perform a worse GOF than the others. This signal was also redundant for the predictive performance of the MISO ARX-models, were an equally good result regarding GOF was achieved without information about relative humidity.

6. Discussion

6.1 Correlation analysis

The mathematical correlation between household electricity consumption and indoor temperature is small. There are several potential reasons for this. One being that the electricity consumption simply has a low impact on indoor temperature. This is intuitively reasonable for rooms such as the bedroom as the electrical devices present yields a marginal heat output and the room is populated primarily during the night, when the electricity usage is low. Therefore, a correlation in the kitchen seemed more plausible given its multiple heat generating devices such as oven, stove and dishwasher. The mathematical results showed otherwise though, which may be due to the used mathematical methods, i.e. Pearson's and Spearman's methods, and their limitations. It is precarious to assign too much meaning to the mathematical correlations as the methods are constructed to find linear, non-time delayed variable relationships, which should not be expected to be found in the three-day data.

Even though the mathematical correlations in the respective rooms were indeed small, visual correlations seemed to be present upon inspection of the variable behaviors for most rooms, except the bedroom. An initial hypothesis was that if electricity is used, this yields a heat output and may also imply that people are present in the apartment. As the indoor temperature seemed to rise during heavy use of electricity for most rooms, it is not too farfetched to conclude that the activities within the apartment (use of electricity, presence of humans, physical activity etc.) do impact the indoor temperature. Disturbances such as outdoor temperature, solar insolation, wind and humidity can of course distort the results but some pattern between electricity consumption and indoor temperature was present.

Another important correlation test was the performance comparison between AR- and ARX-models. A significant improvement in the ARX-models would imply that the electricity data provides useful information. Indeed, the ARX-models outperformed their AR counterparts, both regarding the tuned one-step predictors and the tuned 15-step predictors. It should however be pointed out that the gain was typically marginal. Moreover, the tuned one-step predictors performed better than the corresponding tuned 15-step predictors with the same polynomial structure, which may indicate that useful information exists in the household electricity consumption 15 minutes before prediction instance t.

From the results, we acknowledge a correlation between household electricity consumption and indoor temperature. However, the magnitude and the characteristics of the correlation differs between the rooms. A proof of this is the different time delays estimated for each room. Also, the difference between the input-output models and their corresponding no-input models were small, meaning that good results were achieved without using household electricity consumption.

The visual correlation was strongest for the kitchen, of which room the best performing model was constructed, i.e. the tuned one-step ARX-model. This model outperformed its AR-counterpart with about 2 percentage units, indicating that valuable information exists in using the electricity consumption, even if the difference is marginal.

For the living room, the difference in performance between the input-output models and the no-input models were greater than for the rest of the rooms, which attributes some benefits to knowing the household electricity consumption. However, the visual correlation was weaker compared to the kitchen. An explanation to this can be the behavior of the indoor temperature. For the kitchen, the temperature was of a more oscillating character compared to the living room. This can be explained by the fact that the kitchen has a lot of heat generating appliances affecting the indoor temperature during times when household electricity is used in the apartment. For the living room, it is reasonable to assume that household electricity consumption affects the indoor temperature marginally, due to the lack of heat generating electrical appliances in it. An increase of indoor temperature is rather caused by people present, outdoor temperature, radiator heat and heat leakage from other rooms.

The bathroom has, similarly to the kitchen, heavy heat generating machinery. The electrical components which yields a heat contribution is the dryer and the washing machine. Their respective heat output is however dependent on their energy effectiveness. A newer machine will likely yield less energy waste than an old one, this can explain the marginal effect of electricity on indoor temperature in the bathroom as both machines are fairly modern. The human presence in the bathroom can possibly be higher when household electricity is used but it is reasonable to assume that these visits are time limited and does not affect the indoor temperature significantly, resulting in a relatively constant temperature. The same assumption may hold for the warm water usage in the bathroom.

The bedroom showed no visual correlation in comparison to the other rooms. This may indicate that the household electricity consumption is redundant for predictions in this room, which could be reasonable. There are not many electrical components in the bedroom generating heat and it is likely that people are present when electrical usage is low, i.e. during the night. Still, when people are home and awake and uses electricity, the temperature may rise periodically by heat transfers from other rooms and sporadic human presence.

The middle room had the least reliable measurement data. Partly as a result of the configuration failure making the amount of data less than for the other rooms but also due to the sensor likely being defective, which was discovered after the two-week measurements. Even if so, a visual correlation seem to exist between electricity consumption and indoor temperature, and again the input-output models outperform the no-input models.

6.2 Two-week study

This study has shown that good performing models representing the thermal systems of the rooms of an apartment can be constructed to predict 15 minutes ahead from two weeks of minute-wise samples of indoor temperature and household electricity consumption. The indoor temperature itself is sufficient to achieve good results but by including household electricity consumption as an input it is possible to improve the results depending on room and model type.

For every room, there are small differences in performance between a specific type of input-output model and its corresponding no-input model. More or less equally good predictions of the indoor temperature are obtained using only information about the indoor temperature. For the kitchen, the input-output models consistently outperformed the no-input models and the difference of the best NLARX-model and the NLAR-model was the most significant, in relation to the other rooms. The household electricity consumption seem to have a stronger relation to the indoor temperature in the kitchen than for the other rooms.

The thermal systems studied can be considered non-linear given the complex dynamics observed. With this in mind, a reasonable assumption is that non-linear models captures the dynamics of the apartment better than linear models. Seemingly this was not true in this study, as the non-linear models rarely performed better. The linear models proved competitive and was often better in predicting the indoor temperature which is beneficial in terms of simplicity and compatibility with physical models. Only the best performing model of the kitchen, the NLARX-model, stands out a bit as it outperformed the best linear kitchen model GOF with 2.58 percentage units. Why the NLARX-model works best for the kitchen is not entirely easy to explain. One reason can be the number of electrical appliances. Compared to the other rooms, the kitchen contains a wide range of machinery generating different amounts of heat which can impact the indoor temperature. Also, some components vary in their heat generation over time depending on usage. A good example of this is the oven as, trivially, a lower oven temperature generates a lower heat contribution to the indoor temperature. This may give rise to non-linear patterns in the data as a specific level of household electricity consumption will not deterministically result in a linear influence on the indoor temperature.

Another reason is that some kitchen appliances are used during limited times of the day, which also coincide with the presence of people. It is possible that the different combinations of electrical appliance operating, and human activity, together with other disturbances like solar insolation, makes it hard to approximate a relationship in linear terms. With this in mind, one can assume that the relationship between household electricity consumption and indoor temperature in the kitchen is of a non-linear character. However, it is also possible that better non-linear models could have been modelled for the other rooms if the data was used differently or if more data was collected. But on the other hand, the constructed models do overall perform well, meaning that the amount of data collected can be considered sufficient to generate both good linear- and non-linear models.

For all rooms, the best performing models were tuned one-step predictors which upon validation uses inputs up to $t-1$ and outputs up to $t-15$, i.e. more input data than 15-step predictors. While the difference in performance cannot be considered great (about 1 percentage unit), it is possible that some useful information is contained in the electricity data. However, the best model for the bedroom was a no-input model which excludes the electricity. This can be due to correlation inconsistencies between indoor temperature and electricity consumption in the bedroom data. What can be deduced is that the household electricity consumption is not suitable for predicting the indoor temperature of the bedroom, which is reasonable considering the lack of electrical appliances and the daily patterns of human presence.

The living room is the best modelled room in terms of GOF. The reasons can be many but one explanation is that both the estimation- and validation data contained similar patterns for the indoor temperature data, making the models fit well on validation. This highlights the importance of informative data, as the middle room did not provide data to successfully model the "true" thermal dynamics. The living room is also one of the rooms with the least exposure to solar insolation, which was one of the major sources of measurement disturbance. In contrast, both the bedroom and kitchen are oriented south which added noise to the measurements a few hours each sunny day. This was observed in the data before pre-processing as temporary non-natural peaks seemed to rise instantly, probably because the sensors interpreted the solar insolation as a general rise of temperature in the rooms. As the living room and its sensor was not exposed to the sun, it was probably able to measure the indoor temperature more accurately.

The performance of all models can partly be attributed to each room's character in terms of exposure to disturbances and room size. The disturbances affecting the indoor temperature vary in number and tangibility between the rooms. For instance, the kitchen is primarily exposed to radiator heat, solar insolation, outdoor temperature and heat from both electrical appliances and humans. It is also a small room, potentially making it more sensitive to disturbances than bigger rooms. The living room is in contrast significantly bigger with the main sources of disturbance being radiator heat, outdoor temperature and the presence of humans. One can argue that the size of the kitchen and some of its disturbances contributes to more noisy temperature data, compared to the living room. This can explain why the living room temperature remains less noisy throughout the measuring period (Figure 29). The indoor temperature of the living room may therefore consist of more predictable fluctuations, yielding better GOF on each ranking compared to the other rooms (Table 25). Also, each linear model performed better than the corresponding non-linear model for this room. It is reasonable to assume that the indoor temperature follows a more linear pattern compared to the other rooms.

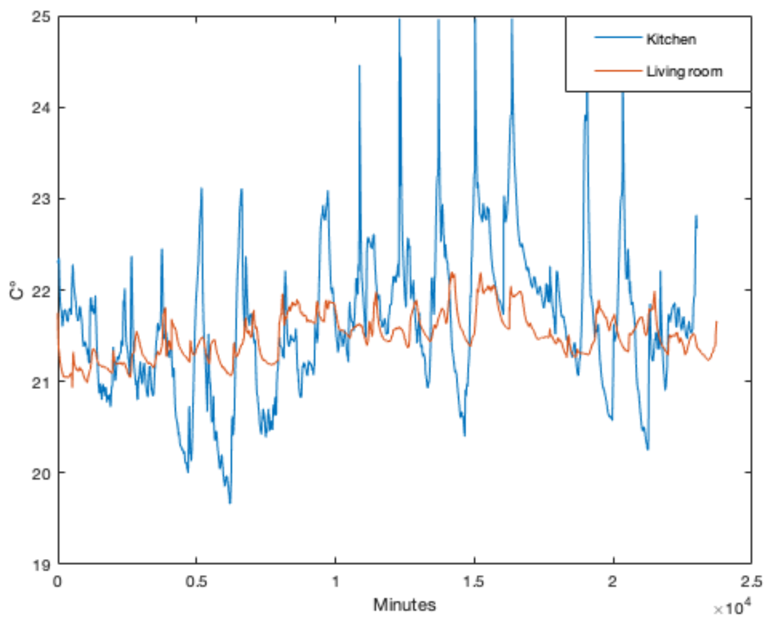

Figure 29. Comparison between the two-week measured indoor temperature of the kitchen and the living room.

The models constructed for the middle room generated predictions with the worst GOF. On every ranking the models lacked the rest of the rooms (Table 25). This may be explained by the size and the exposure to disturbances. The middle room is small and is also branched to the rest of the rooms, enabling heat transfers to and from these locations. Therefore, the middle room may be more sensitive and depend heavier on disturbances than the other rooms. However, the middle room neither contain any considerable heat generating appliances and it is not likely that people spend time in it for longer periods as it works more as a hallway to the rest of the rooms. The modeling thus resulted in more challenges than the other rooms, indicating that the indoor temperature is hard to predict. The indoor temperature data of the middle room is even more noisy than the kitchen data (Figure 30).

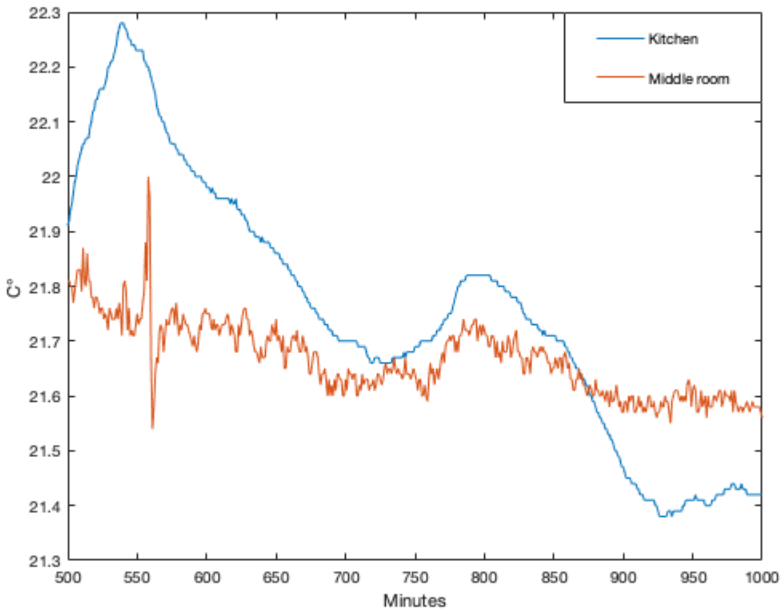

Figure 30. Comparison between the two-week measured indoor temperature of the kitchen and the middle room between minute 500 and 1000.

A possible explanation to the behavior of the indoor temperature of the middle room is a defective Tinytag sensor, resulting in noisy measurements. It is hard to believe that the temperature would have behaved as the measurement data shows as indoor temperature should not fluctuate this fast. All rooms used the same sample time and the sensors were placed on locations similar for all rooms. Therefore, the results of the middle room are likely not accurate.

Overall, the tuned 15-step predictors performed a worse GOF than the equivalent tuned one-step predictors. However, it is not possible to deduce that this difference depends on the usage of the input signal. The idea was that if input-output tuned one-step predictors, using inputs up to $t-1$, performs better than input-output tuned 15-step predictors using inputs up to $t-15$ to predict outputs at time t, then important information exists in the input data the last 15 minutes before prediction. In relation to this, it is easy to overestimate the benefits of using the household electricity consumption when evaluating the results. For instance, regarding the linear input-output models of the kitchen, the tuned one-step ARX-model outperformed the tuned 15-step ARX-predictor with 1.77 percentage units. From this, it would be easy to assume that more information about the input signal results in a better prediction performance.

However, the corresponding linear no-input models of the kitchen has a bigger difference in prediction performance than the input-output models. The tuned one-step AR-model outperforms the tuned 15-step AR-predictor with 2.01 percentage units, a quite similar and even larger difference than for the input-output models. For the other rooms, the differences between no-input and input-output models are also similar. This means that the difference in performance between tuned one-step and tuned 15-step predictors do not necessarily depend on the usage of the household electricity consumption. Instead, one explanation can be the selected structures of the tuned 15-step predictors. They were selected to be the same as for the best performing tuned one-step predictors. In other words, the structures of the tuned 15-step predictors were not optimized.

The purpose of this study was not to find the best models but instead to evaluate the benefits of using the household electricity consumption. If the household electricity consumption had brought more valuable information for predicting the indoor temperature than the results of this study indicates, then a bigger difference would have been observed between the performances of the input-output- and no-input models. Also, if valuable information had existed in the last 15 minutes before prediction, the difference in performance between input-output models (tuned one-step predictors compared to tuned 15-step predictors) would have been smaller than the corresponding difference between the no-input models.

For every room, the non-linear tuned 15-step LAVA-predictors performed worse than the other models. The reason for this is the same as for why the tuned one-step predictors generally outperformed the tuned 15-step predictors, which is the selected structures of the models. For the LAVA-models, the structures were chosen to have the same order as their linear equivalents and better results would have been achieved by optimizing the polynomial orders of the LAVA-models. It was found that, by changing the polynomial orders for the model of the living room, the performance increased about 1 percentage unit per A-polynomial order increase. By using $n_a = 7$ for the LAVA-AR model, a GOF of 90.50 percent was achieved, making it the third best performing model of this room, but at the cost of complexity. The linear tuned one-step predictors were only tested on orders up to four, as higher orders were deemed redundant from the correlation analysis.

Finally, the sensitivity analysis on the kitchen brought interesting results to the study. It was not surprising that the input of the OE-model performing the best GOF was the radiator temperature. This is reasonable as the radiator surface temperature directly affects the indoor temperature in the kitchen, with some time delay. The radiators also operates as the main heating source in the apartment which entails that their surface temperatures are correlated with the indoor temperature. Therefore, it is somewhat surprising that the ARX-model that performed the best results in regard to GOF (by a small margin), used household electricity as input. This gives some merit to the hypothesis that the household electricity consumption, along with the trends following it, is correlated with the indoor temperature and can help predict it, at least for the kitchen.

7. Conclusion

The results of this study show that for the rooms of the apartment, the benefits of knowing the household electricity consumption when predicting indoor temperature are marginal, but nevertheless existing for all rooms except the bedroom. For the kitchen, the knowledge about household electricity consumption seems more valuable than for the other rooms, even if it does not result in any significantly better predictions compared to the no-input models. Regarding the bedroom, the electricity consumption reduces the prediction performance.

The correlation between household electricity consumption and indoor temperature is strongest for the kitchen. A weaker correlation exists for the living room and the bathroom and no significant correlation can be found in the bedroom. The data of the middle room were remarkably noisy and therefore it is hard to evaluate the correlation for this room.

Non-linear models yields no considerable improvements in prediction performance compared to the corresponding linear models. As the linear models are less complex and more compatible with physical models, they are considered favorable for this modeling purpose. However, it is possible to construct better performing non-linear models by optimizing their structures or by using other methods and data.

Tuned one-step-predictors performs better than tuned 15-step-predictors, when validated on a prediction horizon of 15 minutes. However, the difference does likely not depend on the additional information received from the input signal during validation, but instead the selected structures of the models. The information about the household electricity consumption 15 minutes before prediction can be considered superfluous.

8. Future studies

The results of this study are not generalizable. The location of the apartment, the measuring period, the housing conditions and the building characteristics all bias the results. However, some aspects of the relationship between household electricity consumption and indoor temperature may, in future studies, show to be common for Swedish apartments. For instance, it is likely that the correlation is strongest for the kitchen, both for smaller and larger apartments. It would therefore be interesting to observe this relationship for apartments with only one or two rooms. For these apartments, it is common that the kitchen and the living room coincide to one living space. It is possible that the relationship is stronger for these apartments, as all electrical appliances are located in a limited space. When constructing empirical models of these thermal systems, one can deduce where heat is generated when the appliances are being used, in contrast to larger apartments where heat can be generated from electrical appliances in multiple rooms. Also, the size of the apartments entails that when electrical appliances are used, the only room people can occupy is the bathroom or the living space. This likely makes it easier to construct good models as the household electricity consumption can affect the indoor temperature more than for larger apartments.

In the central parts of Stockholm, about 60 percent of the apartments are one- or two-room apartments of older building types (SCB, 2016). In this study, the best input-output model of the kitchen did perform 1.78 percentage units better than its corresponding no-input model. If one can assume that this holds, or that this difference even increases for smaller apartments, there may exist a potential in regarding this relationship when shaping strategies for reducing energy demands and improving energy efficiency. We therefore recommend further studies, similar to ours, to be carried out on smaller apartments.

Today, smart radiator thermostats exists on the market. Some of these technologies considers inputs and disturbances such as indoor temperature, weather, the relative indoor humidity and the residents relative locations (tracking them by mobile phones) to control the indoor temperature towards a dynamic target temperature, which changes depending on the level of occupancy (TADO, 2020). An additional parameter to include in these techniques may be the household electricity consumption, possibly implementing a more proactive approach to the control, suitable for smaller apartments or kitchens.

In this study, we have showed how to construct well-performing predictors without an expensive budget. They were able to predict the indoor temperature 15 minutes in advance with satisfying results from one week of estimation data. We also showed that the household electricity consumption, as an input signal to an ARX-model of the kitchen, resulted in better prediction performance than the relative humidity and that it is possible to receive even better performances by also including the radiator temperatures. In line with these results, further studies are recommended to evaluate if the correlation between household electricity consumption and indoor temperature, can be used in radiator control.

As this study was carried out on an apartment in an older building, it would be interesting to also perform similar measurements on newer low-energy buildings. These buildings have a greater thermal inertia than older ones, meaning the relation between household electricity consumption and indoor temperature may look different. As heat is preserved better in these new buildings it is most likely that the usage of electrical appliances and the presence of people will affect the indoor temperature in another proportion compared to older buildings. It is reasonable to assume that the same amount of heat generated from electrical appliances in an older and less energy efficient apartment would result in heavier impacts on the indoor temperature in newer apartments with less heat leakage. As more energy is needed to heat up older apartments, the heat generated from electrical appliances and people present correspond to a smaller portion of the total energy affecting the indoor temperature. Of course, some of the appliances in new buildings are more energy efficient and therefore yields less heat. Also, many of these buildings do already have smart systems for controlling the indoor climate, similar to the technologies described earlier. Still, it is possible that low-energy buildings are the most sensitive ones regarding the impact of household electricity consumption on indoor temperature and that a potential for energy savings exists.

References

Adetiloye, T., Awasthi, A. (2017), “Predicting Short-Term Congested Traffic Flow on Urban Motorway Networks”, Handbook of Neural Computation, pp. 145-165.

Afroz, Z., Shafiullah, G.M., Urmee, T., Higgins, G. (2018), “Modeling techniques used in building HVAC control systems: A review”, Renewable and Sustainable Energy Reviews, vol. 83, pp. 64-84.

Akiful, M.D. & Qavi, I. (2017), “Natural Convection Heat Transfer from Human Body in different postures”. International Conference on Mechanical Engineering and Renewable Energy, ICMERE2017, Chittagong, Bangladesh.

Alm, S.E. & Britton, T., (2008), ”Stokastik: sannolikhetsteori och statistikteori med tillämpningar”, 1st edition, Liber, Stockholm.

Balan, R., Cooper, J., Chao, K., Stan, S., Donca, R. (2011), “Parameter identification and model based predictive control of temperature inside a house”, Energy and Buildings, vol.43, nr. 2-3, pp. 748-758.

Bourke, P. (1996), “Cross correlation, AutoCorrelation – 2D Pattern Identification”, Available at: http://paulbourke.net/miscellaneous/correlate/. (Retrieved: 2020-04-17).

Carlsson, Lindholm. (2019), Linear regression. Department of information technology, Uppsala University.

Chujai, P., Kerdprasop, N., Kerdprasop, K. (2013), “Time series analysis of household electric consumption with ARIMA and ARMA models”, Proceedings of the International MultiConference of Engineering and Computer Scientists, Hong Kong, vol. 1, pp. 295-300.

Rogers, K. (2012), Scientific modeling. *Britannica.* Available at: https://www.britannica.com/science/scientific-modeling. (Retrieved: 2020-01-21)

Duun-Henriksen, A.K., Schmidt, S., Norgaard, M.R., Rikke, K., B Moller, J., B. Jorgensen, J., Madsen, H. (2013), “Model Identification Using Stochastic Differential Equation Grey-Box Models in Diabetes”, Journal of diabetes science and technology, vol. 7, nr. 2, pp. 431-440.

Energibutiken. (2019), “Logger 2020”, Available at: https://www.energibutiken.se/sv/datalogger/102-logger-2020-startpaket-datalogger.html. (Retrieved: 2020-02-18).

Energimyndigheten. (2019), ”Scenarier över Sveriges energisystem 2018”. ER 2019:07.

Engineering Statistics Handbook (ESH). (2020), “Autocorrelation”, Available at: https://www.itl.nist.gov/div898/handbook/eda/section3/eda35c.htm. (Retrieved: 2020-04-17).

Fielsch, S., Grunert T., Stursberg, M., Kummert, A. (2017), “Model Predictive Control for Hydronic Heating Systems in Residential Buildings”, IFAC-PapersOnLine. vol. 50, nr. 1. pp. 4216-4221.

Firth, S., Lomas, K., Wright, A. & Wall, R. (2008), "Identifying trends in the use of domestic appliances from household electricity consumption measurements", Energy & Buildings, vol. 40, no. 5, pp. 926-936.

Hernández-Molinar, R., Sarmiento-Rebeles, R., F. Méndez-Barrios, C. (2016), “Least Squares Method and Empirical Modeling: A Case Study in a Mexican Manufacturing Firm”, Empirical Modeling and Its Applications, Chapter 3, IntechOpen.

Hietaharju, P., Ruusunen, M., Leiviskä, K. (2018), “A Dynamic Model for Indoor Temperature Prediction in Buildings”, Energies, vol. 11, nr. 6, pp. 1-20.

Huang, H., Chen, L., Mohammadzaheri, M. (2012), “A New Zone Temperature Predictive Modeling for Energy Saving in Buildings”. Procedia Engineering, vol. 49, pp. 142–151.

Hulstaert, L. (2019), “Black-box vs. white-box models”, Available at: https://towardsdatascience.com/machine-learning-interpretability-techniques-662c723454f3. (Retrieved: 2020-01-29).

Intab. (2019), “Tinytag Ultra 2”, Available at: https://intab.se/catalog/products/tinytag-ultra-2-intern-temperatur. (Retrieved: 2020-02-18).

Karlsson, P. (2019), Knowledge manager at Aktea Energy. Interview from November 2019.

Karlsson, J. (2012), “Possibilities of using thermal mass in buildings to save energy, cut power consumption peaks and increase the thermal comfort”, Division of Building Materials, LTH, Lund University.

Kenton, W. (2019), “Sum of squares”, Available at: https://www.investopedia.com/terms/s/sum-of-squares.asp. (Retrieved: 2020-04-24).

Kramer, R., Schijndel, J., Schellen, H.L. (2012), “Simplified thermal and hygric building models: A literature review”, Frontiers of Architectural Research, vol. 1, nr. 4.

LeCun, Y., Denker, J. & Solla, S. (1990), “Optimal brain damage”, Advances in neural information processing systems (NIPS), vol. 2, pp. 598–605.

Ljung, L., Glad, T. (2003), ”Modellbygge och simulering”, 2nd edition, Studentlitteratur AB, Lund.

Mathworks. (2018), “GoodnessOfFit”, Available at: https://se.mathworks.com/help/releases/R2018b/ident/ref/goodnessoffit.html. (Retrieved: 2020-04-13).

Mathworks. (2020a), “Corrcoef”, Available at: https://se.mathworks.com/help/matlab/ref/corrcoef.html. (Retrieved: 2020-01-28).

Mathworks. (2020b), “Corr”, Available at: https://se.mathworks.com/help/stats/corr.html. (Retrieved: 2020-02-20).

Mathworks. (2020c), “Xcorr”, Available at: https://se.mathworks.com/help/matlab/ref/xcorr.html. (Retrieved: 2020-02-24).

Mathworks. (2020d), “Delayest”, Available at: https://se.mathworks.com/help/ident/ref/delayest.html. (Retrieved: 2020-03-02).

Mathworks. (2020e), “Spa”, Available at: https://se.mathworks.com/help/ident/ref/spa.html. (Retrieved: 2020-02-26).

Mathworks. (2020f), “What are non-linear ARX models?”, Available at: https://www.mathworks.com/help/ident/ug/what-are-nonlinear-arx-models.html?fbclid=IwAR03iTWNWKnN8xE9saJ4SEjwA9LsaERrBLy4ASC1yMd5Rec666VUytc18EQ. (Retrieved: 2020-04-14).

Mathworks. (2020g), “Sigmoidnet”. Available at: https://www.mathworks.com/help/ident/ref/sigmoidnet.html. (Retrieved 2020-05-01).

Mattsson, P., Zachariah, D., Stoica, P. (2018), Recursive nonlinear-system identification using latent variables, Automatica, vol. 93, pp. 343-351.

Merriam Webster. (2019), “Model”, Available at: https://www.merriam-webster.com/dictionary/model. (Retrieved: 2020-01-28).

Mustafaraj, G., Chen, J., Lowry, G. (2010), “Development of room temperature and relative humidity linear parametric models for an open office using BMS data”. Energy and Buildings, vol. 42, nr. 3, pp. 348-356.

Naturskyddsföreningen. (2016), ”Faktablad: Energianvändning”.

Oja, H., & Shevlyakov, G.L. (2016), “Robust Correlation: Theory and Applications”, 1st edition, John Wiley & Sons Inc. New York.

Regeringskansliet. (2016), ”Överenskommelse om Sveriges mål för energieffektivisering”, Available at: https://www.regeringen.se/pressmeddelanden/2016/11/overenskommelse-om-sveriges-mal-for-energieffektivisering/. (Retrieved: 2020-01-30).

Rychlik, I., Rydén, J. (2006), ”Probability and risk analysis: an introduction for engineers”, Teknisk-naturvetenskapliga vetenskapsområdet, Uppsala universitet, Matematiska institutionen & Matematisk-datavetenskapliga sektionen 2006, Springer, Berlin.

SCB. (2016), ”Vanligast med två rum och kök på 57 kvadratmeter”, Available at: https://www.scb.se/hitta-statistik/artiklar/2016/Vanligast-med-2-rum-och-kok-pa-57-kvadratmeter/. (Retrieved 2020-05-25).

Schober, P., Boer, C. & Schwarte, L.A. (2018), “Correlation Coefficients: Appropriate Use and Interpretation”. Anesthesia and Analgesia. vol. 126, no. 5, pp. 1763-1768.

Schmidhuber, J. (2014). “Deep learning in neural networks: An overview”. Neural Networks, vol. 61, pp. 85-117.

Schneider, RA. (2009), “Basic Statistics for Social Workers”, Revised edition, University Press of America, UPA, Lanham, MD.

Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Deylon, B., Glorennec, P.Y., Hjalmarsson, H., Judisky, A. (1995), “Nonlinear black-box modeling in system identification: a unified overview”, Automatica, vol. 31, nr. 12, pp. 1691-1724.

Souza, C.B., & Alsaadani, S. (2012), “Thermal zoning in speculative office buildings: discussing the connections between space layout and inside temperature control”, First building simulation and optimization conference, Loughborough, UK.

Suszanowicz, D. (2017), “Internal heat gain from different light sources in the building lighting systems”, E3S Web of Conferences, vol. 19, pp. 1024.

Svensson, A. (2018), ”Föreläsning 9-Parameterskattning i Box-Jenkins modellfamilj, och lite skräddarsydda linjära modeller”. Lecture notes. Empirical modeling 1RT890. Uppsala University.

Svensson, G. (1973), “Dygnsbehovet av tappvarmvatten: behovsfördelningen i flerfamiljshus med gemensam varmvattenberedning”, Stockholm.

TADO. (2020), “Smart Radiator Thermostat- Starter Kit V3+”, Available at: https://www.tado.com/se/products/smart-radiator-starter-kit. (Retrieved 2020-05-26).

Vilela, M., Danuser, G. (2013), “Fluorescence Fluctuation Spectroscopy (FFS) Part B”, Methods in Enzymology, 1st edition, vol. 519, Academic Press, Cambridge, Massachusetts.

Voll, H., De luca, F., Pavlovas, V. (2016), “Analysis of the insolation criteria for nearly-zero energy buildings in Estonia”, Science and Technology for the Built Environment, vol. 47, pp. 939-950.

Wolfram. (2020), “Covariance”, Available at: https://mathworld.wolfram.com/Covariance.html. (Retrieved 2020-04-17).

Yiu, T. 2019. “Understanding Neural Networks”, Available at: https://towardsdatascience.com/understanding-neural-networks-19020b758230. (Retrieved 2020-02-12).

Zavattoni, S, Garcia-Polanco, N, Capablo, J, Doyle, J, Barbato, M. (2014), "Household Appliances Wasted Heat Storage by means of a Packed bed TES with Encapsulated PCM". 13th International conference on sustainable energy technologies. Geneva, Switzerland.

Appendix A- Three-day data codes

A.1. Correlation analysis

```
% Correlation analysis kitchen.
clear all
close all

%Kitchen:
kitchen = importdata('kitchen11to13NEW.txt');
temp_kitchen = kitchen(1:end, 1);
kW_kitchen = kitchen(1:end, 2);

%Plot:
time=[1:length(temp_kitchen)];
plot(time,temp_kitchen)
grid on;
hold on;
plot(time,kW_kitchen)
figure;

%Scatterplot:
scatter(kW_kitchen, temp_kitchen)
figure;

%Correlation coefficients (Pearson¥s).
[correlation_coefficients,p] = corrcoef(kW_kitchen',temp_kitchen');

%Correlation coefficients (Spearman¥s).
[SRHO,Pval] = corr(kW_kitchen,temp_kitchen,'Type','Spearman');

%Iddata:
kitchen_data = iddata(temp_kitchen, kW_kitchen, 60);
%Removing means:
kitchen_data =detrend(kitchen_data,0);
%Detrending:
kitchen_data = detrend(kitchen_data,1);
%Estimation data:
kitchen_data_est = kitchen_data([1:round(length(time)/2)]);
%Validation data:
kitchen_data_val = kitchen_data([(round(length(time)/2))+1:length(time)]);

%Estimating impulse response:
[impulse_response,Corr,sign_level] = cra(kitchen_data,2000,4,1);
figure;

%Estimating time delay:
time_delay=delayest(kitchen_data, 4, 4)

%Estimating Bode-diagram:
bodeplot(spa(kitchen_data));
figure;

%Spectrumplot:
spectrumplot(spa(kitchen_data));
```

```
%Prediction comparison:

%AR:
arOpt = arOptions;
ar2=ar(kitchen_data_est.y,2,'Ts', 60, arOpt);
compare(kitchen_data_val.y,ar2, 15);
figure;

%ARX:
arxOpt = arxOptions;
arx2115 = arx([kitchen_data_est.y, kitchen_data_est.u],[2 1 15],'Ts', 60,
arxOpt);
compare([kitchen_data_val.y, kitchen_data_val.u],arx2115,15);

%AR tuned 15-step predictor:
Y_est = kitchen_data_est.y(17:end);
y_est = kitchen_data_est.y;
Y_val = kitchen_data_val.y(17:end);
y_val = kitchen_data_val.y;
u_est = kitchen_data_est.u;
u_val = kitchen_data_val.u;

Phi_est = [-y_est(2:end-15) -y_est(1:end-16)];
theta_hat_ar = Phi_est\Y_est;
y_hat_est = Phi_est*theta_hat_ar;

%Validation data:
Phi_val = [-y_val(2:end-15) -y_val(1:end-16)];
y_hat_val = Phi_val*theta_hat_ar;

%NRMSE-based goodness of fit:
GOF_est_ar = goodnessOfFit(Y_est,y_hat_est,'NRMSE')
GOF_val_ar = goodnessOfFit(Y_val,y_hat_val,'NRMSE')

%ARX tuned 15-step predictor:
Phi_est_arx = [-y_est(2:end-15) -y_est(1:end-16) u_est(1:end-16)];
Phi_val_arx = [-y_val(2:end-15) -y_val(1:end-16) u_val(1:end-16)];

theta_hat_arx = Phi_est_arx\Y_est;
y_hat_est_arx = Phi_est_arx*theta_hat_arx;

%Validation data.
y_hat_val_arx = Phi_val_arx*theta_hat_arx;

%NRMSE-based goodness of fit:
GOF_est_arx = goodnessOfFit(Y_est,y_hat_est_arx,'NRMSE')
GOF_val_arx = goodnessOfFit(Y_val,y_hat_val_arx,'NRMSE')
```

Appendix B- Two-week data codes

B.1. Best linear tuned one-step models with respect to GOF

```
function[model] = find_model(file, model_type)

%Data:

[temperature, humidity, electricity, radiator_temp]= split_data(file);
time = [1:length(temperature)];
data_length = length(time);

room_data = iddata(temperature, electricity, 60);
 room_data = detrend(room_data,0);
 room_data = detrend(room_data,1);
 room_data_est = room_data([1:floor(data_length/2)]);
 room_data_val = room_data([ceil(data_length/2):data_length]);

 counter = 1;
 delays = zeros(16,1);
 max_na = 4;
 max_nb = 4;

 for na = 1:max_na
   for nb = 1:max_nb
     time_delay=delayest(room_data, na, nb);
     delays(counter, 1) = time_delay;
     counter = counter+1;
   end
 end

 %Most frequent time delay:
 most_freq_time_delay = mode(delays);

  if model_type == "arx"
 %Evaluate the arx-structures with fixed time delay:
    arx_structures = struc(1:4, 1:4, most_freq_time_delay);
    best_fit = 0;
    best_model_index = 0;
     for i = 1:length(arx_structures)
          arx_model = arx(room_data_est, arx_structures(i,:));
          [y,fit,x0] = compare(room_data_val, arx_model, 15);
          fit1 = goodnessOfFit(y.y,room_data_val.y, 'NRMSE')*100;
          if fit1 > best_fit
              best_fit = fit1;
              best_model_index = i;
          end
     end
       opt = arxOptions;
            model = arx(room_data_est, arx_structures(best_model_index,:)
            ,'Ts', 60, opt);
           compare(room_data_val, model, 15)
 end

 if model_type == "armax"
 %Evaluate the armax-structures with fixed time delay:
```

```
    armax_structures = struc(1:4, 1:3, 1:2, most_freq_time_delay);
    best_fit = 0;
    best_model_index = 0;
    for i = 1:length(armax_structures)
             armax_model = armax(room_data_est,    armax_structures(i,:));
            [y,fit,x0] = compare(room_data_val, armax_model);
            [y,fit,x0] = compare(room_data_val, armax_model, 15);
            fit1 = goodnessOfFit(y.y,room_data_val.y, 'NRMSE')*100;
            if fit1 > best_fit
                  best_fit = fit1;
                  best_model_index = i;
            end

    end
             model = armax(room_data_est,
             armax_structures(best_model_index,:), 'Ts', 60);
            compare(room_data_val, model, 15)
end

  if model_type == "bj"
  %Evaluate the bj-structures with fixed time delay:
             bj_structures = struc(1:4, 1:4, 1:4, 1:4, most_freq_time_delay);
            best_fit = 0;
            best_model_index = 0;
        for i = 1:length(bj_structures)
              bj_model = bj(room_data_est, bj_structures(i,:));
              [y,fit,x0] = compare(room_data_val, bj_model, 15);
              fit1 = goodnessOfFit(y.y,room_data_val.y, 'NRMSE')*100;
              if fit1 > best_fit
                  best_fit = fit1;
                  best_model_index = i;
             end
        end
             model = bj(room_data_est, bj_structures(best_model_index,:),
             'Ts', 60);
            compare(room_data_val, model, 15)

  end

  if model_type == "oe"
  %Evaluate the oe-structures with fixed time delay:
        oe_structures = struc(1:4, 1:4, most_freq_time_delay);
        best_fit = 0;
        best_model_index = 0;
        for i = 1:length(oe_structures)
              oe_model = oe(room_data_est, oe_structures(i,:));
              [y,fit,x0] = compare(room_data_val, oe_model, 15);
              fit1 = goodnessOfFit(y.y,room_data_val.y, 'NRMSE')*100;
              if fit1 > best_fit
                  best_fit = fit1;
                  best_model_index = i;
              end

        end
             model = oe(room_data_est, oe_structures(best_model_index,:),
             'Ts', 60);
            compare(room_data_val, model, 15)
```

```
    end

    if model_type == "ar"
        %Evaluate the ar-structures with fixed time delay:
        best_fit = 0;
        best_model_index = 0;
        for i = 1:4
             ar_model = ar(room_data_est.y, i);
             [y,fit,x0] = compare(room_data_val.y, ar_model, 15);
             fit1 = goodnessOfFit(y,room_data_val.y, 'NRMSE')*100;
             if fit1 > best_fit
                 best_fit = fit1;
                 best_model_index = i;
             end

        end
           model = ar(room_data_est.y, best_model_index, 'Ts', 60);
           compare(room_data_val.y, model, 15)

    end

    if model_type == "arma"
        %Evaluate the arma-structures with fixed time delay:
        arma_structures = struc(1:4, 1:4);
        best_fit = 0;
        best_model_index = 0;
        for i = 1:length(arma_structures)
             arma_model = armax(room_data_est.y, arma_structures(i,:));
            [y,fit,x0] = compare(room_data_val.y, arma_model, 15);
            fit1 = goodnessOfFit(y,room_data_val.y, 'NRMSE')*100;
             if fit1 > best_fit
                 best_fit = fit1;
                 best_model_index = i;
            end

        end
            model = armax(room_data_est.y,
            arma_structures(best_model_index,:), 'Ts', 60);
           compare(room_data_val.y, model, 15)

      end

    end
```

B.2. Linear tuned one-step- and 15-step predictors

```
%Kitchen 2-week:
[temperature, humidity, electricity, radiator_temp]=
split_data('KitchenDone.txt');
time = [1:length(temperature)];
data_length = length(time);

room_data = iddata(temperature, electricity, 60);
room_data = detrend(room_data,0);
room_data = detrend(room_data,1);
```

```
room_data_est = room_data([1:floor(data_length/2)]);
room_data_val = room_data([ceil(data_length/2):data_length]);

% Best models:
arx318 = arx([room_data_est.y, room_data_est.u],[3 1 8], 'Ts', 60);
compare([room_data_val.y, room_data_val.u],arx318,15);
figure;
armax1228 = armax([room_data_est.y, room_data_est.u],[1 2 2 8], 'Ts',  60);
compare([room_data_val.y, room_data_val.u],armax1228,15);
figure;
bj31218 = bj([room_data_est.y, room_data_est.u],[3 1 2 1 8], 'Ts', 60);
compare([room_data_val.y, room_data_val.u],bj31218,15);
figure;
oe248= oe([room_data_est.y, room_data_est.u],[2 4 8], 'Ts', 60);
compare([room_data_val.y, room_data_val.u],oe248,15);
figure;
ar3= ar(room_data_est.y,3, 'Ts', 60);
compare(room_data_val.y,ar3,15);
figure;
arma14 = armax(room_data_est.y,[1 4], 'Ts', 60);
compare(room_data_val.y, arma14, 15);

%Evaluation:
resid(room_data_val,arx318);
figure;
resid(room_data_val, armax1228);
figure;
resid(room_data_val, bj31218);
figure;
resid(room_data_val, oe248);
figure;

iopzplot(arx318);
[poles,zeros] = pzmap(arx318)
figure;
iopzplot(armax1228);
[poles,zeros] = pzmap(armax1228)
figure;
iopzplot(bj31218);
[poles,zeros] = pzmap(bj31218)
figure;
iopzplot(oe248);
[poles,zeros] = pzmap(oe248)
figure;

step(impulseest(spa(room_data)),'r', arx318,'b', armax1228, 'g', bj31218,
'm', oe248, 'y')
figure;

bodeplot(spa(room_data),'r', arx318,'b', armax1228, 'g', bj31218, 'm', oe248,
'y');

%AR 15-step predictor:
Y_est = room_data_est.y(18:end);
y_est = room_data_est.y;
Y_val = room_data_val.y(18:end);
y_val = room_data_val.y;
u_est = room_data_est.u;
```

```
u_val = room_data_val.u;

Phi_est = [-y_est(3:end-15) -y_est(2:end-16) -y_est(1:end-17)];
theta_hat_ar = Phi_est\Y_est;
y_hat_est = Phi_est*theta_hat_ar;

%Validation data:
Phi_val = [-y_val(3:end-15) -y_val(2:end-16) -y_val(1:end-17)];
y_hat_val = Phi_val*theta_hat_ar;

% GOF:
GOF_est_ar = goodnessOfFit(Y_est,y_hat_est,'NRMSE')
GOF_val_ar = goodnessOfFit(Y_val,y_hat_val,'NRMSE')

%ARX 15-step predictors:
Phi_est_arx = [-y_est(3:end-15) -y_est(2:end-16) -y_est(1:end-17)
u_est(1:end-17)];
Phi_val_arx = [-y_val(3:end-15) -y_val(2:end-16) -y_val(1:end-17)
u_val(1:end-17)];

theta_hat_arx = Phi_est_arx\Y_est;
y_hat_est_arx = Phi_est_arx*theta_hat_arx;

%Validation data:
y_hat_val_arx = Phi_val_arx*theta_hat_arx;

% GOF:
GOF_est_arx = goodnessOfFit(Y_est,y_hat_est_arx,'NRMSE')
GOF_val_arx = goodnessOfFit(Y_val,y_hat_val_arx,'NRMSE')
```

B.3. NLAR-and NLARX-models

```
file = "KitchenDone.txt";
[temperature, humidity, electricity, radiator_temp]= split_data2(file);

time = [1:length(temperature)];
data_length = length(time);

%Treat data:
room_data = iddata(temperature, electricity, 60);
room_data = detrend(room_data,0);
room_data = detrend(room_data,1);
room_data_est = room_data([1:floor(data_length/2)]);
room_data_val = room_data([ceil(data_length/2):data_length]);

%Opt:
opt = nlarxOptions;
opt.Focus = 'prediction';

arx=[3 4 1];
training_data = room_data_est;

GOF_fit = [];
counter = 1;
for i = 1:20
```

```
    NN_est = nlarx(training_data, arx,sigmoidnet('numberOfUnits',i),
    opt);
    [y, FIT, x0]=compare(room_data_val, NN_est, 15);
    GOF_fit(counter)=FIT;
    counter= counter + 1;
end

%Best model structure:
best_FIT=max(GOF_fit)
index = find(GOF_fit==best_FIT)
```

B.4. LAVA AR-predictors

```
%Read data points with split_data function:
[temperature, humidity, electricity, radiator] =
split_data('KitchenDone.txt');

%Divide data into estimation- and validation data:
data_length = length(electricity);
temperature_est = temperature([1:floor(data_length/2)]);
temperature_val = temperature([ceil(data_length/2):data_length]);
electricity_est = electricity([1:floor(data_length/2)]);
electricity_val = electricity([ceil(data_length/2):data_length]);

%Pre-process data, removing means and autocorrelation:
Y_id = detrend(detrend(temperature_est', 0), 1);
U_id = detrend(detrend(electricity_est', 0), 1);
Y_val = detrend(detrend(temperature_val', 0), 1);
U_val = detrend(detrend(electricity_val', 0), 1);

N = size(Y_id, 2);

% Model parameters:
na = 3;
nb = 0;
M = 3; %Resolution of Laplace basis.
L_vec = 1.1*[ repmat(max(abs(Y_id')),1,na) repmat(max(abs(U_id')),1,nb)];
%Boundaries in Laplace basis

%Construct AR-Predictor for prediction:
Phi_est = [-Y_id(3:end-15); -Y_id(2:end-16); -Y_id(1:end-17)];
Gamma = create_gamma_laplace(Phi_est', M, L_vec);

% Run identification:
% We remove the first max(na,nb) elements from Y_id, due to unknown initial
conditions.

L = 5; %Number of iteration per sample in cyclic minimization
[Theta_hat, Z_hat, D, Sigma, Theta_tilde] = func_onlinesol(Y_id(:,
max(na,nb)+15:end), Phi_est, Gamma, L); %Gamma ska kanske ha transponat!

%Predictor 15 steps ahead, AR:
y_hat_est_ar = Phi_est'*Theta_hat'+(Z_hat*Gamma)';

Phi_val = [-Y_val(3:end-15); -Y_val(2:end-16); -Y_val(1:end-17)];
```

```
Gamma_val = create_gamma_laplace(Phi_val', M, L_vec);


y_hat_val_ar = Phi_val'*Theta_hat'+(Z_hat*Gamma_val)';

%Calculate goodness of fit with NRMSE:
Y_est_stripped = Y_id(18:end);
Y_val_stripped = Y_val(18:end);

nrmse_est_ar = goodnessOfFit(Y_est_stripped',y_hat_est_ar,'NRMSE');
nrmse_val_ar = goodnessOfFit(Y_val_stripped', y_hat_val_ar,'NRMSE');

%Get fit percentage:
fit_val_ar = (1-nrmse_val_ar)*100;
fit_est_ar = (1-nrmse_est_ar)*100;

%Plot predictions:
figure
hold on
plot(y_hat_val_ar)
plot(Y_val_stripped)
legend('Y predictions', 'True Y')
```

B.5. LAVA ARX-predictors

```
%Read data points with split_data function:
[temperature, humidity, electricity, radiator] =
split_data('KitchenDone.txt');

%Divide data into estimation- and validation data:
data_length = length(electricity);
temperature_est = temperature([1:floor(data_length/2)]);
temperature_val = temperature([ceil(data_length/2):data_length]);
electricity_est = electricity([1:floor(data_length/2)]);
electricity_val = electricity([ceil(data_length/2):data_length]);

%Pre-process data, removing means and autocorrelation:
Y_id = detrend(detrend(temperature_est', 0), 1);
U_id = detrend(detrend(electricity_est', 0), 1);
Y_val = detrend(detrend(temperature_val', 0), 1);
U_val = detrend(detrend(electricity_val', 0), 1);

N = size(Y_id, 2);

% Model parameters:
na = 3;
nb = 1;
M = 3; %Resolution of Laplace basis.
L_vec = 1.1*[ repmat(max(abs(Y_id')),1,na) repmat(max(abs(U_id')),1,nb)];
%Boundaries in Laplace basis


%Construct ARX-Predictor for prediction:
```

```
Phi_est = [-Y_id(3:end-15); -Y_id(2:end-16); -Y_id(1:end-17); U_id(1:end-
17)];
Gamma = create_gamma_laplace(Phi_est', M, L_vec);

% Run identification:
% We remove the first max(na,nb) elements from Y_id, due to unknown initial
conditions.
L = 5; %Number of iteration per sample in cyclic minimization
[Theta_hat, Z_hat, D, Sigma, Theta_tilde] = func_onlinesol(Y_id(:,
max(na,nb)+15:end), Phi_est, Gamma, L); %Gamma ska kanske ha transponat!

%Predictor 15 steps ahead, ARX:
y_hat_est_arx = Phi_est'*Theta_hat'+(Z_hat*Gamma)';

Phi_val = [-Y_val(3:end-15); -Y_val(2:end-16); -Y_val(1:end-17); U_val(1:end-
17)];
Gamma_val = create_gamma_laplace(Phi_val', M, L_vec);

y_hat_val_arx = Phi_val'*Theta_hat'+(Z_hat*Gamma_val)';

%Calcluate goodness of fit with NRMSE:
Y_est_stripped = Y_id(18:end);
Y_val_stripped = Y_val(18:end);

nrmse_est_arx = goodnessOfFit(Y_est_stripped',y_hat_est_arx,'NRMSE');
nrmse_val_arx = goodnessOfFit(Y_val_stripped', y_hat_val_arx,'NRMSE');

%Get fit percentage:
fit_val_arx = (1-nrmse_val_arx)*100;
fit_est_arx = (1-nrmse_est_arx)*100;

%Plot predictions:
figure
hold on
plot(y_hat_val_arx)
plot(Y_val_stripped)
legend('Y predictions', 'True Y')
```

Appendix C- Sensitivity analysis codes

C.1. Kitchen analysis

```
[temperature, humidity, electricity, radiator_temp]=
split_data('KitchenDone.txt');
time = [1:length(temperature)];
data_length = length(time);

%Humidity:
room_data_hum = iddata(temperature, humidity, 60);
room_data_hum = detrend(room_data_hum,0);
room_data_hum = detrend(room_data_hum,1);
room_data_hum_est = room_data_hum([1:floor(data_length/2)]);
room_data_hum_val = room_data_hum([ceil(data_length/2):data_length]);

%Electricity:
room_data_el = iddata(temperature, electricity, 60);
room_data_el = detrend(room_data_el,0);
room_data_el = detrend(room_data_el,1);
room_data_el_est = room_data_el([1:floor(data_length/2)]);
room_data_el_val = room_data_el([ceil(data_length/2):data_length]);

%Radiator temperature:
room_data_rad = iddata(temperature, radiator_temp, 60);
room_data_rad = detrend(room_data_rad,0);
room_data_rad = detrend(room_data_rad,1);
room_data_rad_est = room_data_rad([1:floor(data_length/2)]);
room_data_rad_val = room_data_rad([ceil(data_length/2):data_length]);

%Do high levels of humidity coincide with high levels of electricity usage?

plot(time, humidity)
hold on;
plot(time, electricity)
figure;
plot(time, radiator_temp)

%Correlation coefficients (Pearson¥s).
[correlation_coefficients,p] = corrcoef(humidity,electricity);

%Correlation coefficients (Spearman¥s).
[SRHO,Pval] = corr(humidity,electricity,'Type','Spearman');

%Best OE-models for kitchen with regard to goodness of fit:
oe_el= oe([room_data_el_est.y, room_data_el_est.u],[2 4 8], 'Ts', 60);
compare([room_data_el_val.y, room_data_el_val.u],oe_el,15);
figure;

oe_hum= oe([room_data_hum_est.y, room_data_hum_est.u],[3 4 0], 'Ts', 60);
compare([room_data_hum_val.y, room_data_hum_val.u],oe_hum,15);
figure;

oe_rad= oe([room_data_rad_est.y, room_data_rad_est.u],[4 4 0], 'Ts', 60);
compare([room_data_rad_val.y, room_data_rad_val.u],oe_rad,15);
```

```
figure;

%Best ARX-models with reagrd to goodness of fit:
arx_el = arx([room_data_el_est.y, room_data_el_est.u], [3 4 8], 'Ts', 60);
compare([room_data_el_val.y, room_data_el_val.u], arx_el,15);
figure;

arx_hum = arx([room_data_hum_est.y, room_data_hum_est.u], [3 3 0], 'Ts', 60);
compare([room_data_hum_val.y, room_data_hum_val.u], arx_hum,15);
figure;

arx_rad = arx([room_data_rad_est.y, room_data_rad_est.u], [3 2 0], 'Ts', 60);
compare([room_data_rad_val.y, room_data_rad_val.u], arx_rad,15);
figure;

%Multiple input ARX (electricity, humidity, radiator temp.):
arx_misu_all=arx([room_data_el_est.y, room_data_el_est.u room_data_hum_est.u
room_data_rad_est.u],[3 4 3 2 8 0 0], 'Ts', 60);
compare([room_data_el_val.y, room_data_el_val.u room_data_hum_val.u
room_data_rad_val.u ], arx_misu_all,15);
figure;

%Multiple input ARX (electricity, radiator temp.):
arx_misu_el_rad=arx([room_data_el_est.y, room_data_el_est.u
room_data_rad_est.u],[3 4 2 8 0], 'Ts', 60);
compare([room_data_el_val.y, room_data_el_val.u room_data_rad_val.u ],
arx_misu_el_rad,15);
```