UPPSALA UNIVERSITET

UPTEC STS 21023
Examensarbete 30 hp
Juni 2021

Development of grey-box models for simulating heating consumption in buildings

A study applying system identification methodology to a physics-based framework

Zack Klockar

Civilingenjörsprogrammet i system i teknik och samhälle

UPPSALA UNIVERSITET

Development of grey-box models for simulating heating consumption in buildings: A study applying system identification methodology to a physics-based framework

Zack Klockar

Abstract

This study models the energy used for heating in buildings by applying system identification methodology. The model development is grounded in physics to provide guidance and interpretability when evaluated. Time-series of heating demand, outdoor temperature, indoor temperature and solar irradiation are considered for the modelling purpose. Evaluation is done through simulation and relies on quantitative measures, residual analysis and visual inspection of model output. Through model development, the study seeks to extrapolate information of physical properties that drives heating demand in buildings. Seven buildings located in the same geographic area are studied. It is found that linear ARX-models can simulate heating demand with high precision but at times low accuracy. A common system model structure can be identified between buildings, indicating that physical properties shared between buildings can be identified through this methodology. A sensitivity analysis is conducted to derive the contributions from model constituents to simulation results. Two buildings were also modelled as OE-models. These models performed better than the respective ARX-models but were deemed more difficult to use for the purpose of this study. The study finds difficulties in implementing aggregated time-series of indoor temperature, which could be explored further in future studies for more detailed interpretations. An interpretation of the physical properties identified is proposed.

Teknisk-naturvetenskapliga fakulteten

Uppsala universitet, Utgivningsort Uppsala

Handledare: Amanda Fors Ämnesgranskare: Per Mattsson

Examinator: Elísabet Andrésdóttir

Svensk populärvetenskaplig sammanfattning

Det finns ett stort behov av att minska energikonsumtionen i världen. Till stor del är behovet sprunget ur ett ohållbart resursanvändande och klimatgasutsläpp relaterat till energianvändningen. Energiförbrukningen i byggnader är stor och det går åt mycket energi för att värma upp byggnader. I världen, likaså Sverige, står uppvärmningen av byggnader för ungefär en fjärdedel av den totala energiförbrukningen.

Uppvärmningsbehovet i en byggnad beror huvudsakligen av tre aspekter: dess konstruktion, dess användning och vädret som råder. Det sista är alla bekanta med: när det är kallt så värmer vi husen. Det föregående är de flesta också bekanta med: om du öppnar fönstret på vintern så blir det kallt. Men det är också en fråga om att mänsklig aktivitet, såsom kroppsvärme, datorer och lampor, avger värme och minskar uppvärmningsbehovet. Den första aspekten är däremot olika för alla byggnader och beror både på dess ursprungliga utformning samt dess nuvarande skick. Ett sätt att se på en byggnad är att den ska skydda oss ifrån vädrets yttre förhållanden, genom att till exempel hålla oss isolerade i värmen. Dessa tre aspekter kan sammantaget beskrivas med fysik som en balans av energiflöden.

För att både effektivisera och minska uppvärmningsbehovet har modeller av olika slag utvecklats. Deras syften varierar från att designa byggnadens originalutformning, till att kontrollera uppvärmningssystemet på ett effektivt sätt. Det finns därför många olika typer av modeller som har utvecklats och olika lärdomar att dra. Den här studien syftar till att ta fram en modell som kan simulera uppvärmningsbehovet samt ge oss insikt i vad som driver det och hur.

Grovt sätt så kan modeller delas in i två typer. Den ena typen utgår ifrån att modellen byggs helt på förståelse för ett system, en så kallad ”vitlåde”-modell. Till exempel kan fysikaliska lagar beskriva ett känt system. En sådan modell är även helt ”genomskinlig” och vi kan förstå allt som händer i modellen. Den andra typen utgår ifrån att en modell av system ska härledas ifrån observationer av det, så kallade ”svartlåde”-modeller. Dessa modeller bygger helt på våra observationer av ett system och är inte direkt relaterade till fysik eller någon annan förståelse för systemet. På så vis blir de unikt anpassade till systemet i fråga, men behöver inte alls beskriva ett liknande system väl.

Det går att kombinera dessa två typer av modeller och få en så kallad ”grålåde”-modell. Dessa erbjuder förståelsen av en vitlåde-modell, samt den unika anpassningen av en svartlåde-modell. I den här studien utvecklas grålåde-modeller för sju olika byggnader. Modellerna tas fram genom att använda fysik för att utgöra en modellgrund och sedan tillämpa en maskininlärningsmetodik – systemidentifiering – för att anpassa modellerna för vardera byggnaden. På så vis utvecklas anpassade modeller som delar en gemensam och fysiskt tolkningsbar grund. En sådan modell kan möjligen i förlängningen lämpa sig för att undersöka vilken inverkan en energieffektiviserande renovering hade haft på just en specifik byggnad.

De framtagna modellerna beskriver samspelet mellan byggnadens utformning och de yttre väderförhållandena genom att identifiera olika energiflöden och beteenden hos byggnadens uppvärmningssystem. Detta identifieras genom att analysera de framtagna modellerna med hjälp av vad fysik och vad som bör vara fallet. Därigenom beskriver de framtagna modellerna att energikonsumtionen från uppvärmningen drivs av

- en tröghet i uppvärmningssystemet,
- värmeförluster genom fönster och köldbryggor samt
- värmeförluster genom tunga byggnadsmaterial (till exempel väggar),

men även att solens strålar påverkar uppvärmningsbehovet. De framtagna modellerna i studien ska ej ses som uteslutande och framtida forskning föreslås att både utveckla modellerna samt bekräfta de resultat som framställdes av denna studie.

En undersökning av modellernas beståndsdelar kommer fram till att energibehovet stämmer överens dugligt med detaljerade energikartläggningar. Därigenom föreslås även en koppling mellan modellens olika delar och relevanta renoveringsmöjligheter för att minska uppvärmningsbehovet och energikonsumtionen. Därigenom förväntas denna studie att kunna bidra till en minskad energikonsumtion i världen.

Table of contents

1. Introduction

The introduction of this study begins with providing both an academic and practical context. These contexts will motivate the use for system identification modelling as a tool to decrease energy consumption. The purpose and goals of the study are presented. Some delimitations of the thesis are considered. Lastly, a literature review that summarizes relevant academic writings on system identification and modelling energy systems in buildings are presented.

1.1 What is this study?

Heating in commercial and residential buildings account for about a quarter of global energy demand (IEA, 2020a). The respective number in Sweden is similar and further, heating accounts for the majority of the energy consumed within the building sector (Boverket, 2021). According to the IEA (2020b), carbon emissions related to heating, ventilation and air conditioning (HVAC) systems must decrease in order to reach set goals for sustainable development. Further, the IEA states that heating in particular is “not on track” to reach said goals. As carbon emissions are related to the energy consumption of HVAC systems, there are two ways of decreasing them (IEA, 2020b). The first being to move away from using carbon-intensive energy sources, such as fossil fuels. The second way is to reduce the amount of energy used, which is the interest of this thesis.

The energy consumed for heating is dependent on the type of heating system and heating demand. By retrofitting buildings with energy-efficient solutions, the heating demand can be decreased. Retrofitting solutions can be classified as either active or passive solutions. Active solutions depend on how the building is used, for example, heat scheduling or so-called smart systems. Passive solutions, on the other hand, relate to the building’s physical surroundings. These solutions depend on the design of the building and its physical properties, such as insulation, thermal bridges and window surfaces. Passive solutions are the main factors relevant in describing a building as an energy system with energy flows through the building envelope (Abel and Elmroth, 2016).

Buildings and HVAC systems are complex systems dependent on their use, internal design and external context (Abel and Elmroth, 2016). Understanding these systems for the purpose of identifying worthwhile investments that will contribute to energy efficiency can be difficult. It requires both an understanding of the measures invested in and the system they are deployed in, which is often complex. Therefore, modelling buildings as energy systems has been an active research field for decades (Drgoňa et al., 2020). Different tools have been developed and designed for various modelling purposes. Popular software often makes use of comprehensive thermodynamics to model the buildings’ energy systems (e.g. EnergyPlus, Modellica and TRNSYS) (Coakley et al., 2014). However, these models are inherently both complex and difficult to apply as they require detailed knowledge of the specific building.

Based on statistical modelling and machine learning, it possible to model a system through a methodology called system identification (Ljung and Glad, 2004). As machine learning has become more common, system identification has too. In essence, the methodology is to relate a set of samples corresponding to “inputs” and “outputs” of a system through a mathematical description. Therefore, it is not necessary with detailed knowledge of a system's inner workings to describe its behaviour (Ljung and Glad, 2004). However, if the physics that govern the system is known, the mathematical description can be directly related to the real physical system. This type of model is called a grey-box model and has been used to model buildings' energy systems (Coakley et al., 2014; Drgoňa et al., 2020). In general, the empirical data can be said to provide specificity and the integrated physics can be said to provide some generalisability between similar systems. Thus, grey-box modelling can be suitable for an overview assessment of the potential energy savings from retrofitting a building with passive solutions.

The hindrance when attempting to apply system identification methodology to empirically model a system is often a lack of sample data (Ljung and Glad, 2004). The company Mestro AB specialises in collecting and organising data of energy usage in buildings and external meteorological conditions are available as open data sets. Therefore, data similar to what has been used by research like Hietaharju et al. (2018), Jiménez et al. (2008), Lowry and Lee (2004), and Wu and Sun (2012) to model and determine the thermal properties of buildings, is available from Mestro AB. However, in these datasets measured indoor temperature is seldom available. There is hence a possibility to model buildings' heating demand by applying system identification. Such a model could then be used to develop a tool for analysing retrofit potential, generalisable to different buildings.

1.2 Purpose and goals

This thesis sets out to derive a grey-box model suitable for system identification of buildings' heating demand. The derived model should be suitable for physical interpretation and to provide insight into the building energy system to which it would be applied to. The model should mainly make use of data available through a service like that which Mestro AB offers.

The goals set by this thesis are thus to

- through physics develop a model framework suitable for system identification,
- extrapolate information about physical properties through analysis of the derived model, and
- identify the main contributions to heating demand through model simulation.

This would allow for a generalisable systematic method to analyse the potential of retrofitting and passive solutions in a complex thermal system based on commonly available data.

1.3 Delimitations

This thesis has some delimitations with regards to the modelling. Firstly, it considers a set amount of input that have been used by previous studies, see Hietaharju et al. (2018) and Wu and Sun (2012). It is therefore not exhaustive in its modelling with regards to physics. Secondly, the modelling process makes no distinction between walls and other heavy construction, such as roofs. This a common modelling simplification (Verbeke and Audenaert, 2018).

1.4 Literature review

The modelling of buildings' energy systems has been an active field for a long time. Hence today, there are many publications applying different models and methods to serve different purposes and types of systems. Literature reviews have been published which summarise methods and results created for different purposes. Coakley et al. (2014) provide the most comprehensive review regarding energy simulations of whole buildings. The review focused on energy simulations with the purpose of designing and optimising building's energy consumption. It concludes that accurate simulations for real-world building operation are difficult to achieve, however, they are improved by implementing measurements of the building in operation. Thus, motivating the use of what this thesis considers a grey-box methodology. As an extension, the paper provides an assessment of analytical and statistical methods that have been used by practitioners to develop a model both empirically and through physical insight.

Drgoňa et al. (2020) published a literature review regarding modelling buildings to implement active solutions, in particular, predictive control of HVAC systems, to reduce energy consumption. Prior to, and mentioned, in Drgoňa et al. (2020), Afram and Janabi-Sharifi (2014) reviewed modelling methods for HVAC systems. The reviewed literature seeks both to predict behaviour and to study the energy consumption of HVAC systems. Partly, this includes a review of the potential of different modelling techniques to accurately describe a system's thermal behaviour and states. The reviewed literature in Drgoňa et al. (2020) and Afram and Janabi-Sharifi (2014) provide a foundation for modelling HVAC systems and thermal behaviour in buildings.

Both Hietaharju et al. (2018) and Wu and Sun (2012) have published papers designing grey-box models to predict the indoor temperatures of buildings. The respective paper considers large buildings (such as schools, municipal buildings etc.) and office buildings. Their grey-box models are based on the thermodynamic energy balance through the building envelope, linearly incorporating time series of heating power, solar irradiation, outdoor and indoor temperature. Any energy sources or sinks not modelled are considered as noise. Both studies find that their model approaches are suited to be generalised and applied for long-term prediction of indoor temperature. In the case of Hietaharju et al. (2018), the original energy balance is developed further as an AutoRegressive-Moving-Average with eXogenous inputs (ARMAX)-model by using time delays and dynamic behaviours that have been documented in the literature previously. They implement physical properties and information from national building codes to base estimates of some relations and parameters in the model structure. Their study considers buildings located in Oulu and Jyväskylä, Finland, which are similar in climate to Sweden and the buildings studied in this thesis. Further, the paper finds that forecasted outdoor temperature was sufficient for accurate prediction models, it need not be measured on-site. Wu and Sun (2012) also developed their grey-box model as an ARMAX-model and relates to the thermodynamic energy balance.

Jiménez et al. (2008) provided a study to determine thermal properties of buildings by applying system identification methodology. The study examines simplified buildings and provides a method to extract information of thermal properties from the developed model.

Since data of indoor temperature is seldom available for individual buildings, an attempt to recreate the data as a system state was conducted during the project. The work is excluded as a main part of the thesis but is included in Appendix A with a literature review available for future studies.

2. Theory

The theory chapter of this thesis begins with a discussion of the concept of and different approaches to modelling in science. The chapter then highlights the field of system identification. This is done by describing the process of model discretisation, development of common model structures and lastly the process of adjusting and validating models. Followingly, the chapter provides a discussion on data pre-processing and dealing with a non-ideal reality. Finally, the chapter provides a description of the energy flows to create an interpretable model framework with.

2.1 What are models and why are we interested in them?

According to Coakley et al. (2014) buildings are considered complex systems and the thermal behaviour in a building is highly dependent on many external sources. Though complexity makes systems difficult to model accurately, that is simultaneously the purpose of modelling: to provide insight into a system that is too complex to understand in full (Ljung and Glad, 2004). As learnt from the literature review (see Afram and Janabi-Sharifi (2014), Coakley et al. (2014) or Drgoňa et al. (2020)), depending on the application and intended use of a model, it should be designed differently. Its design is governed by the complexity of the system, measurements available, linearity and intended use (for example, short term predictions for control or long-term simulations for design and analysis). It is therefore important to recognize that models are a representation of a real system, designed after a purpose and can highlight certain aspects or dependencies of a system; it is not the actual real system. Coakley et al. (2014) provides the view that modelling relies on the modeller's knowledge, experience, statistical expertise, engineering judgement as well as an iterative process of trial and error.

Models are never an exact representation of a physical reality, but a common idea is that an exact representation

$$y_{sys}(t) = y(t|\theta_0) \tag{1}$$

can be achieved given the exact parametrisation θ_0 and that the measured output

$$y(t) = y(t|\theta_0) + e(t) \tag{2}$$

is subject to a stochastic noise $e(t)$ (Ljung and Glad, 2004). Even though models are never exact, they can still represent system behaviour. Coakley et al. (2014) consider models based on physical laws to be useful to for predicting the behaviour of energy systems in complex buildings.

2.2 What different models are there?

Models can be divided into three paradigms: white-, black- and grey-box models (Drgoňa et al., 2020). A white-box model predicts a system's behaviour by applying a given set of physical laws (gravity, heat transfer, etc) acting under specific properties and conditions. A black-box model works oppositely. It is fully data-driven and infers the system properties solely from observations. A grey-box model is the result of combining the two paradigms. According to the reviews from Afram and Janabi-Sharifi (2014), Coakley et al. (2014) and Drgoňa et al. (2020), there are both benefits and challenges to each paradigm.

A white-box model can allow for generalisability and great understanding of the system. However, the model risks becoming difficult to develop and demands great insight on behalf of the modeller. This means that highly detailed information of many different systems aspects is necessary for producing a white-box model that can recreate a system's behaviour accurately. Information like such is seldom available and the white-box model will be sub-par in its performance and inadequate for its purpose. Still, there is a generalisability to white-box models as they are based on laws, even though they can be lacking in accuracy for a specific system.

A black-box model generally uses fewer parameters but is solely reliant on the observable input and output data. Thus, a pure black-box model can be accurate in the specific case but is not generalisable. This is the result of a black-box model being based purely on data from the specific system.

A grey-box model bases its structure on laws and relationships thought to be important for the system's behaviour, like a white-box model. The system's properties are derived by inference through observable input and output data, like a black-box model. As such, a grey-box model to some extent allows for the generalisability of a white-box model and the accuracy of a black-box model (Afram and Janabi-Sharifi, 2014; Drgoňa et al., 2020; Ljung and Glad, 2004).

System identification is a methodology used to infer system properties from input- and output data (Ljung and Glad, 2004). It has been used by researchers interested in simulations of building's energy demand, see for example Hietaharju et al. (2018) and Wu and Sun (2012). The methodology is commonly deployed in situations where plenty of data is available for developing a black or grey-box model, and a detailed physics-based model is deemed too costly and complex (Coakley et al., 2014; Ljung and Glad, 2004).

2.3 Continuous and discrete models

Physical laws and systems are often described as continuous-time models. For example, Newton's second law of motion $F = m \cdot a$ or the first law of thermodynamics $\Delta U = Q - W$. When data is measured, it is done at discrete points in time. It is therefore common to use discrete-time models in system identification (Ljung and Glad, 2004). Assuming that the sampling data is spaced out uniformly with sampling time T, a point in time can be related as proportional to the sampling time

$$t_k = kT, \tag{3}$$

where k is an integer. Uniform sampling allows for the use of the time-shift operator

$$qu(kT) = u\big((k+1)T\big);\ q^{-1}u(kT) = u\big((k-1)T\big) \tag{4}$$

(Ljung and Glad, 2004). Using the operator, polynomials can be used to relate terms of inputs and outputs at different points in time

$$P(q) = p_0 + p_1 q^{-1} + p_2 q^{-2} + \cdots + p_n q^{-n} \tag{5}$$

such that

$$P(q)u(t) = p_0 u(t) + p_1 u(t-1) + p_2 u(t-2) + \cdots + p_n u(t-n). \tag{6}$$

This enables the coherent use of statistical methods to estimate the parameters of the polynomial given any uniformly sampled dataset. The time-shift operator relates to the Z-transform such that it directly relates to the frequency-domain (Ljung and Glad, 2004). This means that what time-delays are included in the polynomials dictate what frequency content is captured and recreated by the model.

The transition from continuous to discrete representation has consequences for the information contained in the representation. For example, information of high frequency dynamics of the system may be lost in the sampling of data as it may be lost between sample points (Ljung and Glad, 2004). The extent to which the sampled data contains the information is dependent on the sampling time and specific system. Whether the sampled data is sufficient in information to describe the continuous system should be evaluated both before modelling as well as afterwards. A rule of thumb given by Ljung and Glad (2004) is that somewhere around 4 to 8 samples should be available to describe the slope of a system's step response. However, a building is not an ideal system, and its energy consumption depends on many inputs which all have their separate step responses. Judging the step response of a not yet identified such system is therefore difficult. A good sampling time can instead be based on previous knowledge. Both Hietaharju et al. (2018) and Wu and Sun (2012) developed well-functioning long-term prediction models using data with 1h sampling times.

2.4 Black-box models

Commonly occurring model structures within black-box system identification methodology stem from a family of models called Box-Jenkins models (Ljung and Glad, 2004). More specifically, the literature review found that AutoRegressive with eXternal-input (ARX)- and Output Error (OE)-model structures are commonly used to model buildings' energy systems for simulation purposes. ARMAX-models also belong to the Box-Jenkins model family but are typically used for prediction-models (see for example Hietaharju et al. (2018) or Wu and Sun (2012)). All models in the Box-Jenkins family are linear and describe a system using measurements of inputs and outputs (Ljung and Glad, 2004). The term black-box model stems from the fact that the models do not consider what the actual system is, it only cares about relating time-series measurements of signals to the output.

The equation for an ARX-model is written

$$A(q)y(k) = B(q)u(k) + e(k), \tag{7}$$

where

$$A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + \cdots + a_{\mathrm{n}_a} q^{-n_a} \tag{8}$$

and

$$B(q) = b_0 + b_1 q^{-1} + b_2 q^{-2} + \cdots + b_{\mathrm{n}_b} q^{-n_b}, \tag{9}$$

where a_n and b_n are scalar parameters. In the special case that $A(q) = 1$ the structure is called a finite impulse response (FIR)-model. The model parameters, a_n and b_n, are collectively denoted as θ and can be estimated through least-square estimates

$$\hat{\theta} = arg\min_{\theta} \sum_{i=\max(n_a, n_b, n_f)}^{n} \big(y(i) - \hat{y}(i|\theta)\big)^2 = arg\min_{\theta} J(\theta), \tag{10}$$

where the cost function J is minimised with respect to the parameters θ and $\hat{y}$ is the model output as expressed in (11).

The ARX-model can be described as linear regression

$$\hat{y}(k) = \varphi^T(k)\hat{\theta}, \tag{11}$$

where φ is the information vector (Ljung and Glad, 2004). The information vector should be considered as a column vector where each element corresponds to either an input or output sample such that the multiplication describes the structure in (7). Since the ARX-model can be described as linear regression it also means that, given a dataset, the parameters at the global minimum of the cost function J can be found through the normal equation

$$\hat{\theta} = (X^T X)^{-1} X^T Y, \tag{12}$$

where

$$Y = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(n) \end{bmatrix} \tag{13}$$

and

$$X = \begin{bmatrix} \varphi^T(1) \\ \varphi^T(2) \\ \vdots \\ \varphi^T(n) \end{bmatrix}, \tag{14}$$

where n is the number of samples available in the training dataset (Ljung and Glad, 2004).

OE-models are structured differently to ARX-models and are special in the Box-Jenkins family as the noise is not modelled (Ljung and Glad, 2004). This means that the difference between the measured and undisturbed output signal is modelled as just the noise source. The equation for an OE-model is

$$y(k) = \frac{B(q)}{F(q)} u(k) + e(k), \tag{15}$$

where $B(q)$ is eq. (9) and

$$F(q) = 1 + f_1 q^{-1} + f_2 q^{-2} + \cdots + f_{n_f} q^{-n_f}. \tag{16}$$

The model parameters are estimated through least squares estimates, see (10). However, the OE-model structure is not linear with regards to the parameters and cannot be described as linear regression. The normal equation can therefore not be used to estimate the parameters. Iterative gradient-based methods such as Newton-Raphson (described in Ljung and Glad (2004)) are used instead. These methods risk converging towards local minimum of the cost function J. Software designed for system identification, such as MATLAB's *System Identification Toolbox*, include methods to avoid local minimums.

2.5 Model simulation

To simulate a model is to give it inputs and calculate the outputs (Ljung and Glad, 2004). A model that represents the real system's dynamics can therefore be used to simulate system behaviour given some inputs. The output of an ARX-model is simulated as

$$\hat{y}_{sim}(k) = \big(1 - A(q)\big)\hat{y}_{sim}(k) + B(q)u(k) \tag{17}$$

and an OE-model as

$$\hat{y}_{sim}(k) = \frac{B(q)}{F(q)} u(k). \tag{18}$$

Note that simulation of both model types exclude any noise-terms present in (7) and (15).

2.6 How do you validate a model?

2.6.1 What makes a model valid?

A model simulation can be compared to measured data. This allows for the modeller to determine the model's performance. A valid model can, to some extent, be summarised as a simulation that is accurate to the measured output. If the modeller was able to design the model well, then the measured data and model output should appear similar. This raises the question of what can be considered similar. Therefore, the process of validation is important to model development.

To begin, it is important to consider that a model is a representation of reality (see chapters 2.1-2.3) and the simulation mirrors what is represented. It is therefore important to consider what dynamics that the model incorporates. A remark regarding model validation, given by Ljung and Glad (2004), is that sample data and model validity should consider operating conditions. Non-linearities often impact the behaviour of the real system, which renders both the data and identified system inadequate to describe the system under different operating conditions. Inadequate operating conditions will not be overcome by the process of model validation as the problem is inherent to the sampled data used to design the model. It may however be counteracted by developing models with knowledge of operating conditions in mind.

According to Saltelli et al. (2008), science and researchers involved with mathematics and physics may have become accustomed to being able to describe nature and phenomena accurately and unambiguously. The descriptions present themselves as inherent beauty and self-evident truths. Validating a correct model therefore becomes to validate reality (Saltelli et al., 2008). However, the more complex a system is, the more difficult it is to validate. As grey-box models consist of simplified or aggregated relationships fitted to observed data, the aim is not to validate an exact model of reality. Instead, it is to validate whether the model is good or not for its purpose. In accordance with this contrasting perspective, Saltelli et al. (2008, pp. 3) states that "… practitioners of modelling have come to live with the rather unpleasant reality that more than one model may be compatible with the same set of data or evidence". Further, the methods and criteria for validation vary depending on the purpose of the model. This view appears widely accepted within the current landscape of modelling buildings' energy systems (Afram and Janabi-Sharifi, 2014; Drgoňa et al., 2020). In the end, to validate a model is to examine whether it can be accepted and serves the purpose it was designed for. For this thesis, the stated purpose is to develop a model that is suitable for physical interpretation and that can provide insight into a building's energy system.

Through the process of validation, the model can be adjusted such that it serves its purpose best. Adjusting a model to well represent the system that it is designed after has been found to be an iterative process of trial and error based on a combination of knowledge, experience, expertise and engineering judgement (Drgoňa et al., 2020; Ljung and Glad, 2004). The process of validation includes to evaluate measurements and indicators which can provide insight into how best to adjust a model. In chapters **Fel! Hittar inte referenskälla.**-2.6.6 some relevant measurements used in system identification are presented.

2.6.2 Cross-validation and overfitting

In general, it is good practice to validate the model on a different dataset than that of which it has been fitted to (Ljung and Glad, 2004). To divide data into training and validation data is called *cross-validation*, see illustration in Figure 1. Dividing up data can help reveal an overfit to measured data by comparing the model to data it has not been directly fitted to.

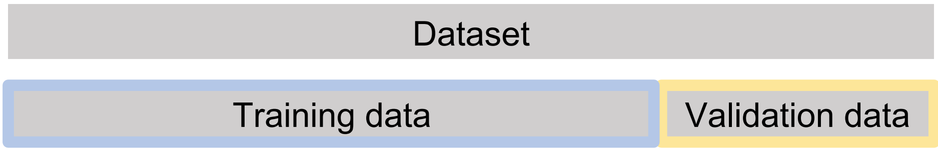

Figure 1. Illustration of division of a dataset into training and validation data.

It is possible to overfit a model that is too complex (Ljung and Glad, 2004). This may mean that the model has been fitted to data that was in actuality noise. Even a model of a complex system can still be too complex and overfit. An overfitted model is illustrated in Figure 2, alongside an underfitted and a well-fitting model. Figure 2 further shows how cross-validation can reveal the overfitted model.

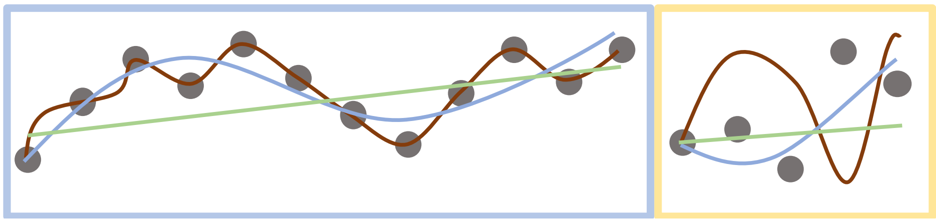

Figure 2. Illustration of an overfitted (brown), underfitted (green) and well-fitting (blue) model. The yellow box illustrates how an overfit model is detected by cross-validation.

To counteract overfitting a model, parsimony is considered as a desirable feature (Saltelli et al., 2008). This is also known as the principle of parsimony (or Occam's razor) and states that a model should always be developed to be as simple as possible whilst encapsulating sought after dynamics. In practice, this means that if two models are equal in performance, the simpler one should be chosen.

In order to decrease the risk of overfitting to a specific data set and operating conditions, a folding method can be applied. The method was previously used by Hietaharju et al. (2018) and entails that the dataset is divided into "folds". A fold is dedicated as validation data and the others are used for fitting the model. The folds are then rotated, such that a different fold is used as validation data. This method reduces reliance on a specific set of data. If the model development is evaluated on all folds, it allows the modeller to consider the full data set and consider the system during multiple operating conditions. The folding method is illustrated in Figure 3.

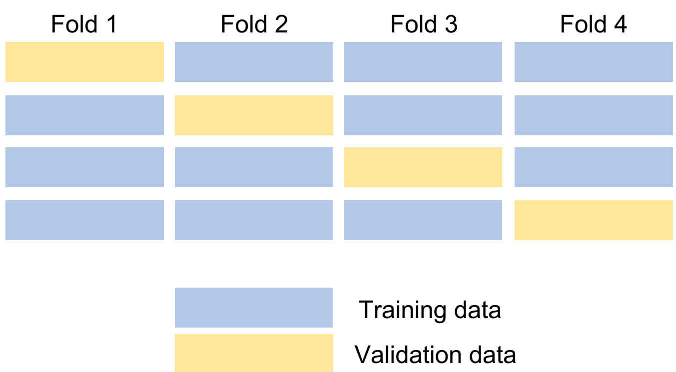

Figure 3. Illustration of the folding method.

2.6.3 Residual error and analysis

The residual error for a model at time t

$$\varepsilon(t|\hat{\theta}) = y(t) - \hat{y}(t|\hat{\theta}) \tag{19}$$

is defined as the difference between the measured output based on the *true system* y and predicted output of the modelled system $\hat{y}$ (Ljung and Glad, 2004). The residual error is the foundational measurement upon which many validation methods are based on.

It is possible to conduct a residual analysis in order to evaluate their relation to the developed model (Ljung and Glad, 2004). This analysis is conducted by looking at two aspects. Firstly, the residual error and the input signals should be independent. If they are not, there are likely system dynamics which have not been described and incorporated into the model. Independence can be analysed by estimating the covariance

$$\hat{R}_{\varepsilon u}(\tau) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon(t+\tau|\hat{\theta})u(t), \quad |\tau| \leq M \tag{20}$$

of the input signals and residuals. The covariance should be close to, and ideally be, zero. If the covariance is non-zero for a value τ_0 it is an indication that the model should include the term $u(t-\tau_0)$ (Ljung and Glad, 2004). Secondly, if the noise signal has been integrated as a part of the model, the autocovariance

$$\hat{R}_{\varepsilon}(\tau) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon(t+\tau|\hat{\theta})\varepsilon(t|\hat{\theta}) \tag{21}$$

of the residuals should be analysed. The residuals should also be independent of themselves.

2.6.4 Bias, variance and mean square error

Bias and variance error are two prevalent measures of error based on the residual error within modelling of buildings' energy systems (Afram and Janabi-Sharifi, 2014). The bias error

$$MBE = \frac{1}{N}\sum_{t=1}^{N} \varepsilon(t|\hat{\theta}) \tag{22}$$

is the mean residual error and represents a model's accuracy (Afram and Janabi-Sharifi, 2014; Kuhn and Johnson, 2013). It can loosely be described as a systematic error due to the model being too simple or trained on data insufficient to describe the real system. The variance error, or residual variance,

$$variance = E\left(\mathrm{y}(t|\hat{\theta}) - E\left(\mathrm{y}(t|\hat{\theta})\right)\right)^2 \tag{23}$$

represents a model's precision (Afram and Janabi-Sharifi, 2014; Kuhn and Johnson, 2013). A model that is too complex and overfits to training data will suffer a larger variance error when evaluated on validation data.

Both errors relate to the mean square error (MSE) as

$$MSE = E\varepsilon^2\big(t|\hat{\theta}\big) = \sigma^2 + MBE^2 + variance, \tag{24}$$

where σ^2 is an error termed "irreducible noise" and can never be reduced through modelling (Afram and Janabi-Sharifi, 2014; Kuhn and Johnson, 2013). Minimizing the MSE entails finding a balance between the bias and variance errors. Analysing the MSE and bias error therefore allows for the model's accuracy and total error to be evaluated.

2.6.5 Absolute error

The absolute error of a model

$$MAE = E\big|\varepsilon\big(t|\hat{\theta}\big)\big| \tag{25}$$

is similar to the bias error, except it is measured by the absolute value of the residual. The absolute error provides an indication of the error compared to each sample point (Afram and Janabi-Sharifi, 2014).

2.6.6 Correlation coefficient

The correlation coefficient is a measurement of how strongly associated the model output and the actual measurements are (Afram and Janabi-Sharifi, 2014). The correlation coefficient

$$CC = \frac{cov\big(y(t), \hat{y}(t)\big)}{\sigma_y \sigma_{\hat{y}}} \tag{26}$$

is calculated from the covariance and standard deviations of the measured output σ_y and the model output $\sigma_{\hat{y}}$ and ranges from 1 to -1. The correlation will be close to 1 if the model and the real system are closely positively correlated, -1 if closely negatively correlated, and near 0 if the model and real system are not correlated. Hietaharju et al. (2018) used the correlation coefficient to determine how well the dynamics of the real system have been captured.

2.7 Pre-processing data

This subchapter provides theoretical background and discussion of pre-processing data, as data often differs from the ideal case. The theory is largely based on the work of two authors (Ljung and Glad, 2004) which have been influential to the field of system identification. The pre-processing of data is seldom described in the publications mentioned in this thesis. It therefore assumed that these publications share a common theoretical framework that relates to pre-processing data.

2.7.1 Missing samples

Due to various reasons, data samples can be missing in both input and output signals. Ljung and Glad's (2004) primary recommendation is to use sections of data with as few missing datapoints as possible. However, in the case that there still are missing data points prevalent, it is recommended that the modeller fills in the missing values using either a predicted or an interpolated value. The best choice for filling in missing values is dependent on the characteristics of the data, which in turn impacts the effects of filling in missing values. Both the method of filling in missing values and the effect of doing so is dependent on context and the modeller's judgement.

2.7.2 Pre-filtering

The sampled data may be subject to high-frequency noise which hides the actual system's behaviour (Ljung and Glad, 2004). There may also be frequencies of larger importance for the described system which should be emphasized in the sampled data. Methods for removing high-frequency noise includes smoothing or filtering the data through a low-pass filter. The process of removing the high-frequency noise is reliant on engineering judgement. If a low-frequency drift appears prevalent it can be appropriate to apply a high-pass filter to the data. Engineering judgement should be applied in an analogous manner to a low-pass filter.

Any filtering to remove noise will inevitably also remove detailed information and removing a drift may remove a system behaviour from the data. Thus, pre-filtering is a weigh-off between removing noise and retaining details. The modeller's judgement also needs to determine what is to be deemed noise and what frequencies should be addressed to filter out the noise. As a continuous system is discretised this means that naturally some frequencies are emphasized over others as frequency content is limited by sampling-time (see chapter 2.3). Ljung and Glad (2004) reasons that modelling a non-linear system may benefit from not only applying a low- or high-pass filter, but a band-pass filter. That is because it allows the informational content of the data to focus on a range of frequencies deemed to be relevant for describing the behaviour of the non-linear system.

2.8 Linear time-invariant systems

The literature and theory for analysing a linear time-invariant (LTI) system is well established (Ljung and Glad, 2004; Oppenheim et al., 2013). As the name suggests, an LTI system is both linear and time-invariant. A linear system has properties of additivity and homogeneity. This means that if inputs are added or scaled, an analogous output is produced. This is written in a single statement as

$$au_1(k) + bu_2(k) \to ay_1(k) + by_2(k). \tag{27}$$

That a system is time-invariant means that its behaviour is not dependent on the absolute moment in time (Oppenheim et al., 2013). This means that a time-shift of the input signal $u(k-d)$ directly corresponds to a time-shifted output $y(k-d)$. Consider the input signals $u_1(k)$ and $u_2(k)$. Let

$$u_2(k) = u_1(k-d) \tag{28}$$

and the output of a system be defined as

$$u_1(k) \rightarrow y_1(k). \tag{29}$$

If the system is time-invariant, then the output should be independent on the point in time k such that

$$u_2(k) \rightarrow y_2(k) = y_1(k-d). \tag{30}$$

If this is not true, then the system is time-variant. In a time-variant system a time-shifted input signal $u(t-d)$ does *not* directly correspond to a $y(t-d)$. The output is dependent on the point in time k such that

$$u_2(k) \rightarrow y_2(k) \neq y_1(k-d). \tag{31}$$

A nonlinear time-variant system is more difficult to analyse since the well-established literature and theory regarding LTI systems are not fully applicable (Ljung and Glad, 2004; Oppenheim et al., 2013). According to Ljung (2001), LTI systems are ideal systems and do not exist in the real world. But many systems can be approximated to be, allowing for valuable insight by applying a theoretical understanding of LTI systems. This demands that care is given to choosing operating conditions, since the system may change with time (Ljung and Glad, 2004).

With regards to time-variance, operational conditions can be considered (Ljung and Glad, 2004). It may be known that the system G is dependent on the point time k in such a way that the system is approximately an LTI systems for sets of time. This is expressed as

$$G(k) = \begin{cases} G_1(k), k \in k_1 \\ G_2(k), k \in k_2 \end{cases}, \tag{32}$$

where the system G dependent on if the time k is a part of set k_1 or k_2. The sample data can be divided and used to identify the system as G_1 or G_2 dependent on the sample time k. Doing so allows for the use of methods for time-invariant systems. Having this knowledge of a system is not guaranteed, and especially if little is known about the system. Another approach is to evaluate the model on different data-sets, which can highlight the system's overall behaviour. Hietaharju et al. (2018) applies a folding method to avoid overly accommodating a small dataset and by extension system behaviour (see subchapter 2.6.2).

2.9 The physical laws that govern a building's heating system

Physical laws relate to the complex behaviour of buildings' thermal systems (Coakley et al., 2014). These laws include thermodynamic and heat transfer laws, which relate to the thermal behaviour of a building. The laws can be used to develop aggregated model frameworks for different purposes. Hietaharju et al. (2018) and Wu and Sun (2012) developed long-term prediction models and Jiménez et al. (2008) developed models for identifying thermal characteristics. Related laws have also been applied in less aggregated models that focus on detailed aspects of HVAC systems, documented by Afram and Janabi-Sharifi (2014) and Drgoňa et al. (2020).

The physical laws are ideal and will not describe a complex system perfectly. In modelling buildings' thermal energy systems, including time-delays have improved models and understanding based on physical laws, see (Hietaharju et al., 2018; Jiménez et al., 2008; Wu and Sun, 2012). Identifying these can increase the physical interpretability of later identified models.

2.9.1 The energy balance

A building is an open system where energy can both enter and exit. The system boundary is commonly defined as the building envelope. Based on the first law of thermodynamics, which state that energy can neither be created or destroyed, the flow of energy through a building's envelope can be described as a balance equation. This results in the energy balance

$$\partial E_{T_{ia}}(t) = P_{heat}(t) + P_{wall}(t) + P_{wd}(t) + P_{sun}(t) + P_{other}(t), \tag{33}$$

stating the change in the building's internal energy $\partial E_{T_{ia}}$ results from an energy flow from heating P_{heat}, convective heat transfer via walls P_{wall} and windows P_{wd}, solar heat radiation P_{sun}, and any additional heat sources, transfers or sinks P_{other}. Expanding on each term allows for insight into the physical relationships that dictate the energy balance. Additional heat transfers not included in (33) will not be considered.

2.9.2 Internal energy

The air inside of a building is considered an ideal gas under atmospheric pressure and constant volume (no work is being done). Therefore, the change of internal energy is expressed as being proportional to the change in indoor air temperature ∂T_{ia}

$$\partial E_{T_{ia}}(t) = \rho_i C_v V \partial T_{ia}(t) \tag{34}$$

dependent on the properties of indoor air density ρ_i, volume-specific heat capacity C_v, and the volume of air indoors. The density and heat capacity are well documented physical constants (tabulated in for example, Physics Handbook by Nordling and Österman (2013)) and considered generalisable under stable conditions. The volume of air is dependent on the size of the building. $\rho_i C_v V$ will be aggregated and written as K from now on.

2.9.3 Heat transfer through walls

The convective heat transfer via the building's walls

$$P_{wall}(t) = h_{wall_i} S_{wall} \left(T_{wall_i}(t) - T_{ia}(t) \right) \tag{35}$$

is proportional to the difference in indoor temperature T_{ia} and the temperature of the wall's inside surface T_{wall_i}, where h_{wall_i} is the convective heat coefficient of the surface of the wall's inside and S_{wall} is the surface area of the wall. Through further analysis of conductive and convective heat transfer, Wu and Sun (2012) relates the surface temperature to the outdoor temperature. The temperature at any point inside the wall is given by

$$T_{wall}(t) = \frac{T_{wall_o}(t) - T_{wall_i}(t)}{L_{wall}} x_{wall} + T_{wall_i}(t) \tag{36}$$

where T_{wall_o} is the temperature on the outside of the wall, L_{wall} is the wall's thickness, and x_{wall} is the coordinate point inside the wall. During stable conditions, the heat convection on both sides of the wall is equal to the heat conduction through the wall. The outside air temperature T_{oa} is then related to the indoor air temperature via

$$\begin{aligned} -k_{wall} \frac{\partial T_{wall}}{\partial x_{wall}} &= h_{wall_i} \left(T_{ia}(t) - T_{wall_i}(t) \right) \\ &= h_{wall_o} \left(T_{wall_i}(t) - T_{oa}(t) \right), \end{aligned} \tag{37}$$

where k_{wall} is the thermal conductivity coefficient of the wall. The surface temperature can then be described through a linear relationship as

$$T_{wall_i}(t) = P_1 T_{ia}(t) + P_2 T_{oa}(t) \tag{38}$$

and thus, the heat transfer through the wall (35) can be rewritten as

$$P_{wall}(t) = h_{wall_i} S_{wall} \left(P_2 T_{oa}(t) - (1 - P_1) T_{ia}(t) \right). \tag{39}$$

$(1 - P_1)$ from now on will just be written as P_1 for increased readability.

According to a review by Verbeke and Audenaert (2018) the heat transfer through walls display a time-delay due to the mass and material of the walls. The delay can be circa 12 hours and is an effect of thermal inertia. The thermal inertia also contributes to a dampening effect of the heat transfer. Tsilingiris (2004) finds that the time-constant ranges from a single hour to dozens, depending on both material and construction, indicating that the effects from thermal inertia is dependent on the building in question.

2.9.4 Heat transfer through windows

The convective heat transfer via the building's windows

$$P_{wd}(t) = h_{wd_i} S_{wd} \left(T_{wd_i}(t) - T_{ia}(t) \right) \tag{40}$$

is proportional to the difference in indoor temperature T_{ia} and the temperature of the windows' inside surface T_{wd_i}, h_{wd_i} is the convective heat coefficient of the surface of the windows' inside and S_{wd} is the surface area of the windows. Analogous to the heat transfer through walls, Wu and Sun (2012) suggests that the heat transfer through the windows (40) can be rewritten as

$$P_{wd}(t) = h_{wd_i} S_{wd} \left(P_4 T_{oa}(t) - (1 - P_3) T_{ia}(t) \right) \tag{41}$$

where $(1 - P_3)$ from now on will just be written as P_3.

As windows do not consist of massive materials as a building's walls, there should not be a time-delayed heat transfer associated as there will be through walls.

2.9.5 Solar irradiation

The solar irradiation of heat

$$P_{sun}(t) = I\psi(t) \tag{42}$$

is proportional to the solar irradiation ψ via the coefficient I. The solar irradiation effects the energy balance in two ways. One being by direct insolation through windows, which heats the inside air (Wu and Sun, 2012). Direct insolation should have an instantaneous and quick response. The other way is via heating building materials as an intermediary medium which heats the inside air via conduction and convection, and depends on the material's solar energy transmittance (Jiménez et al., 2008). Due to thermal inertia, this heat transfer should be associated with a delayed response.

2.9.6 Additional heat sources and sinks

The heat sources and sinks dependent on human activity also impact the heating demand in a building (Abel and Elmroth, 2016; Coakley et al., 2014). For example, the increased use of computers and printers, but also occupants opening windows and etc. Human activity is, however, difficult to predict through a law-based model because it is not possible to describe it through physical laws.

Different approaches could potentially be used to relate human activity to the energy usage. Yarbrough et al. (2015) applied a statistical analysis to determine a coincidence factor. Lowry and Lee (2004) proposed to use time-series of electricity consumption as a proxy for human activity. However, these approaches require additional modelling and is beyond the scope of this thesis. Hence, these heat sources and sinks will be treated as stochastic noise.

2.9.7 Discretised energy balance

The energy balance (33) can be rewritten as

$$\begin{aligned} K\,\partial T_{ia} = P_{heat}(t) + h_{wall_i} S_{wall}\big(P_2 T_{oa}(t) - P_1 T_{ia}(t)\big) \\ + h_{wd_i} S_{wd}\big(P_4 T_{oa}(t) - P_3 T_{ia}(t)\big) + I\psi(t) + P_{other}(t) \end{aligned} \tag{43}$$

and be discretised as

$$\begin{aligned} &K\big(T_{ia}(k) - T_{ia}(\mathrm{k}-1)\big) \\ &\qquad = P_{heat}(k) + h_{wall_i} S_{wall}\big(P_2 T_{oa}(k) - P_1 T_{ia}(k)\big) \\ &\qquad + h_{wd_i} S_{wd}\big(P_4 T_{oa}(k) - P_3 T_{ia}(k)\big) + I\psi(k) + P_{other}(k). \end{aligned} \tag{44}$$

3. Study objects and data pre-processing

In this chapter, the studied buildings are presented firstly. The pre-processing of the data is presented in the following subchapter. Lastly, a visual inspection which guides removal of outliers and division of data into folds is shown.

3.1 Study objects

Seven different buildings are considered. All studied buildings analysed are located in a coastal city in the southern part of Sweden. The buildings are all old, but the property manager has implemented energy efficient retrofitting solutions in the last 20 years. Building #3 was certified as energy efficient, less than 15 years ago, for consuming 25% less energy than the current norms. All buildings are constructed with heavy building materials of some kind (e.g. stone, brick, cement, etc.). All objects studied are used as offices and some have extra functionality as well. Building sizes can be seen as larger or smaller. Buildings #1, #4 and #7 are larger, between 4400m^2 to 5700m^2. Buildings #2, #3, #5 and #6 are smaller, between 1200m^2 to 1800m^2. The building information is summarised in Table 1.

Table 1. Building specifications and usage based on information from the landlord company.

Building	Decennia of construction	Energy efficiency certification	Area [m^2]	Usage
#1	1880	No	5700	Offices and exhibits
#2	1880	No	1500	Offices
#3	1880	Yes	1600	Offices
#4	1880	No	4500	Offices
#5	1920	No	1200	Offices and garage
#6	1880	No	1800	Offices and restaurant
#7	1920	No	4400	Offices

3.2 Pre-processing

The data used for system identification are series of measured outdoor temperature, solar irradiation, indoor temperature and heating power. The raw data goes through pre-processing to be used for modelling. The process is founded in the theoretical considerations presented in chapter 2.7. Finally, the data is evaluated and divided into folds for training and validation data sets prior to being used for system identification.

3.2.1 Meteorological data

Since all studied buildings are located in the same geographic area, the same meteorological dataset was be used for all models. All meteorological data comes from the Swedish Meteorological and Hydrological Institute (2021) open data. All measurements prior to the 11th of December 2020 have been validated by the same institution. The data that had yet to be validated is still thought to be accurate given that Hietaharju et al. (2018) used forecasted meteorological data.

Outdoor temperature data had very good coverage, except for a period during the month of December 2020, where the sampling frequency was only every 6th hour. Outdoor temperature is considered to be predictable intra-day. Therefore, interpolated values are deemed accurate enough for the purpose as they still convey the change in outdoor temperature intra-day.

The solar irradiation did not need particular treatment as data had full coverage and hourly sampling.

3.2.2 Indoor temperature

Measured indoor temperature data was available for buildings #1 and #2. The studied buildings contain multiple rooms. Therefore, indoor temperature is collected from multiple rooms as multiple data series in an ad-hoc manner. To apply the theoretical background of this thesis, the data series is combined into one aggregated series. This is done by aligning the data series based on their time stamps and then an aggregated series is created from the mean values of the ad-hoc series. This method does not take into account potential weighting factors such as room size when creating an aggregated series of the indoor temperature of a large building. Compared to previous studies, Hietaharju et al. (2018) makes use of less aggregated series whilst Wu and Sun (2012) do not aggregate the indoor temperature at all.

Indoor temperature had sampling frequency of every 10th minutes. However, some series were only sampled for a period of time. Values are only interpolated within the sampled period. Data outside of said period is ignored.

3.2.3 Heating power

The heating energy consumed in buildings is measured as the standing on a cumulative energy meter. The heating power was calculated as the difference between two measurements. The energy meter data series does not have full 100% coverage but does have a sampling frequency of every 10th minute which enables accurate interpolation intra-hour.

The energy meter data from building #5 was deemed noisy and not representative of actual behaviour of an HVAC-system and was therefore low-pass filtered. The data series were therefore put through a $0\text{-}0.6\ rad/s$ low-pass filter to remove what was considered high-frequency noise, as shown in Figure 4. No low-frequency drifts could be found which was not expected seasonal behaviour.

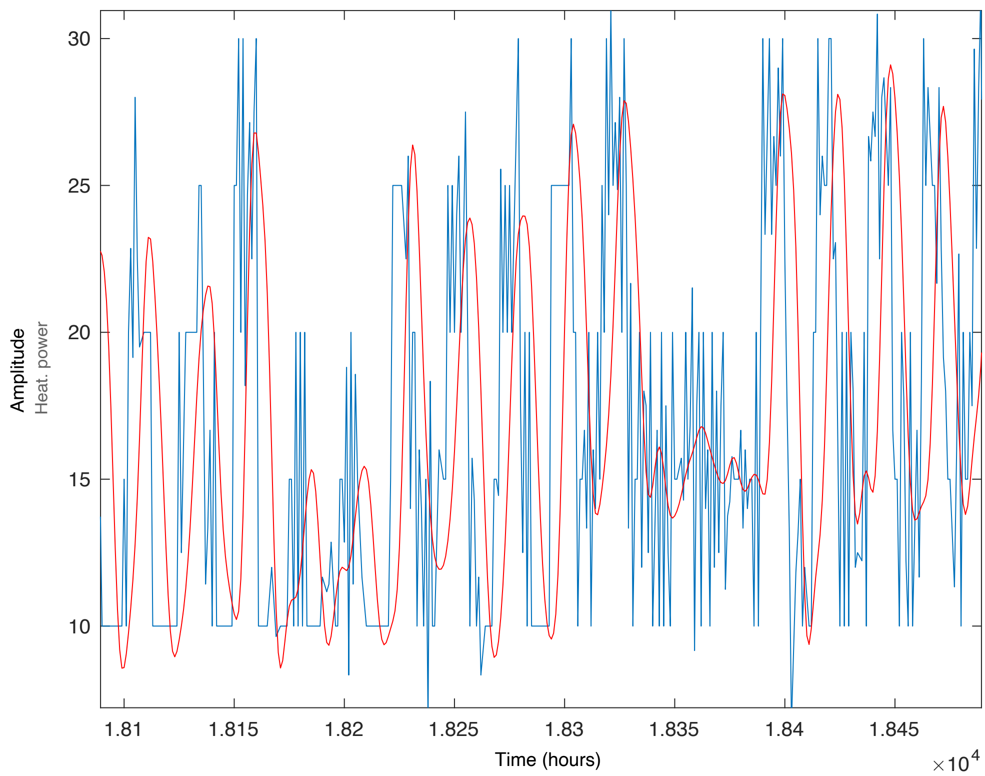

Figure 4. A section of samples from building #5. The original data (blue) has been low-pass filtered (red) to remove high-frequency noise.

3.2.4 Division of data and outlier removal

In Figure 5 the data from building #2 is visualised. From the visualisation, both potential outliers and suitable datasets for validation could be evaluated. Outliers were removed with MATLAB's function *rmoutliers,* which was adapted to achieve results that were sought after. Processed output data and folds are shown in Figure 6. In Table 2 it is shown how sections of data were divided into folds and if the data was filtered. All folds from the same building contain the same number of samples.

Only data from periods with heating demand was chosen for modelling. This means that the solar irradiation is low in magnitude for all data series used for modelling, which is expected due to the limited sunlight in Sweden during the winter half year. The choice to model solely on data sampled from the heating season was made to limit data to similar operating conditions. As the power used in a heating system cannot be negative, there is a threshold which makes the system nonlinear in the transition between heating demand and no heating demand.

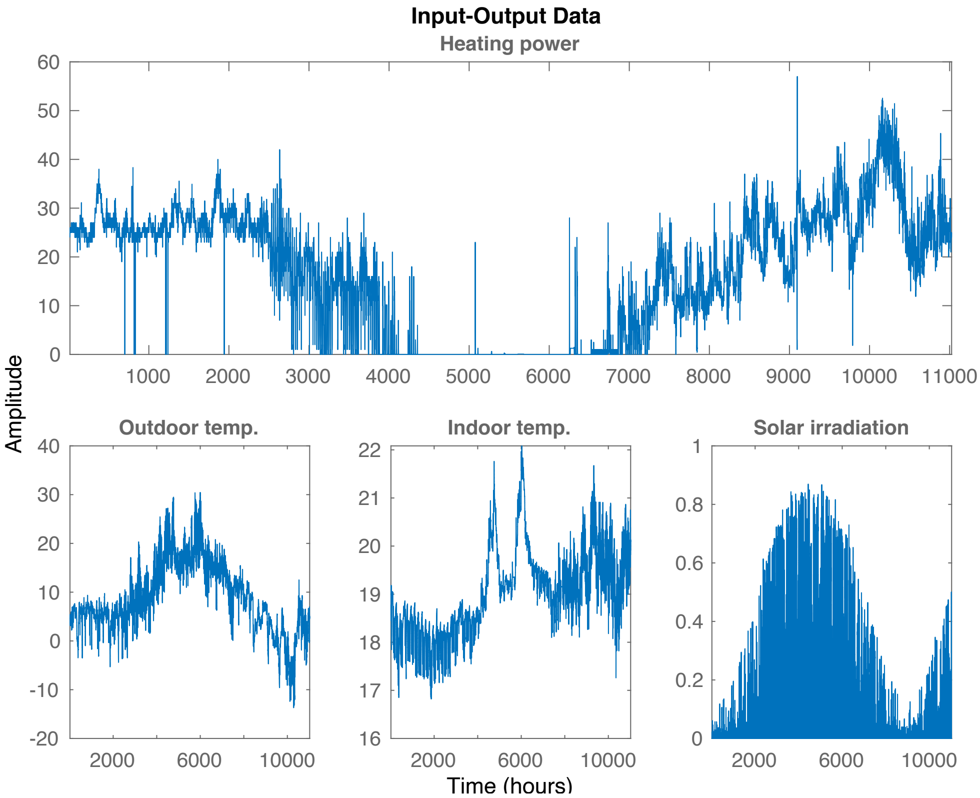

Figure 5. Unprocessed data for building #2. The series contains hourly measurements from 12.12.19 to 15.03.21.

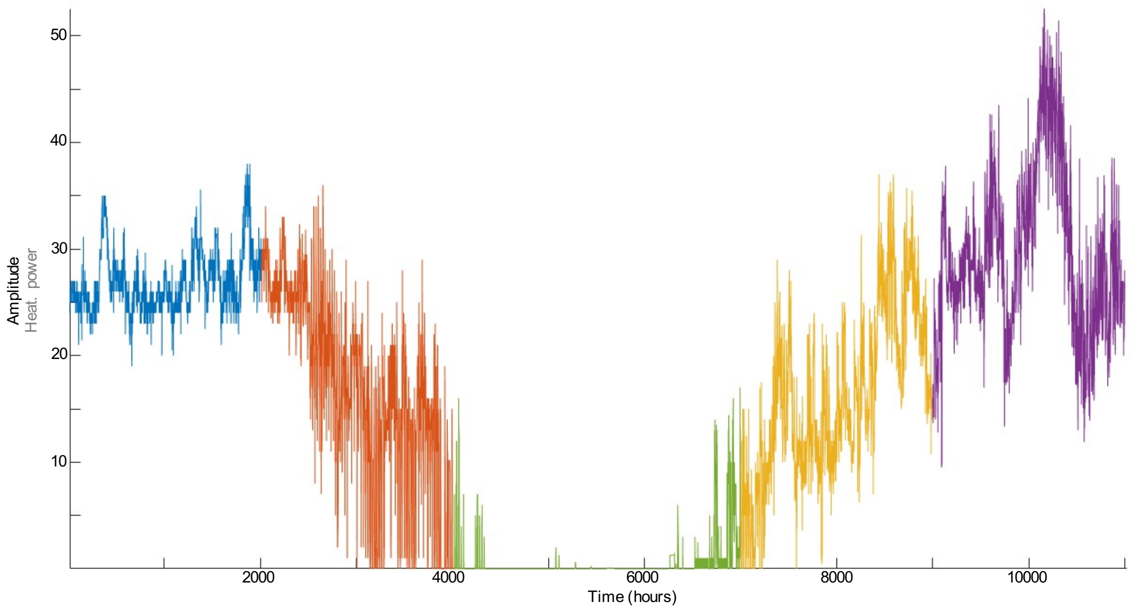

Figure 6. Processed output data from building #2. Outliers have been removed and the data has been split into folds, which is represented by different colours. The green marked data (middle section) is excluded from modelling as it is not heating season. The data contains hourly measurements from 12.12.19 to 15.03.21.

Table 2. How each building's data was divided into folds as well as if it was filtered.

Building	Sampling period	Fold #1 [Sample hour]	Fold #2 [Sample hour]	Fold #3 [Sample hour]	Fold #4 [Sample hour]	Fold #5 [Sample hour]	Filtered
#1	13.11.19-15.03.21	1-2000	2001-4000	7001-9000	9001-11000	-	No
#2	12.12.19-15.03.21	16-2015	2016-4015	7001-9000	9001-11000	-	No
#3	01.10.19-15.03.21	1-2000	2001-4000	4001-6000	8701-10700	10701-12700	No
#4	21.09.19-15.03.21	1-2000	2001-4000	4001-6000	8981-10980	10981-12980	No
#5	04.01.19-15.03.21	1-2000	7001-9000	9001-11000	17201-19200	-	Yes
#6	25.09.18-15.03.21	1-2500	2501-5000	9001-11500	11501-14000	19001-21500	No
#7	25.09.19-15.03.21	1-2000	2001-4000	4001-6000	8701-10700	10701-12700	No

In Table 3 correlations between validation data series are shown. It can be seen that outdoor temperature is the most highly correlated with heating energy for all buildings and that all inputs, except for indoor and outdoor temperature, are weakly correlated with each other. Note that only buildings #1 and #2 have measured indoor temperature available.

Table 3. Correlation between data series used for modelling.
ψ denotes solar irradiation.

Building	Output (P_{heat})			Outdoor temperature (T_{oa})		Indoor temperature (T_{ia})
	T_{oa}	T_{ia}	ψ	T_{ia}	ψ	ψ
#1	-0.819	-0.436	-0.200	0.470	0.339	0.338
#2	-0.863	0.065	-0.325	-0.208	0.347	-0.217
#3	-0.895	-	-0.375	-	0.382	-
#4	-0.681	-	-0.292	-	0.322	-
#5	-0.593	-	0.025	-	0.189	-
#6	-0.765	-	-0.235	-	0.305	-
#7	-0.867	-	-0.340	-	0.345	-

4. Establishing a model framework

This chapter presents the establishment of a model framework based on the energy balance and data available. A subchapter is dedicated to recreating indoor temperature as a system state, which was begun due to a lack of data and later cancelled. Finally, the framework is developed into both ARX- and OE-models.

4.1 Modelling based on the energy balance

In order to apply system identification methodology based on the energy balance (see (44)) for the purpose of this thesis, further considerations are needed. One aspect is that there are similarly structured heat convection relationships that drive the heat convection through walls and windows respectively. These must be separable to provide insight into how the model relates to heat convection through different structures such as walls and windows. Another aspect is that the indoor temperature is seldom measured. Alternate ways to incorporate the indoor temperature is therefore considered.

4.1.1 Heat convection relationships

P_1, P_2, P_3 and P_4 (see (44)) are all constants based on physical building properties and stem from a relation of the surface temperatures of walls and windows. If these values cannot be specified, the respective regressors representing the convective heat exchange through the window and wall will be indistinguishable. Detailed information of buildings may allow some certainty and to decide these values but are not available. This information would make the expressions for heat transfer through the wall (39) and window (41) linearly independent and separable. The methodology used in Jiménez et al. (2008) would be applicable to derive information of the properties, but is beyond feasibility for this thesis.

Another option is to analyse heat convection relationships by exploring time delays. This is motivated by observations of thermal inertia, see for example (Tsilingiris, 2004, 2002; Verbeke and Audenaert, 2018). However, it cannot be expected to find exact same time-delays reoccurring in all buildings. The time-delays are dependent on the construction. Highlighting specific time-delays is easily done within ARX-models by adding specific regressors to the information vector φ (see chapter 2.4).

Given the data available of the buildings, exploring time-delays is more feasible. It can also be done independently of knowledge of the technical specifications. The parameters P_{1-4} will therefore be excluded during model development and considering relevant time-delays will be in focus.

4.1.2 Unmeasured indoor temperature

Author's note: Measured indoor temperature was not available until a while into the thesis project.

The indoor temperature is not always available, but it must be related to the energy balance for a functioning model to be developed in this framework. Considering the energy balance, the indoor temperature can be either constant or changing. Arguing the case that as little heating as possible is used, it should be likely that the indoor temperature is kept as low as is legally allowed. The Swedish Work Environment Authority (2020) regards 20°C to be the lower limit on what a reasonable temperature is for offices, schools and similar workplaces.

4.1.2.1 Constant indoor temperature

If the indoor temperature is considered to be constant, then $T_{ia}(k+1) = T_{ia}(k) = T_{ia}$ and the model that follows from the energy balance (see (44)), should be written as

$$0 = P_{heat}(k) + h_{wall_i} S_{wall}(T_{oa}(k) - T_{ia}) + h_{wd_i} S_{wd}(T_{oa}(k) - T_{ia}) + I\psi(k) + P_{internal}(k). \tag{45}$$

It can further be rewritten as

$$P_{heat}(k) = -\left(h_{wall_i} S_{wall} + h_{wd_i} S_{wd}\right)(T_{oa}(k) - T_{ia}) + I\psi \sin\theta_s(k) + P_{internal}(k) \tag{46}$$

given that heating is considered the output signal in the system.

4.1.2.2 Observing indoor temperature

Another option could be to attempt to recreate the indoor temperature as a bi-process of the modelling. An attempt was made to observe indoor temperature whilst estimating a state-space model using a ES-RLS algorithm proposed by Ding (2014). The reader is referred to Appendix A for a description of the algorithm, process, results and short analysis. Due to inconclusive results and the later availability of measured indoor temperature, it is not seen fit to include simultaneous estimation to a large extent in this thesis. However, it is thought to be valuable for further studies and is therefore included in Appendix A.

4.2 Established model framework

Based on the energy balance (see (44)) regressors were formed. Data was further modified and combined in a manner that results in regressors which are related to the energy balance. The ARX model framework is established as

$$A(q)y(k) = B_0\varphi_0 + B_1(q)\varphi_1(k) + B_2(q)\varphi_2(k) + B_3(q)\varphi_3(k), \tag{47}$$

where φ_0 is a constant and the regressors φ_{1-3} represent $T_{ia}(k) - T_{ia}(k-1)$, $T_{oa}(k) - T_{ia}(k)$ and $\psi(k)$ respectively. The information vector φ used in eq. (12) is constructed as a combination of the output and regressors, for example

$$\varphi(k) = \begin{bmatrix} -y(\mathrm{k}-1) \\ -y(\mathrm{k}-2) \\ \varphi_0 \\ \varphi_1(\mathrm{k}-2) \\ \varphi_2(k) \\ \varphi_2(\mathrm{k}-3) \\ \varphi_3(k) \end{bmatrix}, \tag{48}$$

dependent on the relevant model structure.

OE-models use the same regressors, but a different structure

$$y(k) = B_0\varphi_0 + \frac{B_1(q)}{F_1(q)}\varphi_1(k) + \frac{B_2(q)}{F_2(q)}\varphi_2(k) + \frac{B_3(q)}{F_3(q)}\varphi_3(k), \tag{49}$$

with both nominator B_i and denominator F_i polynomials. OE-models have to be fit through gradient descent algorithms (see chapter 2.4) and therefore does not make use of an information vector.

5. Model development

This chapter first presents the resulting development of ARX- and OE-models using the presented system identification methodology. The first part focuses on quantitative measurements and residual analysis. A visualisation of model simulation is then presented and used as a tool to evaluate the developed models' performance. The chapter will present the resulting model performance for all studied buildings and highlight the process of some. This will result in a continuous comparison between the development of ARX- and OE-models.

5.1 ARX-model development using varying indoor temperature

Table 4 shows that by implementing the change in indoor temperature in ARX-models, little improvement or even increased MSE:s are obtained. By comparing Table 4 to Table 9 (presented in Appendix B) it can be seen that replacing the measured indoor temperature with a constant decreases MSE:s for all folds in both building #1 and #2. In Figure 7 and 8, a residual analysis of the residuals and φ_1 regressor (the change in indoor temperature, see (48)) is shown for models of buildings #1 and #2 respectively. The cross-correlation is almost unchanged, independent of the model structure.

Table 4. MSE of model simulations for buildings #1 and #2, implementing different structures of the B_1 polynomial to model the change in indoor temperature. The other polynomial structures are kept as $B_2(q) = b_{2,0}$ and $B_3(q) = b_{3,0}$ for all structures.

Building	$A(q)$	$B_1(q)$	MSE			
			Validation fold			
			1	2	3	4
#1	1	-	375.22	1941.15	781.34	1140.82
	1	$b_{1,0}$	375.33	1959.23	782.39	1148.47
	1	$b_{1,0} + b_{1,1}q^{-1}$	375.18	1971.74	783.32	1152.38
	$1 + a_1q^{-1}$	-	317.76	1325.65	707.66	992.64
	$1 + a_1q^{-1}$	$b_{1,0}$	319.36	1347.10	733.29	991.80
	$1 + a_1q^{-1}$	$b_{1,0} + b_{1,1}q^{-1}$	320.21	1352.70	738.90	992.10
#2	1	-	60.41	45.05	37.50	31.75
	1	$b_{1,0}$	60.60	45.06	37.52	31.75
	1	$b_{1,0} + b_{1,1}q^{-1}$	59.65	45.05	37.18	32.60
	$1 + a_1q^{-1}$	-	50.33	27.64	41.58	28.55
	$1 + a_1q^{-1}$	$b_{1,0}$	50.71	27.62	41.64	28.57
	$1 + a_1q^{-1}$	$b_{1,0} + b_{1,1}q^{-1}$	50.37	27.69	41.43	28.57

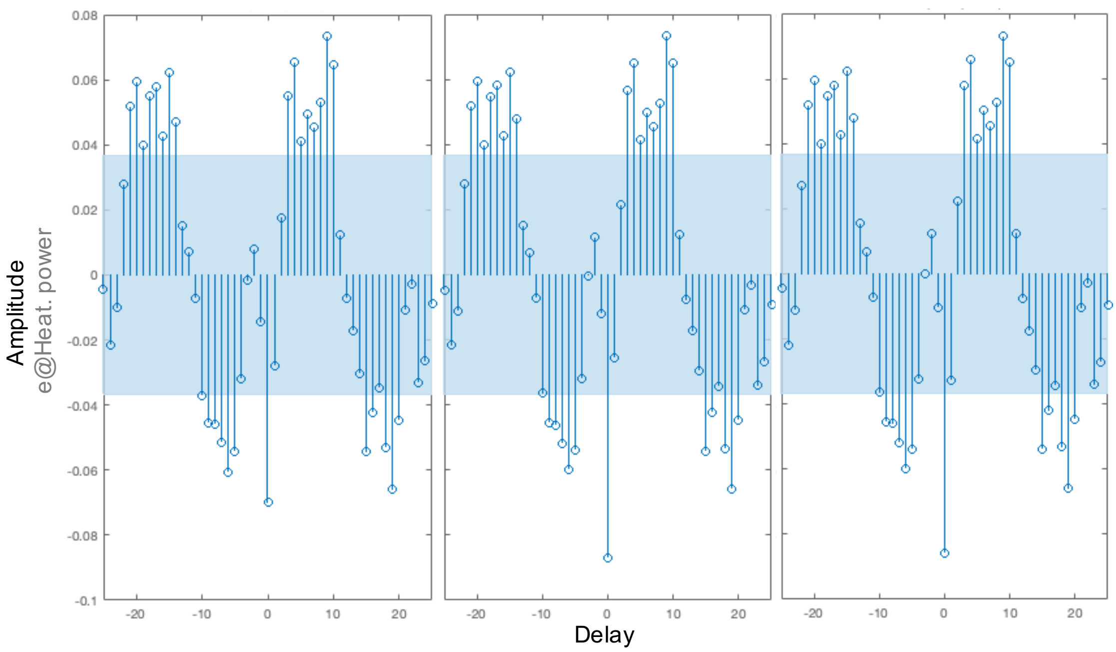

Figure 7. Cross-correlation between residuals and the φ_1 input signal for building #1, fold 4. All model structures contain $A(q) = 1, B_2(q) = b_{2,0}, B_3(q) = b_{3,0}$. B_1 varies between $b_{1,0}$, $b_{1,0} + b_{1,1}q^{-1}$ and $b_{1,0} + b_{1,1}q^{-1} + b_{1,2}q^{-2}$, depicted from left to right. The blue-shaded area depicts the 95% confidence interval.

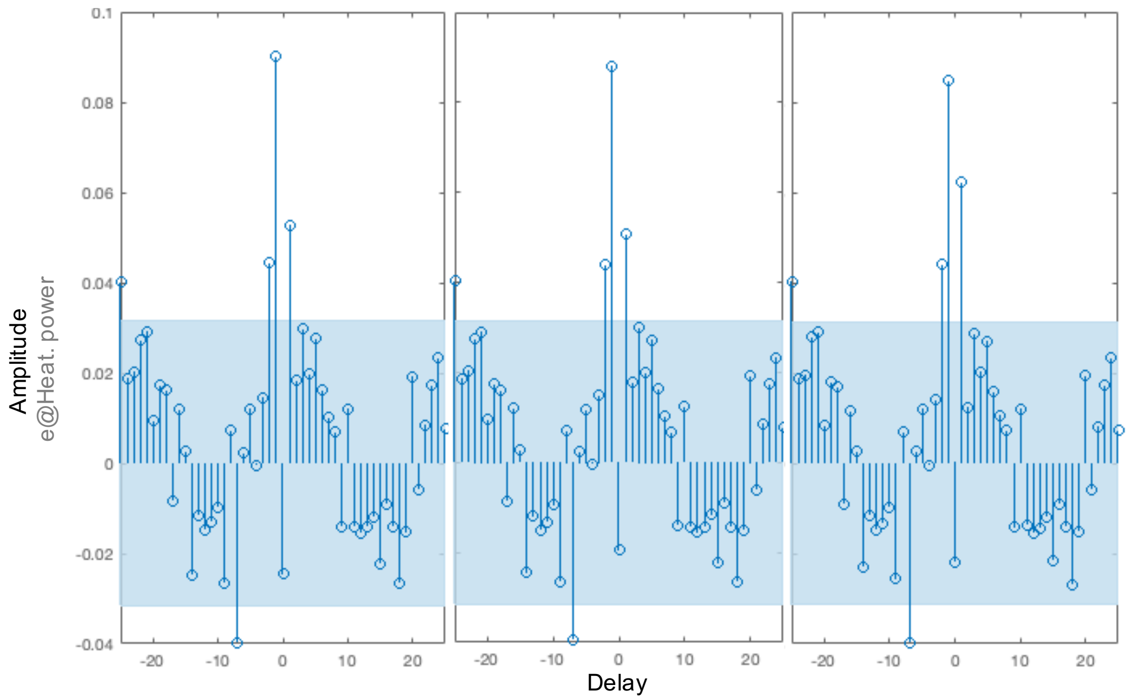

Figure 8. Cross-correlation between residuals and the φ_1 input signal for building #2, fold 4. All model structures contain $A(q) = 1, B_2(q) = b_{2,0}, B_3(q) = b_{3,0}$. B_1 varies between $b_{1,0}$, $b_{1,0} + b_{1,1}q^{-1}$ and $b_{1,0} + b_{1,1}q^{-1} + b_{1,2}q^{-2}$, depicted from left to right. The blue-shaded area depicts the 95% confidence interval.

5.2 ARX-model development using constant indoor temperature

The results presented onwards were obtained with the indoor temperature set to constant 20°C, considered as an average temperature in offices. Keeping the indoor temperature constant enabled model development based on a larger set of data from different buildings. The ARX-model development was led by conducting a residual analysis. This allowed for a statistical analysis comparing the dynamics of the modelled and real system.

In Figures 9 and 10 the residual analysis of a simple and a more complex (and better performing) ARX-model structure, applied to building #1, fold 1, is shown. The simpler model is structured as

$$\hat{y}(t) = \theta_2(t)\big(b_{2,0}\big) + \theta_3(t)\big(b_{3,0}\big) \tag{50}$$

and the more complex one as

$$\begin{aligned}\hat{y}(t)(1 + a_1q^{-1} + a_2q^{-2} + a_{12}q^{-12})& \\ = \theta_2(t)\big(b_{2,0} + b_{2,1}q^{-1} + b_{2,8}q^{-8} + b_{2,10}q^{-10}\big)& \\ + \theta_3(t)\big(b_{3,0}\big).&\end{aligned} \tag{51}$$

The results from the residual analyses provided guidance through model development, as they provide an analysis of time-delayed signals relate to the dynamics of the modelled system.

In Figure 11, the residual analysis for the model with the lowest MSE found for building #1, fold 1, is shown. It can be seen that the residuals are mostly uncorrelated with the inputs, with some exceptions regarding autocorrelation and cross-correlation with solar irradiation. The correlations found with the residual analysis were not always translatable to linear dependencies that could be encapsulated by the ARX-model framework. For example, expanding the B_3 polynomial with time-delays around 12 hours did not make for a better performing model; neither by looking at quantitative measures nor residual analysis.

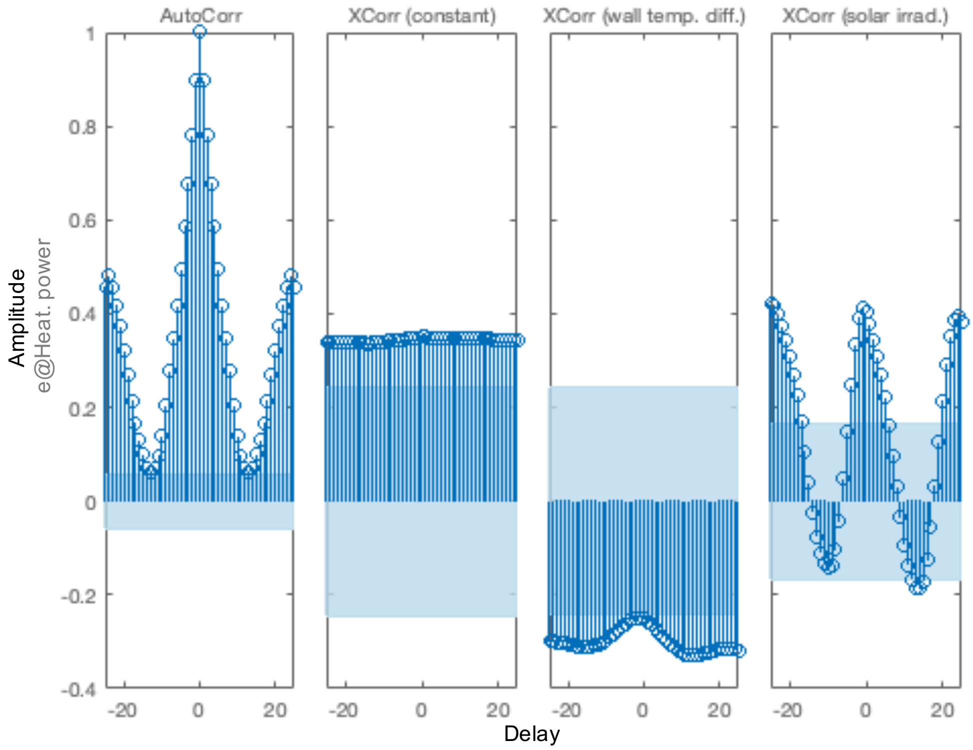

Figure 9. Residual analysis of the simplest ARX-model structure: $A(q) = 1$, $B_2(q) = b_{2,0}$ *and* $B_3(q) = b_{3,0}$, *for building #1, fold 1. The blue-shaded area depicts the 95% confidence interval.*

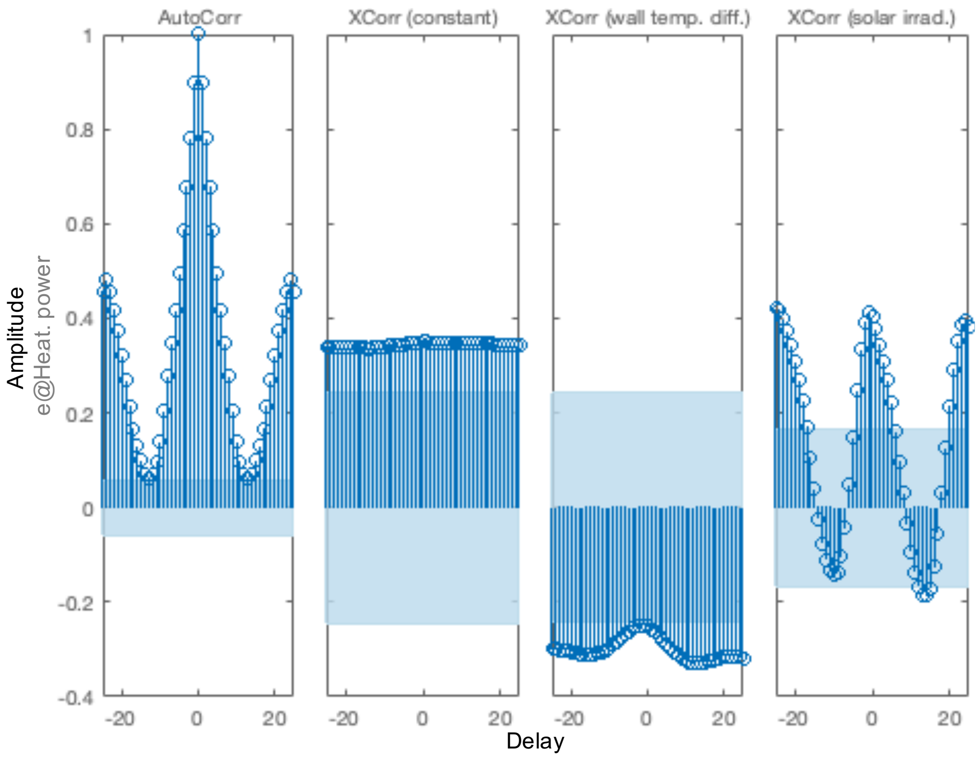

Figure 10. Residual analysis of mid-performing ARX-model structure: $A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + a_{12} q^{-12}$, $B_2(q) = b_{2,0} + b_{2,1} q^{-1} + b_{2,8} q^{-8} + b_{2,10} q^{-10}$ *and* $B_3(q) = b_{3,0}$, *for building #1, fold #1. The blue-shaded area depicts the 95% confidence interval.*

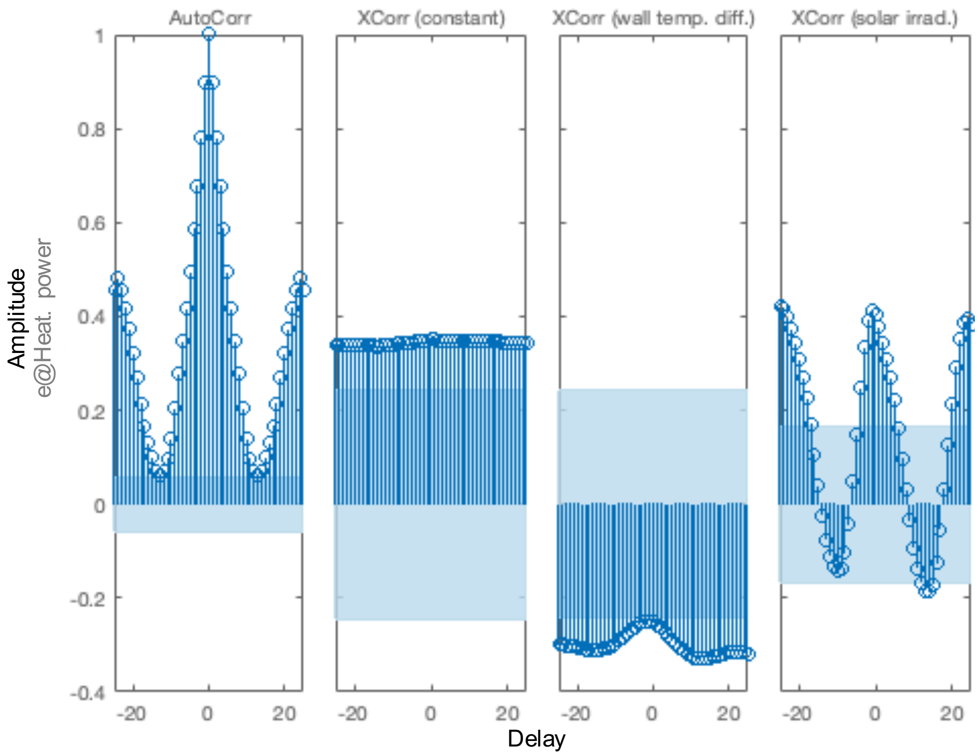

Figure 11. Residual analysis of the ARX-model structure with the lowest MSE found: $A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + a_{24} q^{-24}$, $B_2(q) = b_{2,0} + b_{2,1} q^{-1} + b_{2,2} q^{-2} + b_{2,12} q^{-12}$ *and* $B_3(q) = b_{3,0} + b_{3,8} q^{-8}$, *for building #1, fold 1. The blue-shaded area depicts the 95% confidence interval.*

Table 9 (in Appendix B) shows the best ARX-models found based on MSE of simulations. Different model structures were found to be the best for different buildings and folds. The best model structures found are generalised and summarised in Table 5. Note that the groupings of time-delays are coherent for most folds, with some exceptions, notably building #2, fold 4, and building #6, fold 3.

Table 5. Summary of contents of the three ARX-model structures found with the lowest MSE for each studied building and validation fold. Short delays (S.d.:s) are all time delays 5 hours or shorter, and long delays (L.d.:s) are all delays that are 6 hours or longer. The shade indicates if the structure was a part of the best performing models, dark = all three; white = none. 'X' denotes that the model structure found with the lowest cross-validation MSE included the grouping of time-delays.

Building	Val. fold	S.d. AR [$\hat{y}$]	L.d. AR [$\hat{y}$]	S.d. temp. diff. [φ_2]	L.d. temp. diff. [φ_2]	S.d. solar irr. [φ_3]	L.d. solar irr. [φ_3]
#1	1	X	X	X	X	X	X
	2	X	X	X	X	X	
	3	X	X	X	X	X	
	4	X	X	X	X	X	X
#2	1	X	X	X		X	
	2	X	X	X		X	
	3	X	X	X		X	
	4	X		X		X	
#3	1	X	X	X	X	X	
	2	X	X	X	X	X	X
	3	X	X	X	X	X	X
	4	X	X	X	X	X	
	5	X	X	X	X	X	
#4	1	X		X	X	X	X
	2	X		X	X	X	X
	3	X		X	X	X	X
	4	X	X	X	X	X	X
	5	X		X		X	X
#5	1	X	X	X		X	X
	2	X		X	X	X	X
	3			X		X	
	4	X		X		X	
#6	1	X	X	X	X	X	X
	2	X	X	X	X	X	X
	3			X		X	
	4	X	X	X	X	X	
	5	X	X	X	X	X	X
#7	1	X	X	X		X	
	2	X	X	X	X	X	X
	3	X	X	X		X	
	4	X	X	X	X	X	X
	5	X	X	X	X	X	X

5.2.1 Residual analysis of developed ARX-models

Residual analysis was also used to evaluate the performance of the models developed. The ARX-models developed which had the lowest MSE can be found in Table 9 (presented in Appendix B). In a good model structure, the residuals should ideally be uncorrelated to any signals. The residual analysis of a model structure depended on the fold. This indicates that the heating system behaves differently both dependent on building and the season. Some structures had better residual analyses overall, indicating a generalisability between multiple folds.

5.2.1.1 Residual analysis of ARX-models of building #1

In Figures 12-15 the residual analysis of the model structures with the lowest MSE for building #1 is presented. One model structure is presented per figure. The residual analysis is shown for all folds.

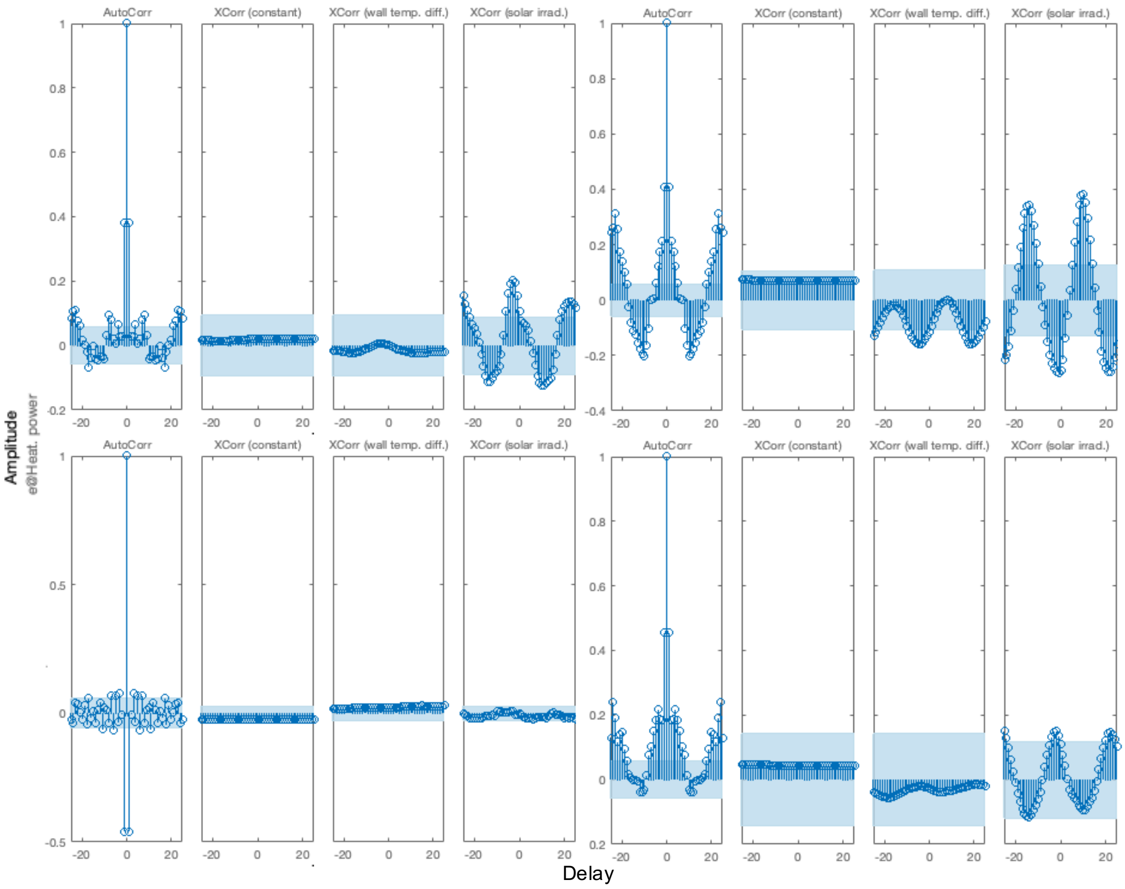

Figure 12. Residual analysis of simulations of all validation folds, building #1, with ARX-model structure $A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + a_{24} q^{-24}$, $B_2(q) = b_{2,0} + b_{2,1} q^{-1} + b_{2,6} q^{-6} + b_{2,7} q^{-7} + b_{2,8} q^{-8} + b_{2,20} q^{-20}$ *and* $B_3(q) = b_{3,0} + b_{3,10} q^{-10}$. *The structure was, measured by MSE, found to be the best one for fold 1. The blue-shaded area depicts the 95% confidence interval.*

With regards to autocorrelation, all model structures presented have notable autocorrelation at time delays 0 and 1, for all folds. Some structures have autocorrelation with other time delays as well.

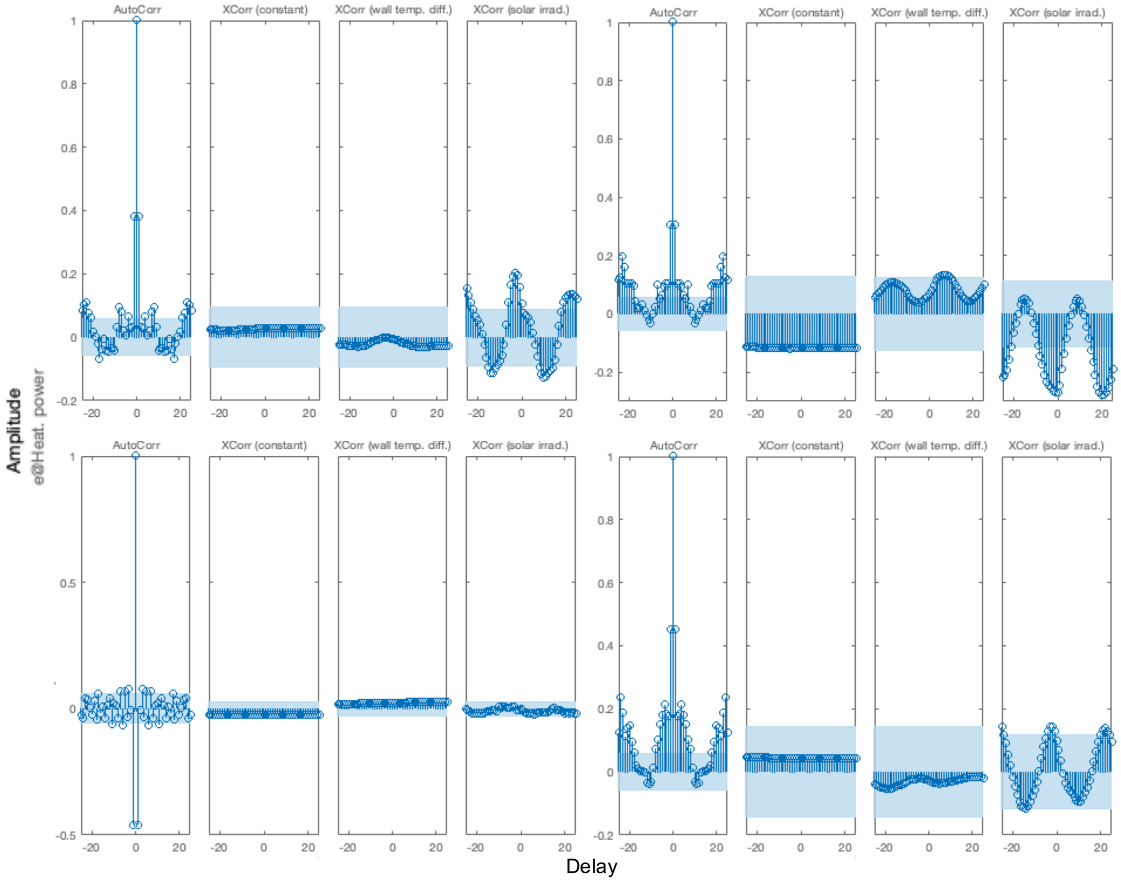

Figure 13. Residual analysis of simulations of all validation folds, building #1, with ARX-model structure $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_{24}q^{-24}$, $B_2(q) = b_{2,0} + b_{2,1}q^{-1} + b_{2,6}q^{-6} + b_{2,7}q^{-7} + b_{2,8}q^{-8} + b_{2,20}q^{-20}$ *and* $B_3(q) = b_{3,0}$*. The structure was, measured by MSE, found to be the best one for fold 2. The blue-shaded area depicts the 95% confidence interval.*

With regards to the constant term φ_0, no structure presented has correlated residuals, except the structure best fit for fold 4 (Figure 15). With regards to the temperature difference φ_2, the structures best fit to fold 1 and 3 (Figures 12 and 14) have a notable cross-correlation for fold 2.

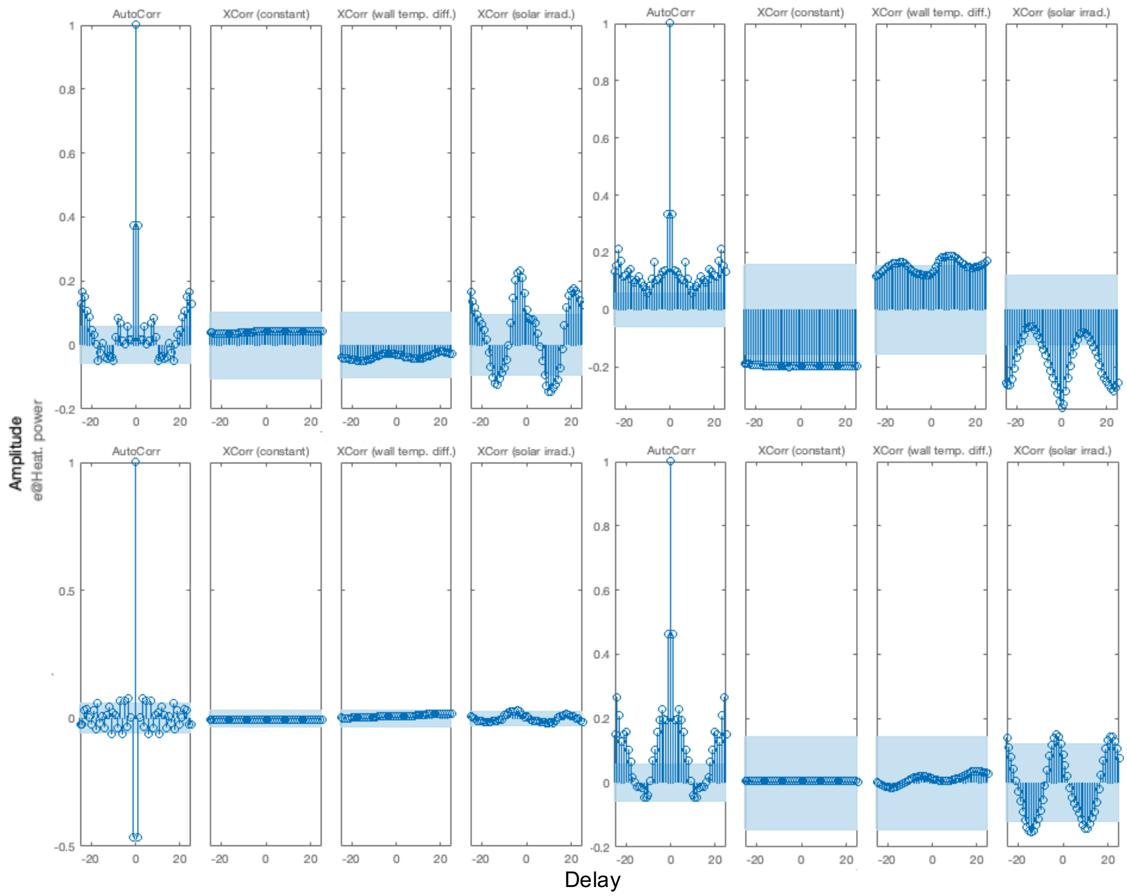

Figure 14. Residual analysis of simulations of all validation folds, building #1, with ARX-model structure $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_{24}q^{-24}$, $B_2(q) = b_{2,0} + b_{2,1}q^{-1} + b_{2,6}q^{-6} + b_{2,7}q^{-7} + b_{2,8}q^{-8}$ and $B_3(q) = b_{3,0}$. The structure was, measured by MSE, found to be the best one for fold 3. The blue-shaded area depicts the 95% confidence interval.

Finally, with regards to the solar irradiation φ_3, all structures have correlated residuals, for all folds. The strength of the cross-correlation between the residuals and the solar irradiation varies. The structures that fit fold 1 and 2 the best (Figures 12 and 13) have the lowest cross-correlation between the residuals and the solar irradiation overall. The structure that best fit fold 2 has a simpler polynomial related to the solar irradiation, which is a deciding aspect according to the principle of parsimony.

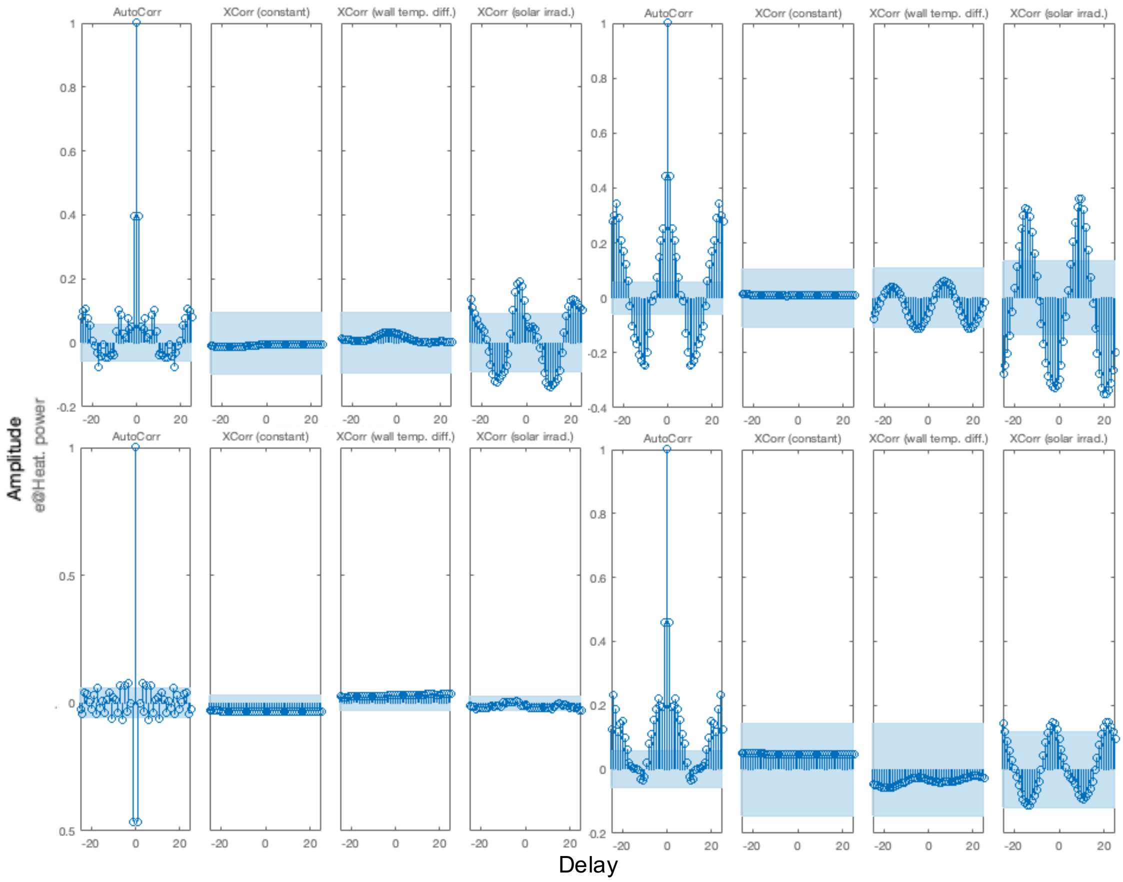

Figure 15. Residual analysis of simulations of all validation folds, building #1, with ARX-model structure $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_{24}q^{-24}$, $B_2(q) = b_{2,0} + b_{2,1}q^{-1} + b_{2,6}q^{-6} + b_{2,20}q^{-20}$ and $B_3(q) = b_{3,0} + b_{3,10}q^{-10} + b_{3,20}q^{-20}$. The structure was, measured by MSE, found to be the best one for fold 4. The blue-shaded area depicts the 95% confidence interval.

5.2.1.2 Residual analysis of ARX-models of building #2

In Figures 16-18, the residual analysis of the model structures with the lowest MSE for building #2 is presented. Overall, the residual analysis of models of building #2 is more pronounced than that corresponding to building #1. The structure with the lowest MSE for fold 1 (Figure 16) shows a notable autocorrelation and cross-correlations with all input signals for all folds.

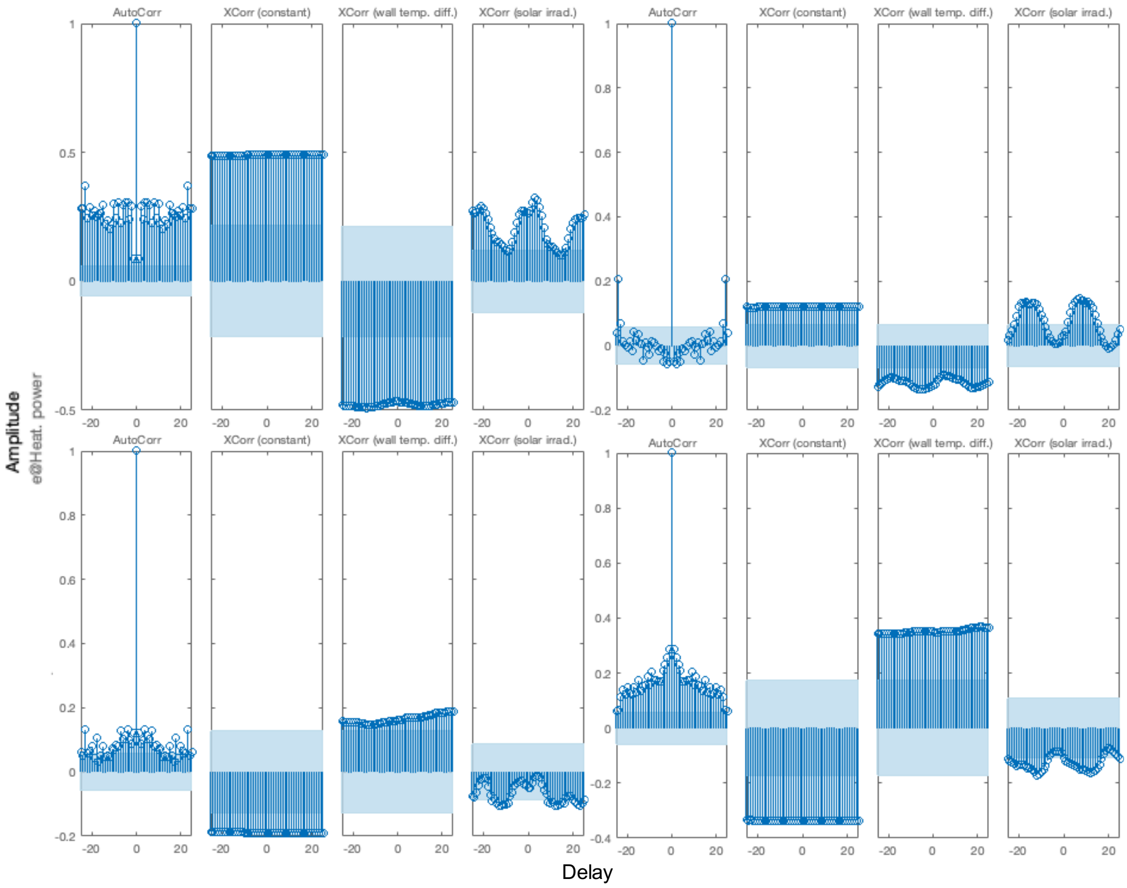

Figure 16. Residual analysis of simulations of all validation folds, building #2, with ARX-model structure $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_3q^{-3} + a_{24}q^{-24}$, $B_2(q) = b_{2,0} + b_{2,1}q^{-1}$ and $B_3(q) = b_{3,0} + b_{3,1}q^{-1} + b_{3,10}q^{-12}$. The structure was, measured by MSE, found to be the best one for fold 1. The blue-shaded area depicts the 95% confidence interval.

The model structure that fit the best to both fold 2 and 3 (Figure 17) show little autocorrelation and cross-correlations for those two folds. There is some correlation of the residuals but including the respective time-delay does not improve model performance and thus, the principle of parsimony is relevant. The residual analysis of folds 1 and 4 is similar to the other structures presented.

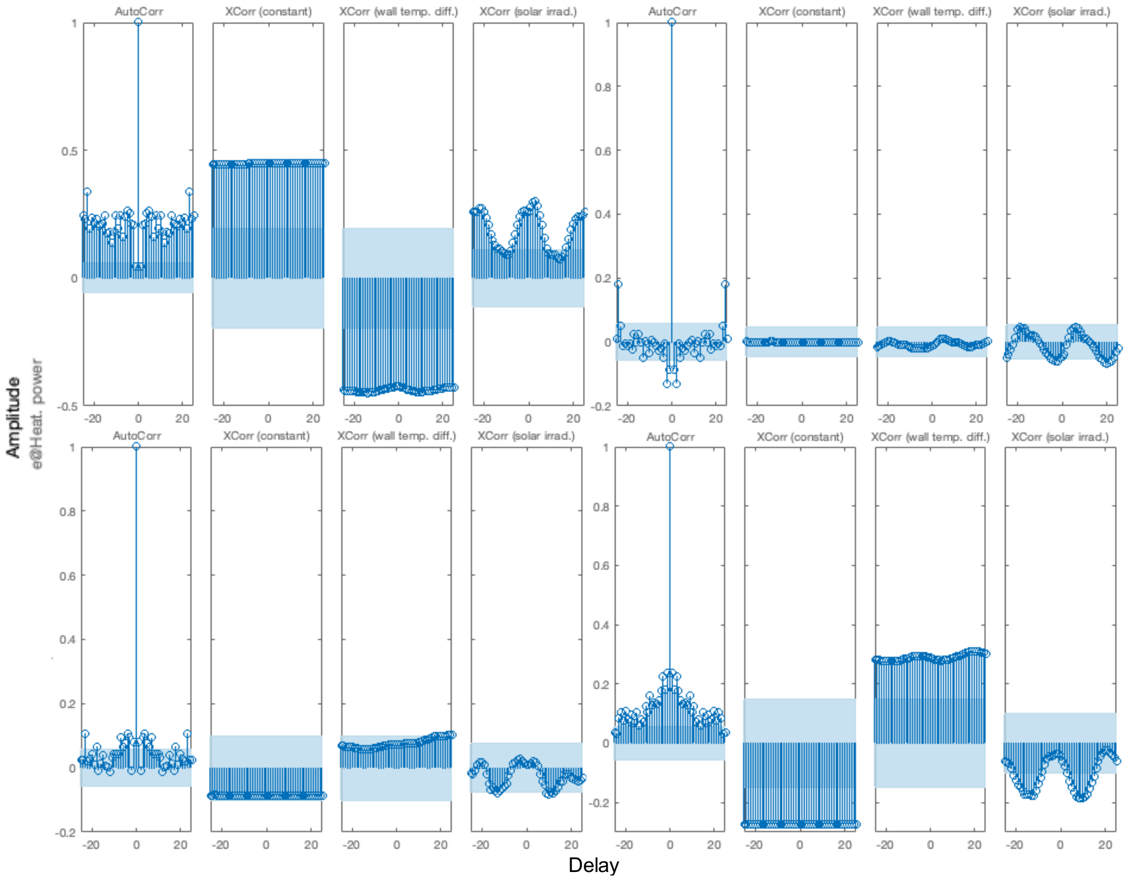

Figure 17. Residual analysis of simulations of all validation folds, building #2, with ARX-model structure $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_{24}q^{-24}$, $B_2(q) = b_{2,1}q^{-1} + b_{2,2}q^{-2}$ and $B_3(q) = b_{3,0}$. The structure was, measured by MSE, found to be the best one for both fold 2 and 3. The blue-shaded area depicts the 95% confidence interval.

The structure best fitting to fold 4 (Figure 18) is notable as very simple. The structure also has considerable cross-correlations with all inputs for all folds but fold 2.

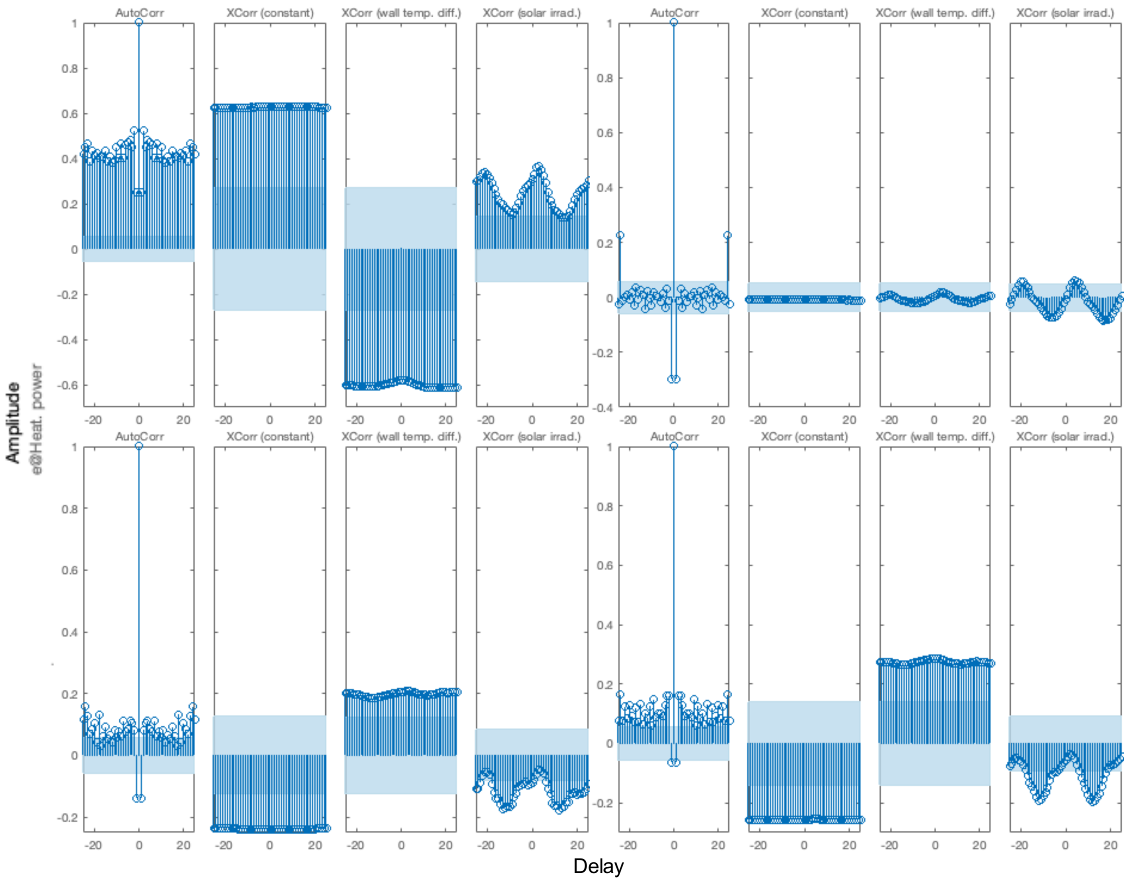

Figure 18. Residual analysis of simulations of all validation folds, building #2, with ARX-model structure $A(q) = 1 + a_1q^{-1}$, $B_2(q) = b_{2,0}$ and $B_3(q) = b_{3,0}$. The structure was, measured by MSE, found to be the best one for fold 4. The blue-shaded area depicts the 95% confidence interval.

5.3 OE-model development

The OE-models were developed for buildings #1 and #2 only. They included all time-delays up to the longest one, unlike the ARX-models that could include specific time-delays. This was due to the way software was designed, which decreased the number of possible combinations and simplified model development. An exhaustive development was not possible due to time constraints and becomes impractical for the purpose of this thesis. It was done for time delays up to $n_b = [-7\ 9]$, $n_f = [-9\ 9]$ for building #1 and took around two days of execution time. The OE-models developed consider constant indoor temperature.

In Table 6 the developed OE-model structures with the lowest MSE are presented. The indoor temperature was considered constant during all OE-model development. For building #1 (developed exhaustively), it can be seen that similar model structures are well performing for a specific fold. Comparing between folds, it appears that different model structures are the better performing for each fold. There is a greater similarity of the structures with regards to the denominator polynomials than the nominators. For building #2 (not developed exhaustively), the better performing models differ for specific folds. Some model structures are well performing for more than a single fold.

Comparing Table 6 to Table 9 (presented in Appendix B) it can be seen that the OE-models perform better than the ARX-models in regards of MSE and CC. However, MAE and MBE measures are similar between ARX- and OE-models. MAE and MBE whey slightly in favour for OE-models overall, but the standard-errors are overlapping. It cannot be said the performance measured in MAE and MBE differ noticeably between ARX- and OE-models.

Table 6. Quantitative measures from cross-validating simulation results of the three best performing OE-model structures for different buildings and folds. Indoor temperature is considered constant. Building #1 was subject to an exhaustive search up to time-delays $n_b = [-\ 7\ 9]$, $n_f = [-\ 9\ 9]$.

Building	Val. fold	n_b	n_f	MAE	(Std.)	MBE	(Std.)	MSE	CC
#1	1	[- 7 6]	[- 7 5]	9.87	8.01	-0.77	12.69	161.62	0.70
		[- 3 9]	[- 7 9]	9.99	8.13	-1.43	12.80	165.76	0.70
		[- 5 9]	[- 8 9]	10.01	8.12	-1.55	12.80	166.10	0.70
	2	[- 2 0]	[- 0 4]	12.12	10.15	-0.83	15.79	249.78	0.92
		[- 2 0]	[- 0 2]	12.19	10.23	-1.84	15.81	253.11	0.92
		[- 4 0]	[- 4 2]	12.25	10.16	-0.64	15.91	253.28	0.92
	3	[- 6 5]	[- 8 8]	16.84	16.13	2.12	23.22	543.49	0.80
		[- 5 7]	[- 6 8]	16.80	16.25	1.98	23.30	546.33	0.80
		[- 2 6]	[- 8 9]	17.09	16.18	2.90	23.35	553.56	0.80
	4	[- 4 9]	[- 5 9]	16.93	13.13	1.71	21.36	458.88	0.86
		[- 4 9]	[- 3 9]	17.17	13.40	-1.43	21.74	474.43	0.85
		[- 3 8]	[- 6 8]	17.62	13.09	1.63	21.90	481.90	0.85
#2	1	[- 24 24]	[- 24 24]	1.95	1.47	-0.83	2.29	5.96	0.71
		[- 3 5]	[- 7 14]	2.35	1.70	-1.21	2.63	8.39	0.65
		[- 0 11]	[- 0 11]	2.76	1.66	-1.84	2.64	10.36	0.64
	2	[- 0 11]	[- 0 11]	3.17	2.48	-0.21	4.02	16.19	0.89
		[- 4 4]	[- 4 4]	3.18	2.55	-0.25	4.07	16.59	0.89
		[- 3 3]	[- 7 7]	3.21	2.54	0.26	4.09	16.80	0.89
	3	[- 3 5]	[- 7 14]	2.92	1.99	0.03	3.54	12.49	0.89
		[- 4 4]	[- 7 7]	2.95	1.98	0.01	3.55	12.62	0.89
		[- 11 11]	[- 23 23]	2.97	1.96	-0.03	3.56	12.66	0.89
	4	[- 24 24]	[- 24 24]	2.54	1.62	0.00	3.02	9.10	0.91
		[- 3 3]	[- 7 3]	2.70	1.66	0.05	3.17	10.03	0.90
		[- 4 4]	[- 11 11]	2.75	1.73	0.29	3.23	10.54	0.90

5.3.1 Residual analysis of OE-models

In Figures 19 and 20, the residual analysis of OE-models with orders $n_b = [-\ 2\ 0]$, $n_f = [-\ 0\ 4]$ and $n_b = [-\ 24\ 24]$, $n_f = [-\ 24\ 24]$ are highlighted. The models highlighted were deemed to be good models of the real system given the quantitative measures.

The residual analysis show that all simulations have residual autocorrelation for all folds. The autocorrelation is larger than for the ARX-models developed. With regards to the constant term φ_0, the residuals are not correlated, with the exception of fold 3, building #1 and fold 1, building #2. These folds also show cross-correlation with the temperature difference φ_2. The simulation of building #1, fold 3, has an MSE of 669.24, MAE of 20.05 and MBE of 4.54, which is higher than for the other structures simulated. However, the simulation of building #2, fold 1, has an MSE of 5.96, MAE of 1.95 and MBE of -0.83, the lowest of any structure developed. The cross-correlation between the residuals and solar irradiation is notable for most folds but varies.

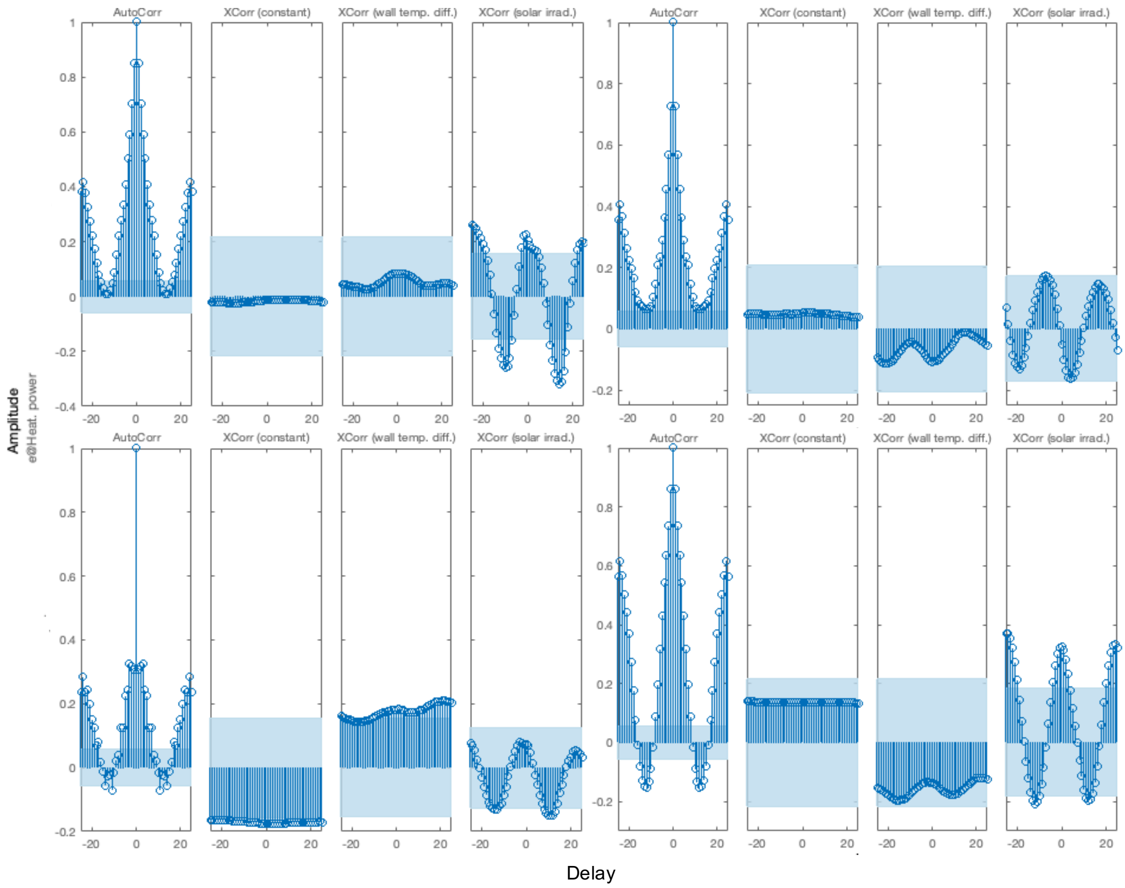

Figure 19. Residual analysis of simulations of all validation folds, building #1, with OE-model structure $n_b = [-2\ 0]$, $n_f = [-0\ 4]$. *The structure was found to be the best for fold #2, by all quantitative measures. The blue-shaded area depicts the 95% confidence interval.*

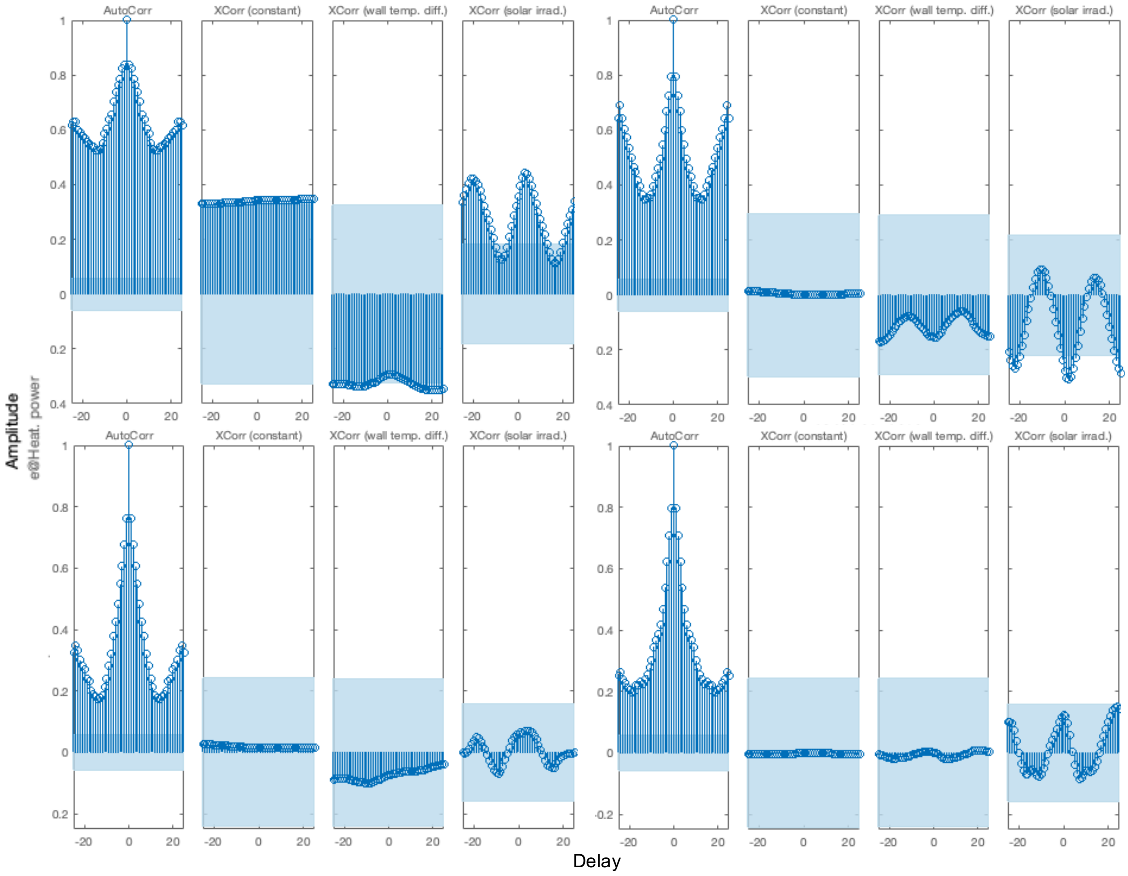

Figure 20. Residual analysis of simulations of all validation folds, building #2, with OE-model structure $n_b = [-\ 24\ 24]$, $n_f = [-\ 24\ 24]$. The structure was found to be the best model structure for fold 1 and 4, by measuring MSE. The blue-shaded area depicts the 95% confidence interval.

5.4 Visual inspection of model output

Following the results of the model development, a visual inspection of the developed models was conducted. All models inspected make use of constant indoor temperature.

5.4.1 Visual inspection of ARX-models

The ARX-model structure that best fit to building #1, fold 2, was

$$\begin{aligned}\hat{y}(t)(1 + a_1q^{-1} + a_2q^{-2} + a_{24}q^{-24}) \\ = b_0 \\ + \theta_2(t)\left(\mathrm{b}_{2,0} + \mathrm{b}_{2,1}q^{-1} + \mathrm{b}_{2,6}q^{-6} + \mathrm{b}_{2,7}q^{-7} + \mathrm{b}_{2,8}q^{-8}\right. \\ \left. + \mathrm{b}_{2,20}q^{-20}\right) + \theta_3(t)\left(b_{3,0}\right)\end{aligned} \tag{52}$$

and simulation of it is presented in Figure 21. The structure showed low correlations between residuals and signals for all folds according to the residual analysis conducted in section 5.2.

A close look at Figure 21 reveals that simulation of model structure captures the overall trend of all folds, with slight exceptions for fold 2. This is notable, as the structure was the best performing one with regards to that specific fold. It can be seen that there is a low-frequency movement in the measured data which is not recreated by the simulation. Notably, in the first 400 sample hours the measured data is slightly higher, whilst in the last 700 it is slightly lower. Overall, the simulations appear representative of the average movements. The structure is not able to capture the variance in the data, with some exceptions. The simulation captures the variance in fold 2 and beginning of fold 3 to an extent. Taking all development and evaluation into consideration, the ARX-model structure (52) can be considered to be a good representation of the system modelled in building #1.

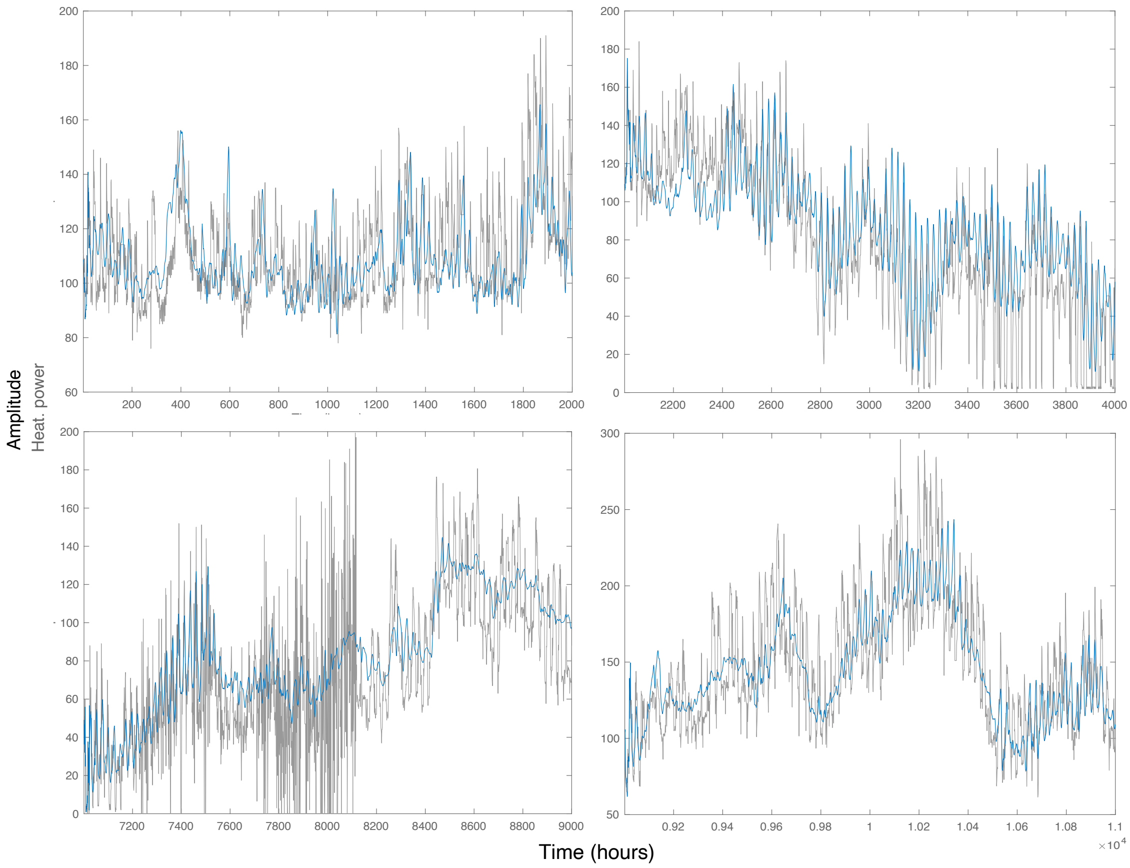

Figure 21. Simulation results (blue) from ARX-model $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_{24}q^{-24}$, $B_2(q) = b_{2,0} + b_{2,1}q^{-1} + b_{2,6}q^{-6} + b_{2,7}q^{-7} + b_{2,8}q^{-8} + b_{2,20}q^{-20}$ *and* $B_3(q) = b_{3,0}$, *for all validation folds (grey) from building #1. The structure was through ARX-model development found to be a good representation of the system modelled in building #1.*

The ARX-model structure simulation showed in Figure 22 was derived through the same method, but for building #2. The model structure is

$$\begin{aligned}\hat{y}(t)(1 + a_1q^{-1} + a_2q^{-2} + a_{12}q^{-12} + a_{24}q^{-24})\\ = b_0 + \theta_2(t)\big(b_{2,1}q^{-1} + b_{2,2}q^{-2}\big) + \theta_3(t)\big(b_{3,0}\big).\end{aligned} \tag{53}$$

The structure was the best fitting ARX-structure for building #2, fold 2 and 3, of all model structures developed. The model structure follows the dynamics well for all validation folds, but there is a constant offset when fitted to and simulating folds 1 and 4. Similar to the visual inspection of simulations of building #1, the structure (53) manages to simulate the variance in the data for fold 2 and the beginning of fold 3 well. It also appears to be the case for fold 4, but with an offset. The offsets are notable in the residual analysis as well. In Figure 17 a notable cross-correlation with the constant input could be seen. All aspects considered, the model structure (53) is regarded to be a good representation of the system modelled.

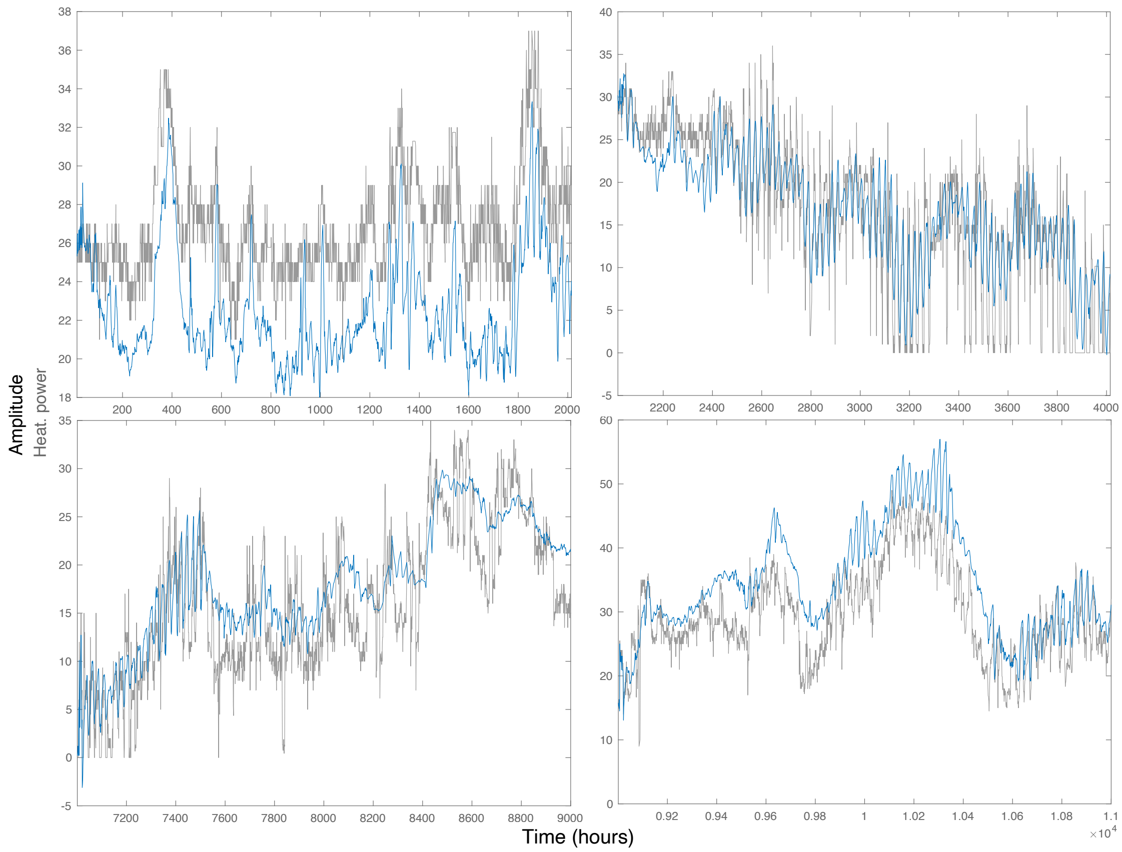

Figure 22. Simulation results (blue) from ARX-model structure $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_{12}q^{-12} + a_{24}q^{-24}$, $B_2(q) = b_{2,1}q^{-1} + b_{2,2}q^{-2}$ *and* $B_3(q) = b_{3,0}$, *for all validation folds (grey) from building #2. The structure was through ARX-model development found to be a good representation of the system modelled in building #2.*

In Figure 23 the same model structure (53) is compared to a very simple model

$$\hat{y}(t) = \theta_2(t)\big(b_{2,0}\big). \tag{54}$$

There are similarities between the simulated model structures, but also a difference in dynamics and mean. The dynamics of the simpler model can be described as more pronounced.

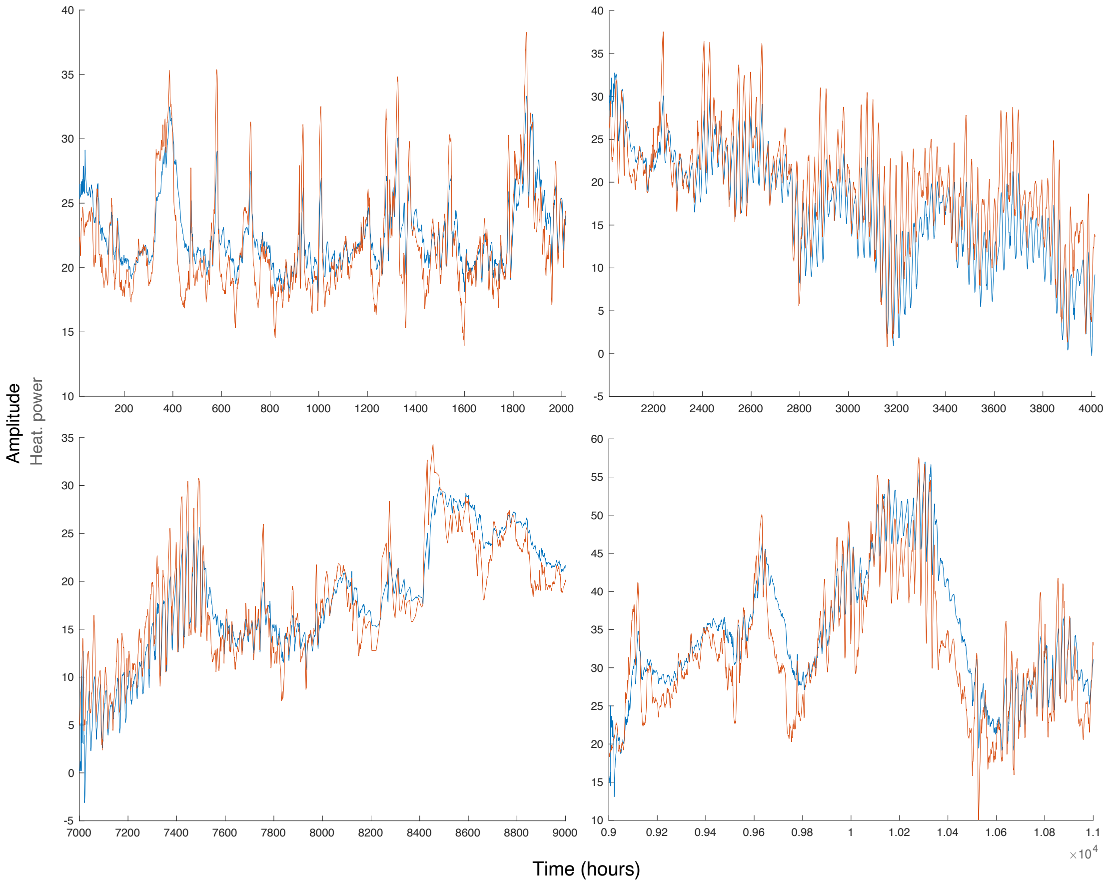

Figure 23. Simulation of building #2, of a developed model (blue): $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_{12}q^{-12} + a_{24}q^{-24}$, $B_2(q) = b_{2,0} + b_{2,1}q^{-1} + b_{2,2}q^{-2}$ *and* $B_3(q) = b_{3,0}$ *and simplest possible model structure (red):* $A(q) = 1,\ B_2(q) = \theta_2(t)\big(b_{2,0}\big)$ *and* $B_3(q) = 0$.

5.4.2 Visual inspection of OE-models

The simulation results from OE-model structures with the lowest MSE for each fold is presented. In Figures 24-27 the simulations for building #1 can be seen and the ones for building #2 in Figures 28-30.

Visually comparing the simulation results of building #1, the model best fitting fold 2 (Figure 25) seems to describe all other models well. Hence, all validation measures combined; the structure $n_b = [-\ 2\ 0]$, $n_f = [-\ 0\ 4]$, can be considered a good representation of the modelled system. The structure best fit to fold 3 (Figure 26) is generally well fitting. The most notable error is in the beginning of fold 2, where the simulation is lower than validation data. The structure best fit to fold 1 (Figure 24) shows a clear overfit to fold 2. The structure best fit to fold 4 (Figure 27) also appear to suffer a slight overfit when simulating fold 2. Notably, the simulation is also quite fluctuating during the first sample hours of fold 4, which it fit best to of all structures.

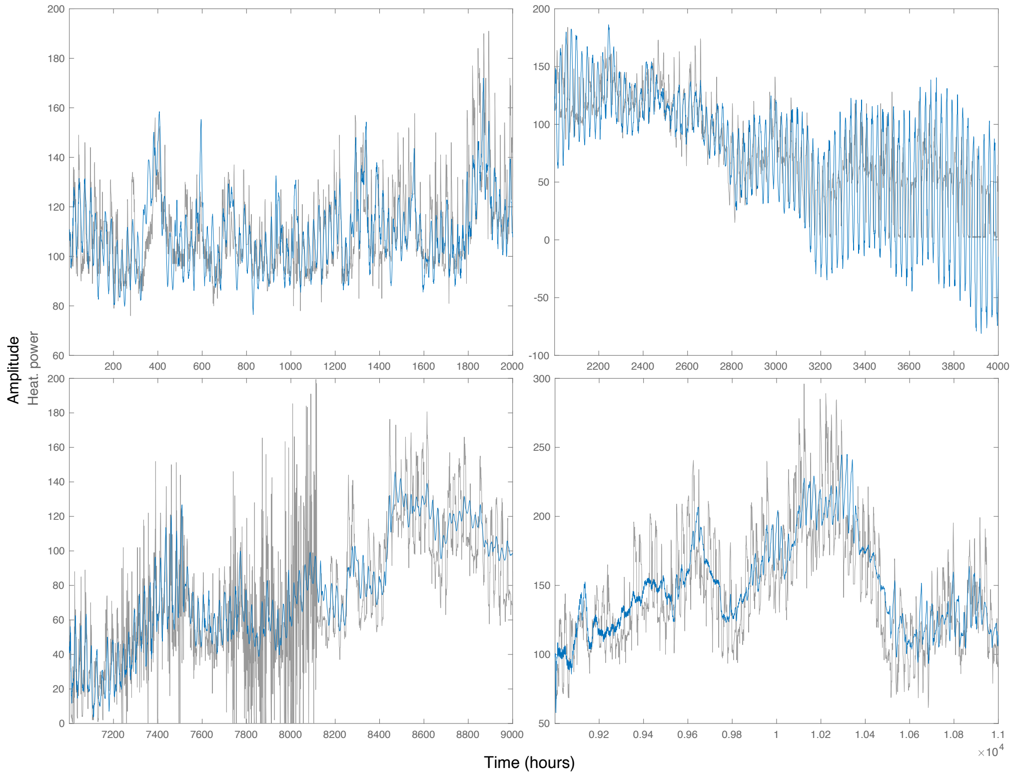

Figure 24. Simulation results (blue) from OE-model structure $n_b = [-7\ 6]$, $n_f = [-7\ 5]$, *for all validation folds (grey) from building #1. The structure was, by measures MAE and MSE, found to be the best model structure for fold 1.*

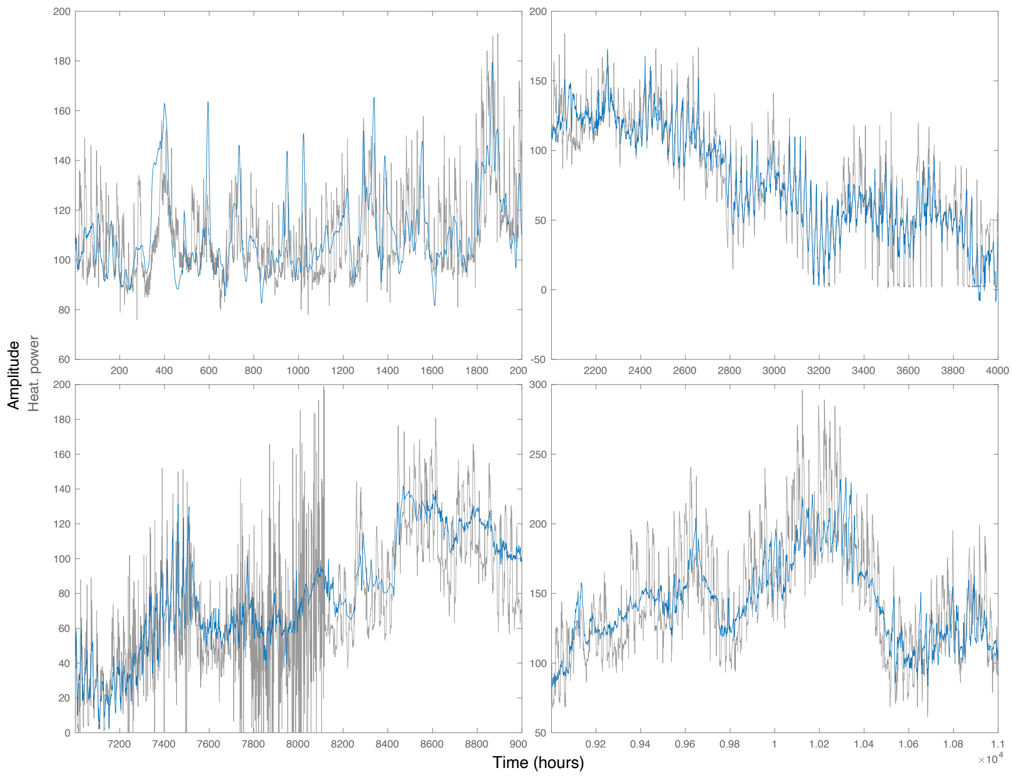

Figure 25. Simulation results (blue) from OE-model structure $n_b = [-2\ 0]$, $n_f = [-0\ 4]$, *for all validation folds (grey) from building #1. The structure was, by measures MAE and MSE, found to be the best model structure for fold 2.*

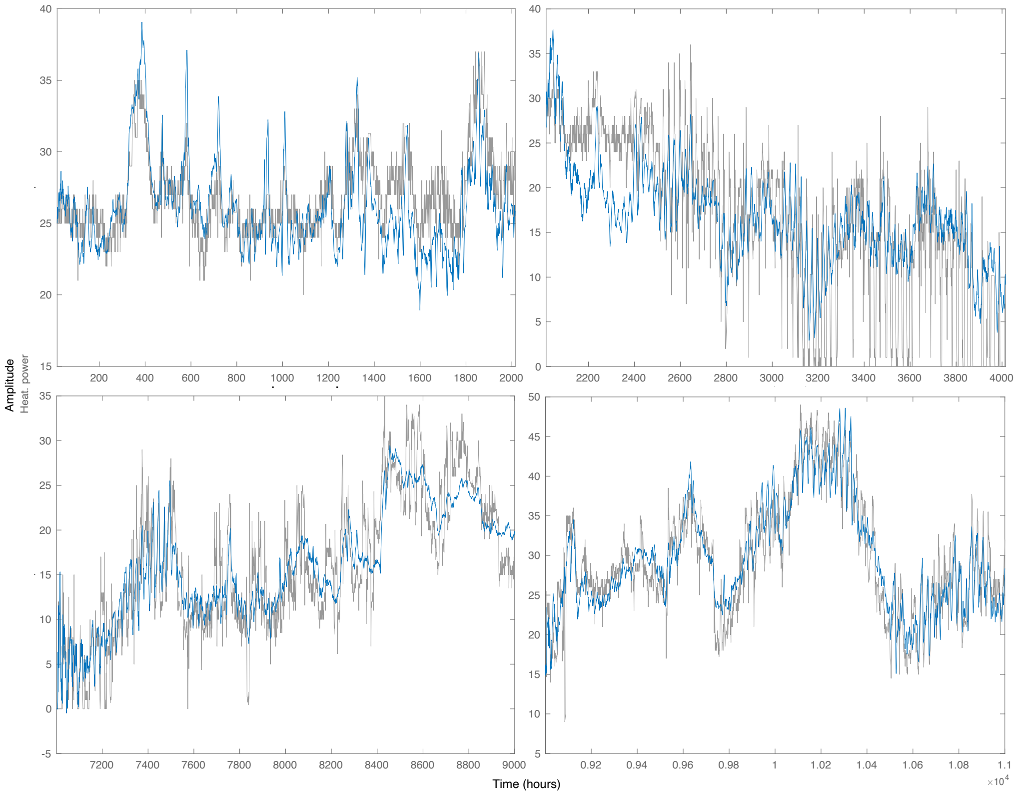

Figure 26. Simulation results (blue) from OE-model structure $n_b = [-6\ 5]$, $n_f = [-8\ 8]$, for all validation folds (grey) from building #1. The structure was, by MSE, found to be the best model structure for fold 3.

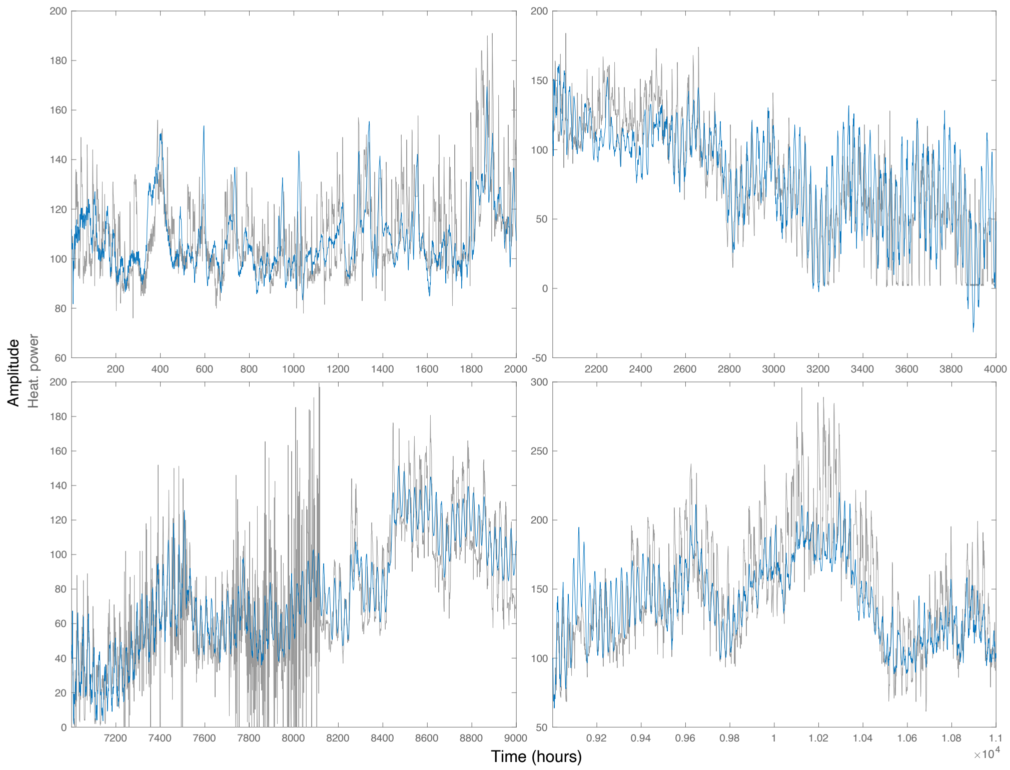

Figure 27. Simulation results (blue) from OE-model structure $n_b = [-4\ 9]$, $n_f = [-5\ 9]$, for all validation folds (grey) from building #1. The structure was, by measures MAE and MSE, found to be the best model structure for fold 4.

Comparing the simulation of model structures for building #2 (Figure 29), no signs of notable overfit can be seen through visual inspection. The simulation results are generally good but differ some for fold 2. Between sample hours 2000-2400, only the model structure that fit particularly well to that specific fold results in a precise simulation. The same model structure appears to be the best performing one overall. All simulations adhere to the measured values in the validation fold and there is no apparent overfitting. All validation measures considered, the structure that best fit to fold 1 and 4, $n_b = [-\ 24\ 24]$, $n_f = [-\ 24\ 24]$, is considered the best representation of the modelled system in building #2.

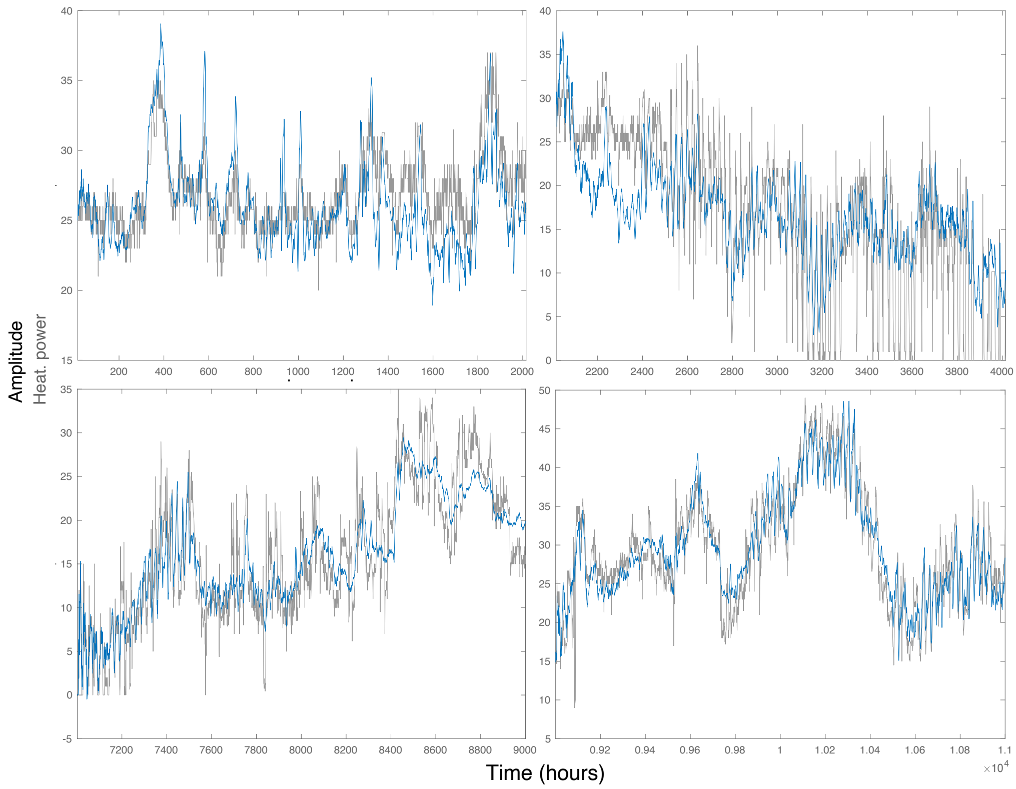

Figure 28. Simulation results (blue) from OE-model structure $n_b = [-\ 24\ 24]$, $n_f = [-\ 24\ 24]$, *for all validation folds (grey) from building #2. The structure was, by all quantitative measures, found to be the best model structure for fold 1 and 4.*

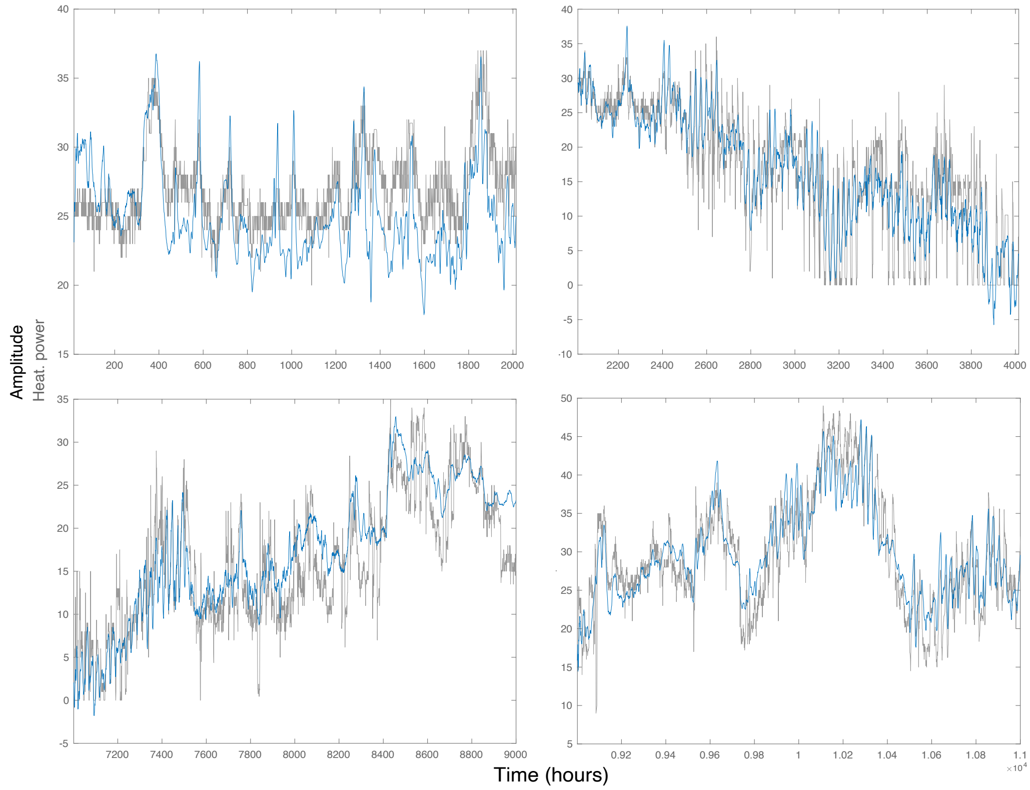

Figure 29. Simulation results (blue) from OE-model structure $n_b = [-\ 0\ 11]$, $n_f = [-\ 0\ 11]$, *for all validation folds (grey) from building #2. The structure was, by all quantitative measures, found to be the best model structure for fold 2.*

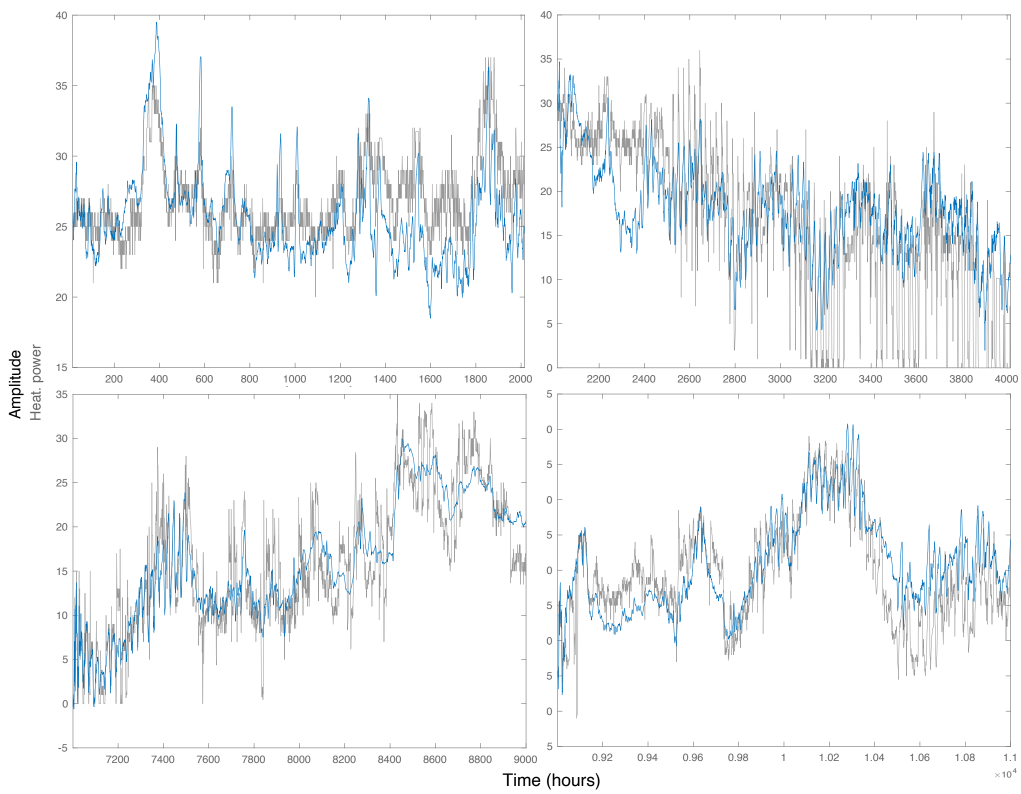

Figure 30. Simulation results (blue) from OE-model structure $n_b = [-\ 3\ 5]$, $n_f = [-\ 7\ 14]$, *for all validation folds (grey) from building #2. The structure was, by measures of MSE and CC, found to be the best model structure for fold 3.*

6. Sensitivity analysis

A sensitivity analysis is conducted with a one-at-a-time (OAT) approach. This allows for the robustness and contribution of different parts to be determined in a linear model (Saltelli et al., 2008). The sensitivity analysis is done by increasing and decreasing the parameters of a part of the model by 25%, whilst keeping the other parameters set to its least squares estimate. For example, if a model is estimated as

$$B(q) = b_0 + b_1 q^{-1} + b_2 q^{-2} + b_{10} q^{-10} \tag{55}$$

the sensitivity analysis of inputs of long delays would look as

$$B_s(q) = b_0 + b_1 q^{-1} + b_2 q^{-2} + (1 \pm 0.25) b_{10} q^{-10}. \tag{56}$$

The sensitivity is then evaluated by the output of the model given that B is exchanged for B_s in the model.

The OAT approach is appreciated amongst the modelling community for its simplicity and features (Saltelli and Annoni, 2010). Three of those features are 1) a baseline is established which all evaluation is referred to, 2) the analysis is isolated to a single factor and 3) the approach only detects a changed output from relevant factors. However, the approach can notably not dismiss irrelevant factors by there not being a change in output. Overall, it is possible to draw conclusive results by applying to OAT approach.

Though the OAT approach is the most commonly used sensitivity analysis by modellers, it is subject of criticism (Saltelli and Annoni, 2010). The main criticism raised by Saltelli and Annoni (2010) is that the approach is inept for conducting a global analysis. It is not possible to conclude overall model sensitivity using the approach. They further argue that an *elementary effects* approach is about as simple but allows for a considerably more thorough analysis. The criticism raised by Saltelli and Annoni (2010) is similar to the emphasis on considering operating conditions by Ljung and Glad (2004). Such consideration has already been made during model development with the data selection process and folding method. It is therefore argued that another ethos of Ljung and Glad (2004) should be considered: to model considering the purpose of the model. The OAT approach allows for the simple integration and clear evaluation of different changes to the model, which is in line with the purpose and goals of this thesis. Therefore, the OAT approach is considered applicable to ARX-models to obtain the results and evaluation that is sought from the sensitivity analysis.

The approach is also applicable to the OE-models. However, as they are more difficult to interpret than the ARX-models, it does not serve the same purpose to group regressors. It is therefore not possible to motivate by the same ethos as applying it to ARX-models. If a similar sensitivity analysis where to be conducted on OE-models, it could perhaps make use of filters to enhance dynamics of certain frequencies. Such a sensitivity analysis would not be dependent on there being groupings of regressors.

6.1 Results of sensitivity analyses

In Figures 31-33, sensitivity analyses of some ARX-model structures are shown. The structures are found to be good representations of the building they were modelled after. Changing the value of the different polynomial parameters changes the simulation output. The variance of the output seems to be maintained for all independent of parameters changed, but the low-frequency output appears slightly changed after decreasing the s.d. AR parameters. Notably, increasing the s.d. AR parameters makes the system unstable and is not included in the visualisation. The most notable changes in model output can be seen in changing the parameters of the AR parameters and s.d. temp. diff.

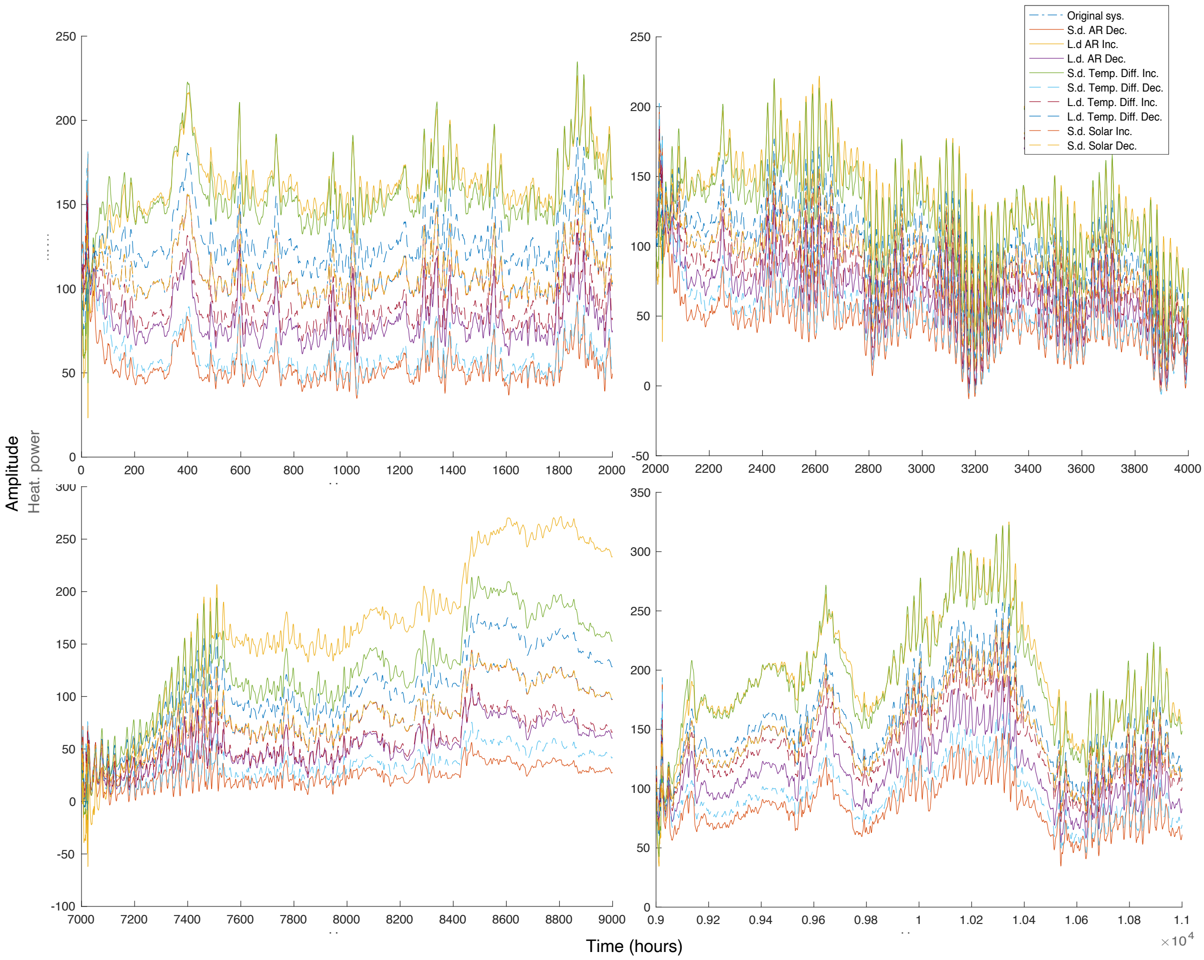

Figure 31. Sensitivity analysis of ARX-model structure: $A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + a_{24} q^{-24}$*,* $B_2(q) = b_{2,0} + b_{2,1} q^{-1} + b_{2,6} q^{-6} + b_{2,7} q^{-7} + b_{2,8} q^{-8} + b_{2,20} q^{-20}$ *and* $B_3(q) = b_{3,0}$*. The analysis is done on data from building #1, which the structure has been found to be a good representation of. Groups of time-delays have been increased or decreased by 25%.*

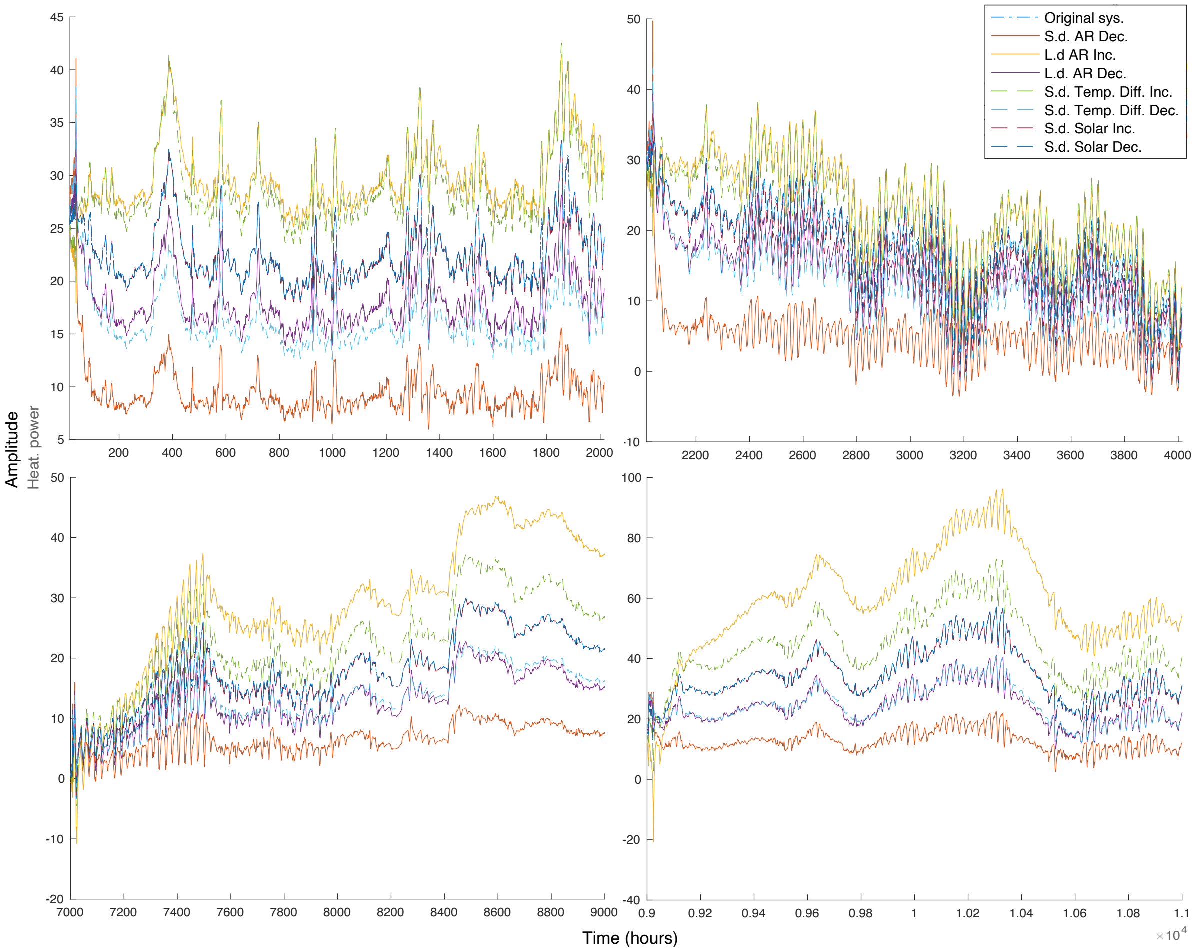

Figure 32. Sensitivity analysis of ARX-model structure: $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_{12}q^{-12} + a_{24}q^{-24}$, $B_2(q) = b_{2,0} + b_{2,1}q^{-1} + b_{2,6}q^{-6} + b_{2,7}q^{-7} + b_{2,8}q^{-8} + b_{2,20}q^{-20}$ *and* $B_3(q) = b_{3,0}$. *The analysis is done on data from building #2, which the structure has been found to be a good representation of. Groups of time-delays have been increased or decreased by 25%.*

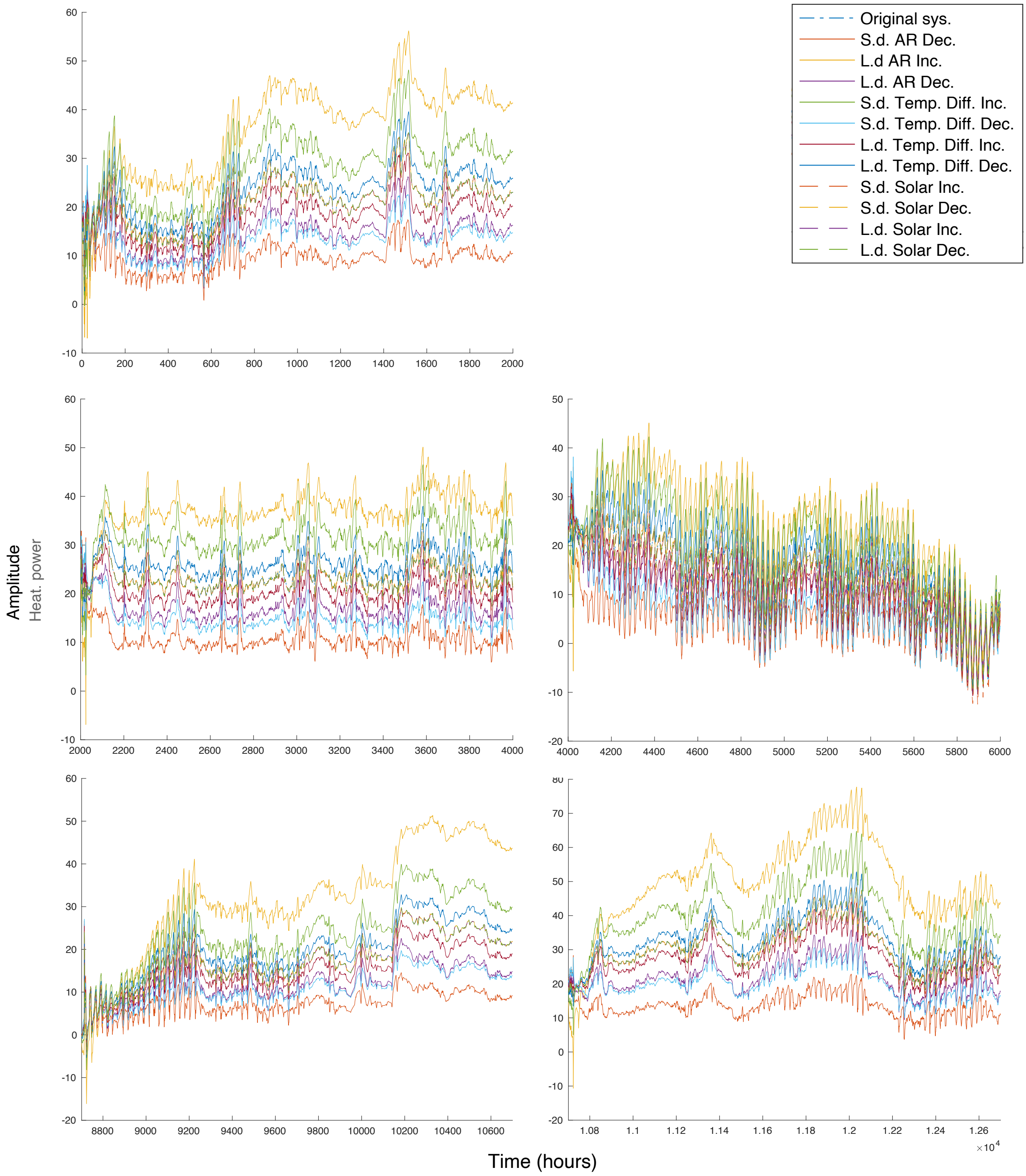

Figure 33. Sensitivity analysis of ARX-model structure: $A(q) = 1 + a_1q^{-1} + a_2q^{-2} + a_3q^{-3} + a_{24}q^{-24}$, $B_2(q) = b_{2,0} + b_{2,1}q^{-1} + b_{2,2}q^{-2} + b_{2,8}q^{-8} + b_{2,24}q^{-24}$ *and* $B_3(q) = b_{3,0} + b_{3,1}q^{-1} + b_{3,2}q^{-2} + b_{3,8}q^{-8}$*. The analysis is done on data from building #3, which the structure has been found to be a good representation of. Groups of time-delays have been increased or decreased by 25%.*

In Table 7, the results of the sensitivity analysis are summarised. It is further notable that the results from the sensitivity analysis are analogous between the models shown.

Table 7. Results of OAT sensitivity analysis. Time-delays 5 hours and shorter are considered short, 6 hours and longer are considered long.

Parameters	Time-delays	Change	Result	Comment
AR ($A(q)$)	Short	Increase	The system becomes unstable.	
		Decrease	Decreases simulation output the most of any change. Low frequency dynamics appear slightly affected.	
	Long	Increase	Increases model output the most of any change.	The increase is similar to s.d. temp. diff. during simulation output with much variance, but not during output with little variance.
		Decrease	Decreases model output the third most of any change.	
Temp. diff. ($B_2(q)$)	Short	Increase	Increases model output the second most of any change.	The increase is similar to l.d. AR during simulation output with much variance, but not during output with little variance.
		Decrease	Decreases model output the second most of any change.	
	Long	Increase	Decrease model output the fourth most of any change.	
		Decrease	Increases model output the third most of any change.	
Solar irradiation ($B_3(q)$)	Short	Increase	Negligable	
		Decrease	Negligable	
	Long	Increase	Negligable	
		Decrease	Negligable	

7. Discussion

This chapter will discuss the resulting model development in relation to the goals and purpose of this thesis. It will begin with a discussion of the functionality of measuring indoor temperature, which was dismissed from model development based on early results. The chapter will continue with a comparison of the developed ARX- and OE-models. It will raise a comparison of quantitative measures, model dynamics and the model development of both structure types. Based on the model development, a general model structure will be discussed. Lastly, the model development will be discussed through a physical interpretation.

7.1 Indoor temperature

According to the energy balance, the indoor temperature plays an important role in explaining the heating demand. A warmer room demands more heat than a colder room and the energy balance explains why this is. Including the indoor temperature should therefore make for a better model which allows for more insight into the dynamics of the modelled system (see chapters 2.9.1, 2.9.4 and 4.1.1). Early model development did however not indicate an improved model by including the indoor temperature. Instead, the model development of this thesis shows that considering the indoor temperature to be constant results in better models.

The lesser performance obtained by using measured indoor temperature can perhaps be attributable to difficulties in aggregating the data series, which were collected in an ad-hoc manner. Previous studies have successfully implemented the indoor-temperature in the respective models, but they did not have the same method for collecting and pre-processing data as this thesis. Wu and Sun (2012) studied smaller building sections and did not aggregate any measurements. Hietaharju et al. (2018) aggregated at most two series.

7.2 Comparing ARX- to OE-models

7.2.1 Quantitative measures

As is seen by comparing Table 6 and Table 9 (presented in Appendix B), OE-models are better performing by quantitative measures. However, both types of models are indistinguishable with regards to MAE and MBE. This means that the OE-models capture more of the variance in the simulation (see the relationship between bias, variance and MSE in (24)). This finding was also found by visually comparing model outputs. The results indicate that OE-models are better than ARX-models for simulating a more detailed heating demand.

7.2.2 Model dynamics

Comparing the results found in chapters 5.2.1, 5.3.1 and 5.4, the different dynamics encapsulated by ARX- and OE-models can be seen. Visual inspection and quantitative measurements further indicate that OE-models are better than ARX-models at simulating the fast-changing heating demand intra-day. Both models can be considered to capture the overall dynamics well given that there is little MAE, MBE and a high CC in general.

The residual analysis of ARX- and OE-models differs some. Both types of models generally have residuals uncorrelated to the difference in temperature (φ_2). This indicates that the heating demand created by the difference in temperature indoors and outdoors can be described by the linear ARX- and OE-model structures.

No model is without cross-correlation between the residuals and solar irradiation (φ_3) for all folds, but the OE-models show less cross-correlation. This means that there are dynamics in the solar irradiation which are difficult to encapsulate with the linear ARX- and OE-models. It could be the case that the relationship is nonlinear or time-variant. However, the sensitivity analysis presented in chapter 6.1 shows that the solar irradiation has little impact on the simulated heating demand. It is an expected result given that there is little sunlight in Sweden during the heating season. On the other hand, including the solar irradiation improved the simulations, so it cannot be dismissed from the model structures. The dynamics related to the solar irradiation cannot be considered fully modelled.

Comparing the autocorrelation of the residuals, it can be seen that the inclusion of AR dynamics does remove some residual dependence previous outputs. This indicates that there is an AR behaviour inherent to the system, which by definition cannot be incorporated into the OE-model.

7.2.3 Comparing model development

The process of developing ARX- and OE-models is different. As ARX-model parameters are obtained through LS-estimates using the normal equation (see (12)), the model structure is easily altered and can incorporate specific time-delays. OE-model parameters are however obtained through gradient-based algorithms, which required more complex methods compared to an LS-estimate. As a result, this thesis has used pre-made software for developing OE-models. The software used includes all time-delays up to the longest one. However, this is in theory not a necessity.

Encapsulating important time-delays in an OE-model is therefore more difficult and important time-delays risk being hidden amongst unnecessary ones that are included. There is also a risk of overfitting the model, which was shown for OE-models. During ARX-model development, no overfit models were encountered. This could be due to the guided approach applied to ARX-model development. Applying the same residual analysis-guided approach to OE-models is more difficult because the structure includes both nominator and denominator polynomials. Doing an exhaustive search is in theory simple. But in practice, it takes days of computational time, far beyond what can be considered practical.

7.3 Finding a general model structure

Finding a general model structure is an endeavour that must grapple with multiple models describe a system well, as noted by Saltelli et al. (2008). Since all the models developed essentially can be seen as a linearisation of a complex system, no model represents the exact system. Instead, a model structure which works well for multiple folds indicates that the building's heating demand is the product of coherent dynamics encapsulated by the model structure. If multiple buildings are described by similar model structures, it means that the dynamics are reoccurring between buildings. This is to be expected from a system driven by physics.

The evaluation of model structures for specific folds has made much use of MSE. However, the measure varies between buildings and folds. To some extent, this is because models describe the system differently well. But to a more significant extent, this is because the magnitudes of the signals are different. Absolute MSE is therefore not a useful measure to compare model structures on different buildings or folds. It will only indicate how well a model fits to the specific data. The basis of a general model structure is better evaluated by the dynamics simulated by the models, making use of residual analysis, visual inspection and measures such as MAE, MBE and CC.

7.3.1 Comparing different buildings

Overall, the developed model structures (ARX and OE) have close to zero MAE and MBE values for all folds. There is uncertainty to the measures, indicated by the standard deviation. Since the low MAE and MBE are reoccurring between folds and buildings, it indicates that the true values are indeed low. Further, the low MAE and MBE indicate that the model structures manage to simulate the low-frequency dynamics and long-term behaviour of the system.

The CC:s are high for most well performing structures, meaning that the model output is highly correlated to the measured outputs. This result is expected from a model that well encapsulates the dynamics the real system. There is only a single fold for which no model simulation had a CC below 0.6. No other quantitative measures are outliers for this fold. The majority of CC:s are higher than 0.8.

Finally, illustrated by Table 5, groupings of time-delayed regressors can be found reoccurring between folds and buildings. This is taken as indication that ARX-model structures can be generalised between buildings. These results could not be found for OE-model structures. It should be remembered that OE-model development was conducted different to ARX-models and could not include specific time-delays. Only buildings #1 and #2 were modelled with OE-models.

7.4 General model structure and comparison to physical relationships

ARX-models can be interpreted in the time-domain and are therefore rather simple to understand the dynamics of. This also makes the model structure fit the purpose of this thesis. As they are considered to model the system well, the principle of parsimony further motivates the use of them. An aggregated general structure for ARX-models can be found through model development. All buildings benefit from including regressors with short time-delays, and most also from including regressors with long time-delays. The general ARX-model structure found, by quantitative measures, was

$$heating\ demand = s.d\ AR + l.d.AR + s.d.temp.difference + l.d.temp.difference + s.d.solar\ irradiation + l.d.solar\ irradiation. \tag{57}$$

However, l.d. solar irradiation was the least reoccurring grouping and residual analysis signed that the grouping was not of much importance.

Interpreting the reoccurring pattern found amongst ARX-models through the lens of physics, the model describes the following thermal dynamics in a heating system

$$heating\ demand = heating\ system\ inertia + inter\text{-}day\ trends + quick\ heat\ losses + heat\ losses\ through\ material\ with\ a\ large\ thermal\ mass + direct\ insolation + heating\ of\ material\ with\ a\ large\ thermal\ mass. \tag{58}$$

Interpreting the ARX-model structure as such allows for energy flows to be identified. It is beyond the scope of this thesis to determine energy flows with certainty, but through the results, coupled with previous literature and documentation, a probable description of energy flows can be portrayed. In Table 8, the different parts of the model structure are coupled to probable energy flows and motivation is given. Note that the coupling is not exhaustive and there may be specific phenomena coupled to a specific energy flow. The purpose of Table 8 is to discuss how energy flow relates to the identified models. The results from the sensitivity analysis are also discussed in the light of physical interpretation.

Coupling energy flows to the developed OE-models is possible, but also more difficult since the denominator polynomials requires that the model is understood through the frequency-domain. Since the models do not highlight specific time-delays, the specific energy flows need to be analysed through frequency analysis and is left out of this thesis.

Table 8. Possible interpretations of generalised model structures.

Group of time-delays	Interpretation	Motivation	Sources
S.d. AR	Inertia in the heating system.	Inertia would cause the heating system to act slower and less precise, leading to heat waste. The inertia could stem from factors such as slow reaction to demand, addressable through model predictive control, or the heating system being under-dimensioned. Based on the sensitivity analysis, the inertia of the heating system plays a large	(Abel and Elmroth, 2016; Drgoňa et al., 2020; Hietaharju et al., 2018; Holm, 2015; Wu and Sun, 2012)

		part in how much energy is needed to be consumed.	
L.d. AR	Inter-day trends	The current heating demand is dependent on the energy consumed the previous day. This could be related to the thermal inertia of the building interacting with seasonal temperature change.	(Verbeke and Audenaert, 2018)
S.d. temperature difference	Heat losses through windows and thermal bridges	Windows and thermal bridges constitute the main heat loss from a warm indoor climate to a cold outside climate. The thermal inertia of these objects is small given that windows generally have low mass, and the thermal bridges tend to have high thermal diffusivity. This gives rise to quick changes in heating demand. The sensitivity analysis shows that the s.d. temperature difference is a notable factor for the energy consumed. Energy mappings have shown the same for heat losses through thermal bridges and windows. The main tool for decreasing these energy losses is targeted insulation.	(Abel and Elmroth, 2016; Holm, 2018; Kurkinen, n.d.; Tsilingiris, 2004)
L.d. temperature difference	Heat loss through walls and thermal inertia	Walls, in particular in stone and brick buildings, have a lot of thermal inertia. The thermal inertia gives rise to time-constants, which in turn is dependent on how the insulation is designed. High thermal inertia is thought to have a smoothing effect on indoor temperature and provide thermal comfort. However, the potential to decrease energy consumption	(Holm, 2018; Kurkinen, n.d.; Tsilingiris, 2004; Verbeke and Audenaert, 2018)

		through increased thermal inertia has proven to be limited. Energy mappings have shown that conductive heat loss through walls and roof is small but can be decreased. The sensitivity analysis states that the l.d. temperature difference has an effect on the overall energy consumption.	
S.d. solar irradiation	Solar insolation	The effect of solar insolation is dependent on how much sunlight shines through the windows. Solar insolation helps to heat up the indoor climate. However, it is very limited during the heating season in Sweden. The insolation is also dependent on the windows' and building's overall position. If it is shaded, the sunlight will not reach. This would also give rise to time-variance, which the residual analysis gives hints of. The sensitivity analysis concludes that the contribution of solar insolation is minor.	(Abel and Elmroth, 2016; Hietaharju et al., 2018; Swedish Meteorological and Hydrological Institute, 2021; Wu and Sun, 2012)
L.d. solar irradiation	Solar heating of walls and opaque outside material	Unlike sunlight that shines through a clear material, sunlight heats up opaque material which in turn gives off heat to its surroundings, decreasing heating demand. The heating of the material itself gives rise to a time-delay, especially large if the material has large thermal inertia. However, the solar irradiation is very limited	(Jiménez et al., 2008; Swedish Meteorological and Hydrological Institute, 2021; Tsilingiris, 2004, 2002)

during the heating season in Sweden.

The heat absorption from solar irradiation is dependent on the colour and texture of the outer material.

The sensitivity analysis shows that the contribution to the energy consumption from the l.d. solar irradiation is minor.

8. Conclusions and future studies

Based on a building's thermal energy balance, ARX- and OE-models were developed for simulation purposes. The models made use of time-series of heating demand, indoor temperature, outdoor temperature and solar irradiation. Early model development showed that including data series of indoor temperature did not to improve the model, likely due to the aggregation of measurements that was made. As the indoor temperature is seldom recorded in most buildings, this result is not considered impractical. The main loss is that the developed models become less insightful than it could have been otherwise. Better models were found by considering the indoor temperature to be constant.

Both ARX- and OE-models can be considered to encapsulate the dynamics of the buildings' heating demand, indicated by visual inspection, residual analysis and measures such as MAE, MBE and CC. OE-models perform better for simulation purposes given most measures, but are more difficult to interpret than ARX-models and more difficult to develop guided by residual analysis. A particular difference in simulation performance is that OE-models are better at simulating intra-day dynamics.

A general structure of ARX-models could be found based on groupings of time delays. The general model structure found was discussed in relation to physics related to heating demand. Based on the OAT sensitivity analysis conducted, the interpreted physics are in line with documented energy flows in buildings. This indicates that information about physics and energy flows can be extrapolated from the ARX-models. The sensitivity analysis state that the two main factors driving energy consumption for heating is the design of the heating system and heat losses through windows and thermal bridges.

The concluding statement on the purpose and goals of this thesis is that it is possible to:

- through physics develop a model framework suitable for system identification,
- extrapolate information about physical properties through analysis of the derived model, and
- identify the main contributions to heating demand through model simulation.

As the goals are considered to be met, the method applied in this thesis can may be proposed to for use in analysing retrofitting potential and analysis of energy flows.

8.1 Future studies

The thesis showed that OE-models are better at simulating energy demand. Therefore, future studies should attempt to discuss the physics in relation to developed OE-models. As it appears AR dynamics are relevant to the model, exploring full Box-Jenkins models can be of interest as well.

For the purpose of developing upon the goals set by this thesis, it is recommended for future studies to elaborate on the physical interpretation of the heating demand. An example would be to include more meteorological inputs, such as wind speeds. It could also be interesting to implement coincidence factors (see (Yarbrough et al., 2015)) or proxy series (like electricity usage, as noted by (Lowry and Lee, 2004)) to include heating demand related to human activities.

Lastly, it is recommended to build upon this study and cement its findings. It is proposed that future studies take a qualitative or comparative approach for this purpose. Generalisable system knowledge should be sought and future models should strive towards becoming more open and “white”.

References

Abel, E., Elmroth, A., 2016. Byggnaden som system. Studentlitteratur, Lund.

Afram, A., Janabi-Sharifi, F., 2014. Review of modeling methods for HVAC systems. Applied thermal engineering 67, 507–519.

Boverket, 2021. Bygg- och fastighetssektorns energianvändning uppdelat på förnybar energi, fossil energi och kärnkraft [WWW Document]. URL https://www.boverket.se/sv/byggande/hallbart-byggande-och-forvaltning/miljoindikatorer---aktuell-status/energianvandning/ (accessed 5.24.21).

Coakley, D., Raftery, P., Keane, M., 2014. A review of methods to match building energy simulation models to measured data. Renewable and Sustainable Energy Reviews 37, 123–141. https://doi.org/10.1016/j.rser.2014.05.007

Ding, F., 2014. Combined state and least squares parameter estimation algorithms for dynamic systems. Applied Mathematical Modelling 38, 403–412. https://doi.org/10.1016/j.apm.2013.06.007

Drgoňa, J., Arroyo, J., Cupeiro Figueroa, I., Blum, D., Arendt, K., Kim, D., Ollé, E.P., Oravec, J., Wetter, M., Vrabie, D.L., Helsen, L., 2020. All you need to know about model predictive control for buildings. Annual Reviews in Control 50, 190–232. https://doi.org/10.1016/j.arcontrol.2020.09.001

Glad, T., Ljung, L., 2003. Reglerteori: flervariabla och olinjära metoder. Studentlitteratur, Lund.

Hietaharju, P., Ruusunen, M., Leiviskä, K., 2018. A Dynamic Model for Indoor Temperature Prediction in Buildings. Energies 11, 1477. https://doi.org/10.3390/en11061477

Holm, H., 2018. Uppvärmning [WWW Document]. Uppvärmning. URL http://www.energimyndigheten.se/energieffektivisering/jag-vill-energieffektivisera-hemma/inkop-av-produkter/uppvarmning/

Holm, H., 2015. Dimensionera ditt värmesystem rätt [WWW Document]. Energimyndigheten. URL (accessed 5.10.21).

IEA, 2020a. Heating [WWW Document]. Heating. URL https://www.iea.org/fuels-and-technologies/heating (accessed 5.24.21).

IEA, 2020b. Heating. IEA, Paris.

Jiménez, M.J., Madsen, H., Andersen, K.K., 2008. Identification of the main thermal characteristics of building components using MATLAB. Building and Environment 43, 170–180. https://doi.org/10.1016/j.buildenv.2006.10.030

Kalman, R.E., 1960. A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering 82, 35–45. https://doi.org/10.1115/1.3662552

Kuhn, M., Johnson, K., 2013. Measuring Performance in Regression Models, in: Kuhn, M., Johnson, K. (Eds.), Applied Predictive Modeling. Springer New York, New York, NY, pp. 95–100. https://doi.org/10.1007/978-1-4614-6849-3_5

Kurkinen, E.-L., n.d. Guide Tilläggsisolering [WWW Document]. Informationscentrum för hållbart byggande. URL https://ichb.se/innehall/guider/guide-tillaggsisolering/ (accessed 5.10.21).

Ljung, L., 2001. Estimating Linear Time-invariant Models of Nonlinear Time-varying Systems. European Journal of Control 7, 203–219. https://doi.org/10.3166/ejc.7.203-219

Ljung, L., Glad, T., 2004. Modellbygge och simulering. Studentlitteratur, Lund.

Lowry, G., Lee, M.-W., 2004. Modelling the passive thermal response of a building using sparse BMS data. Applied Energy 78, 53–62. https://doi.org/10.1016/S0306-2619(02)00164-2

Nordling, C., Österman, J., 2013. Physics handbook for science and engineering, 8. ed. ed. Studentlitteratur, Lund.

Oppenheim, A.V., Willsky, A.S., Hamid Nawab, S., 2013. Signals and systems, 2. ed., Pearson new international Edition. ed, Always learning. Pearson Education, Harlow, Essex.

Pavelková, L., Kárný, M., 2014. State and parameter estimation of state-space model with entry-wise correlated uniform noise: ESTIMATION OF UNIFORM MODEL. Int. J. Adapt. Control Signal Process. 28, 1189–1205. https://doi.org/10.1002/acs.2438

Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. Environmental Modelling & Software 25, 1508–1517. https://doi.org/10.1016/j.envsoft.2010.04.012

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global sensitivity analysis: the primer. John Wiley & Sons.

Swedish Meteorological and Hydrological Institute, 2021. SMHI Öppna data [WWW Document]. Öppna data. URL https://www.smhi.se/data/oppna-data

Swedish Work Environment Authority, 2020. Arbetsplatsens utformning (No. AFS 2020:1).

Tsilingiris, P.T., 2004. On the thermal time constant of structural walls. Applied Thermal Engineering 24, 743–757. https://doi.org/10.1016/j.applthermaleng.2003.10.015

Tsilingiris, P.T., 2002. On the transient thermal behaviour of structural walls — the combined effect of time varying solar radiation and ambient temperature. Renewable Energy 27, 319–336. https://doi.org/10.1016/S0960-1481(01)00167-7

Verbeke, S., Audenaert, A., 2018. Thermal inertia in buildings: A review of impacts across climate and building use. Renewable and Sustainable Energy Reviews 82, 2300–2318. https://doi.org/10.1016/j.rser.2017.08.083

Wang, X., Ding, F., 2016. Joint Estimation of States and Parameters for an Input Nonlinear State-Space System with Colored Noise Using the Filtering Technique. Circuits Syst Signal Process 35, 481–500. https://doi.org/10.1007/s00034-015-0071-z

Wenzel, T.A., Burnham, K.J., Blundell, M.V., Williams, R.A., 2006. Dual extended Kalman filter for vehicle state and parameter estimation. Vehicle System Dynamics 44, 153–171. https://doi.org/10.1080/00423110500385949

Wu, S., Sun, J.-Q., 2012. A physics-based linear parametric model of room temperature in office buildings. Building and Environment 50, 1–9. https://doi.org/10.1016/j.buildenv.2011.10.005

Xu, L., Ding, F., Gu, Y., Alsaedi, A., Hayat, T., 2017. A multi-innovation state and parameter estimation algorithm for a state space system with d-step state-delay. Signal Processing 140, 97–103. https://doi.org/10.1016/j.sigpro.2017.05.006

Yarbrough, I., Sun, Q., Reeves, D.C., Hackman, K., Bennett, R., Henshel, D.S., 2015. Visualizing building energy demand for building peak energy analysis. Energy and Buildings 91, 10–15. https://doi.org/10.1016/j.enbuild.2014.11.052

Appendix A

In this appendix an attempt to simultaneously estimate indoor-temperature and model parameters is presented. The attempt was cancelled during the thesis project because measurements of indoor temperature became available. Therefore, the results are inconclusive. However, both the process and results are deemed valuable for future studies. This appendix ends with a literature review valuable for future studies.

State-space representation

State-space models are commonly used in grey-box modelling as they allow the modeller to parametrise differential equations that provide a physical description of the system (Afram and Janabi-Sharifi, 2014; Ljung and Glad, 2004). The models relate the inputs and outputs of a system through a set of system states. Written as

$$x(t+1) = f\big(x(t), u(t)\big) = Ax(t) + Bu(t) + Nv_1(t) \tag{59}$$

and

$$y(t) = h\big(x(t), u(t)\big) = Cx(t) + Du(t) + v_2(t), \tag{60}$$

a state-space representation describes a change in system states $x(t)$ which influence the output of the system is created, where A, B, C, D and N are matrices containing parameters relating states, inputs, outputs and noise $v_1(t)$ and $v_2(t)$. State-space models allow the modeller to design the systems' structure using the physical laws that govern it whilst using statistical measurements to estimate the elements of the parameter matrices. Further, the modeller can directly describe any parameter using physical insight, not needing to estimate it if not necessary (Ljung and Glad, 2004).

State observation

The states in a state-space model can have physical interpretations. Ideally, the system states are measured, but that is not always the case (Wenzel et al., 2006). They may be important for analysing the system and ensure that it is working and behaving as expected. States that can, for example, resemble quantities such as temperature or velocity which may be importance for safety reasons. Thus, there has been a lot of work and many publications regarding what is called *state observation*.

A system is observable if its observability matrix

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \tag{61}$$

has full rank, where n is the length of matrix A (Glad and Ljung, 2003). If so, then what is called an *observer* can be used to estimate the system states given an initial estimate and the measured input- and output signals. The standard observer can be introduced

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) + K\big(y(t) - \hat{y}(t)\big) \tag{62}$$

and

$$\hat{y}(t) = C\hat{x}(t) + Du(t), \tag{63}$$

where $\hat{x}$ and $\hat{y}$ denotes estimations and K is an error gain determining how much the prediction error should influence the state-estimation (Glad and Ljung, 2003).

The Kalman filter

When doing a state-estimation the error gain K must be determined. Kalman (1960) developed a solution to choosing the value of K by analysing the error-covariance. For a linear time invariant system with Gaussian noise and known covariance, Kalman's method is an optimal estimator and the Kalman filter (KF) has thus become a staple in both modelling and control theory (Glad and Ljung, 2003). Essentially, the KF adapts the uncertainty of the model to estimate the state value. The filter works in a model by predicting a state-estimate given the previous estimate and input

$$\hat{x}(t|t-1) = A\hat{x}(t-1|t-1) + Bu(t) \tag{64}$$

and its covariance

$$P(t|t-1) = AP(t-1|t-1)A^T + Q, \tag{65}$$

where Q is the covariance of the process noise $v_2(t)$. The innovation residual

$$\tilde{y}(t) = z(t) - C\hat{x}(t|t-1) \tag{66}$$

is the residual given the initial state-estimate, where z is the observation of the states. The innovation covariance is defined as

$$S(t) = CP(t|t-1)C^T + R, \tag{67}$$

where R is the observation covariance. Finally, the *Kalman gain*

$$K = P(t|t-1)C^T S^{-1}(t), \tag{68}$$

is calculated and used to update the initial state-estimate by combining the Kalman gain K and model residual $\tilde{y}(t)$

$$\hat{x}(t|t) = \hat{x}(t|t-1) + K\tilde{y}(t), \tag{69}$$

the state-estimate is updated (or corrected) by analysing both covariance and residuals.

Estimated states-based recursive least squares algorithm

State-estimation using the KF works from a defined model. However, the purpose of this thesis is to develop model. The challenge, therefore, is twofold as both states and parameters are to be estimated.

Ding (2014) presented an algorithm for simultaneous estimation of both states and parameters of a state-space model. This is very useful as states may not be possible to measure in real life and parameters may be unknown, as was with this study up to a point during the project. Ding presented an estimated states-based recursive least squares (ES-RLS) algorithm where the states are estimated with a KF using parameters that have been obtained as a least squares estimate. The algorithm was designed for single-input observable canonical state-space systems. However, as this thesis seeks to retain the physical interpretations and has two inputs, the algorithm is analogously adapted to the given state-space system.

Model representation

A linear function is derived such that the output can be described by a vector multiplication

$$y(t) = \varphi^T(t)\theta + v_2(t), \tag{70}$$

of the parameter vector

$$\theta = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix}, \tag{71}$$

where $\boldsymbol{a}$ and $\boldsymbol{b}$ are column vectors containing unknown elements in the A and B matrices respectively, and information vector $\varphi(t)$ contains relevant input- and state values at time t.

State-estimation

The state-estimation is based on the KF, but as there are unknown elements in the matrices the KF is not fully defined. Ding (2014) proposes to estimate the parameter vector $\hat{\theta}$ and to construct estimates of the matrices with unknown elements. The KF using estimated parameters (adapted for this thesis) is

$$\begin{gathered} \hat{x}(t+1) = \hat{A}(t)\hat{x}(t) + \hat{B}(t)u(t) + K_2(t)\big[y(t) - \hat{C}(t)\hat{x}(t)\big], \\ \hat{x}(1) = \frac{\mathbf{1}_n}{p_0}, \end{gathered} \tag{72}$$

$$K_2(t) = \hat{A}(t)P_2(t)\hat{C}\big[1 + \hat{C}(t)P_2(t)\hat{C}^T(t)\big]^{-1} \tag{73}$$

and

$$P_2(t+1) = \hat{A}(t)P_2(t)\hat{A}^T(t) - K_2(t)\hat{C}(t)P_2(t)\hat{A}^T(t), \;\; P_2(1) = \boldsymbol{I}_n, \tag{74}$$

where $\widehat{\;}$ denotes an estimate.

Parameter estimation

By least squares estimation, the estimate of parameters θ at time t is obtained by minimizing the squared error

$$\hat{\theta}(t) = \underset{\theta}{arg\,min} \sum_{j=1}^{t} [y(j) - \varphi^T(t)\theta]^2, \tag{75}$$

with respect to the parameter vector.

From the least squares estimate a recursive algorithm is obtained. The estimate of θ at time t is updated from its previous estimate

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P_1(t)\varphi(t)\big[y(t) - \varphi^T(t)\hat{\theta}(t-1)\big], \qquad \hat{\theta}(0) = \frac{\mathbf{1}_m}{p_0}, \tag{76}$$

$$P_1^{-1}(t) = P_1^{-1}(t-1) + \varphi(t)\varphi^T(t), \qquad P_1(0) = p_0\boldsymbol{I}_m, \tag{77}$$

where $\mathbf{1}_m$ denotes a m-dimensional column vector consisting of ones, m is the dimension of the information vector and p_0 is chosen as a large positive number such as 10^6.

The recursive algorithm still depends on knowing the information vector, but since it contains unmeasured states that is not possible. Instead, it is replaced by an estimate $\hat{\varphi}(t)$ to rewrite (76) and (77) to

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P_1(t)\hat{\varphi}(t)\big[y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)\big], \qquad \hat{\theta}(0) = \frac{\mathbf{1}_m}{p_0}, \tag{78}$$

and

$$P_1^{-1}(t) = P_1^{-1}(t-1) + \hat{\varphi}(t)\hat{\varphi}^T(t), \qquad P_1(0) = p_0\boldsymbol{I}_m. \tag{79}$$

Further, the inverse of $P_1(t)$ eq. (79) can be rewritten using the matrix inversion lemma

$$(A - BC)^{-1} = A^{-1} - A^{-1}B(\boldsymbol{I} + CA^{-1}B)^{-1}CA^{-1} \tag{80}$$

to

$$\begin{aligned} P_1(t) = P_1(t-1) \\ - P_1(t-1)\hat{\varphi}(t)[1 + \hat{\varphi}^T(t)P_1(t-1)\hat{\varphi}(t)]^{-1}\hat{\varphi}^T(t)P_1(t \\ - 1), \end{aligned} \tag{81}$$

obtaining a formula for the covariance of the parameter estimates $P_1(t)$. The gain vector $K_1(t)$ is introduced as

$$K_1(t) = P_1(t)\hat{\varphi}(t) = P_1(t-1)\hat{\varphi}(t)[1 + \hat{\varphi}^T(t)P_1(t-1)\hat{\varphi}(t)]^{-1}, \tag{82}$$

which gives

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K_1(t)\big[y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)\big], \qquad \hat{\theta}(0) = \frac{\mathbf{1}_m}{p_0} \tag{83}$$

and

$$P_1(t) = [\boldsymbol{I}_m - K_1(t)\hat{\varphi}^T(t)]P_1(t-1). \tag{84}$$

Combined state and parameter estimation algorithm

By combining eq. (82), (83) and (84), the ES-RLS algorithm was adapted for the state-space model relevant in this thesis. Eq. (83) recursively estimates the unknown parameters $\hat{\theta}(t)$ at time t using the states $\hat{x}(t)$ in the estimation vector $\hat{\varphi}(t)$ and in turn uses the estimated parameters to compute the estimate of the state vector. The algorithm is as follows:

1) Initialise by letting $t = 1$, $\hat{\theta}(0) = \frac{\mathbf{1}_m}{p_0}$, $P_1(0) = p_0 \boldsymbol{I}_m$, $\hat{x}(1) = \frac{\mathbf{1}_n}{p_0}$, $P_2(1) = \boldsymbol{I}_n$ and $p_0 = 10^6$.
2) Form the information vector $\hat{\varphi}(t)$ from the measured inputs and estimated states.
3) Compute the gain vector $K_1(t)$ and covariance matrix $P_1(t)$ with eq. (82) and (84) respectively. Then use the gain vector and covariance matrix to compute the parameter estimate $\hat{\theta}(t)$ using eq. (83).
4) Read the values from $\hat{\theta}(t)$ to construct matrices $\hat{A}(t)$, $\hat{B}(t)$ and $\hat{C}(t)$.
5) Compute the state gain vector $K_2(t)$ and covariance matrix $P_2(t+1)$ with eq. (73) and (74) respectively. From the state gain vector, covariance matrix, matrices $\hat{A}(t)$, $\hat{B}(t)$ and $\hat{C}(t)$, measured inputs and outputs calculate the state-estimate $\hat{x}(t+1)$ using eq. (72).
6) Increase t by 1 and go to step 2.

State-space model representation

Alternatively, the indoor temperature can be treated as a varying system state and as such, warranting a state-space model representation. The energy balance is a differential equation and can therefore be written as a state-space representation. Derived from the energy balance (44) the state vectors and parameter matrices in the state-space representation (eq. (62) and (63)) are constructed as

$$x(k) = \begin{bmatrix} P_{heat}(t) \\ T_{ia}(t) \\ T_{ia}(t-1) \end{bmatrix}, \tag{85}$$

$$u(k) = \begin{bmatrix} T_{oa}(t) \\ \psi(t) \end{bmatrix}, \tag{86}$$

$$v_1(t) = P_{internal}(t) = e(t), \tag{87}$$

$$v_2(t) = e(t), \tag{88}$$

$$\hat{y}(t) = \hat{P}_{heat}(t), \tag{89}$$

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ a_{21} & a_{22} & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{cases} a_{12} = k + h_{wa}S_{wa}P_1 + h_{wd}S_{wd}P_3 \\ a_{13} = -k \\ a_{21} = -\dfrac{1}{a_{13}} \\ a_{22} = \dfrac{a_{12}}{k} \end{cases}, \tag{90}$$

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ 0 & 0 \end{bmatrix}, \begin{cases} b_{11} = h_{wa}S_{wa}P_2 + h_{wd}S_{wd}P_4 \\ b_{12} = I \\ b_{21} = b_{11}k \\ b_{22} = b_{12}k \end{cases} \tag{91}$$

and

$$C = [1 \quad 0 \quad 0], \tag{92}$$

where $e(t)$ is white noise. This state-space representation describes the heating energy and indoor temperature as system states and relates it to the output of the system (the energy consumption). It can be noted that the system is not linear with regards to the parameters. Written on this form, the system's observability matrix

$$\mathcal{O} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a_{12} & a_{13} \\ a_{12}a_{21} & a_{12}a_{22} + a_{13} & 0 \end{bmatrix} \tag{93}$$

has full rank as long as elements $\mathcal{O}_{32}$ and $\mathcal{O}_{23}$ are nonzero and thus the system should be observable. Elements $\mathcal{O}_{32}$ and $\mathcal{O}_{23}$ are nonzero if $a_{13} \neq 0$ and $a_{13} \neq -a_{12}a_{22}$. The states (indoor temperature) should then be observable in the state-space representation.

Applying the algorithm derived from Ding (2014) the state-space representation

$$\hat{x}(t+1) = \hat{A}\hat{x}(t) + \hat{B}u(t) + K\big(y(t) - C\hat{x}(t)\big) \tag{94}$$

and

$$\hat{y}(t) = C\hat{x}(t) + v_2(t) \tag{95}$$

includes both estimates of states

$$\hat{x}(t) = \begin{bmatrix} \hat{P}_{heat}(t) \\ \hat{T}_{ia}(t) \\ \hat{T}_{ia}(t-1) \end{bmatrix} \tag{96}$$

and parameter matrices

$$\hat{A} = \begin{bmatrix} 0 & \hat{a}_{12} & \hat{a}_{13} \\ \hat{a}_{21} & \hat{a}_{22} & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{97}$$

and

$$\hat{B} = \begin{bmatrix} \hat{b}_{11} & \hat{b}_{12} \\ \hat{b}_{21} & \hat{b}_{22} \\ 0 & 0 \end{bmatrix}, \tag{98}$$

where the relationships from (90) and (91) are assumed to translate such that $\hat{a}_{21} = -\frac{1}{\hat{a}_{13}}$, $\hat{a}_{22} = \frac{\hat{a}_{12}}{k}$, $b_{21} = b_{11}k$ and $b_{22} = b_{12}k$.

Results

The ES-RLS algorithm applied to the energy balance (44) gives asymptotic parameter estimates, leading to unstable state-estimates. Different combinations of relating the parameters to each other (see (91)) at best gives unstable estimates, at worst approaches dividing by zero and creates an error.

The output of the model can be kept similar to the measured output if the parameter estimation is constricted by pre-determining the value of k. This indicates that the feedbacked model works as intended and that state observation and parameter estimation balance each other. However, if parameters are pre-determined, there is no reason to estimate them.

Discussion

Given the results, it seems unlikely that the ES-RLS algorithm can be applied in a practical manner to estimate both the indoor temperature and parameters in the energy balance simultaneously. There could be many reasons for this being the case and two reasons will shortly be discussed. Firstly, it is not an ideal system. It is likely that there are both non-linearities and noise due to human interaction with the building. This may produce residual errors which makes the ES-RLS impractical. Secondly, the model is non-linear with regards to the parameters. This perhaps seems to engage a volatile parameter estimation leading to asymptotic results.

There is however possibility for development of the state-space and simultaneous estimation algorithm. Firstly, the model is developed after the ideal energy balance. As can be seen in the main results of this thesis, there are time delays inherent to the system which are not included in the state-space representation, but perhaps could be. Secondly, the ES-RLS algorithm is designed for linear estimates and has been proven in simple and controlled environments (Ding, 2014). Better results may be obtained if algorithms developed for more complex systems were used. For example, algorithms developed for non-linear systems or that include probabilities of states. To end this appendix, a literature review of simultaneous estimation is provided.

Literature review for future studies

Methods for simultaneous estimation includes Ding (2014), Pavelková and Kárný (2014), Wang and Ding (2016), Wenzel et al. (2006) and Xu et al. (2017). Ding (2014) presented a modest algorithm to simultaneously estimate both states and parameters of a linear state-space system using a KF and recursive least squares estimates (the ES-RLS algorithm applied in this appendix). More complex KF-based algorithms designed for non-linear systems were presented by Wang and Ding (2016) and Wenzel et al. (2006). The previous applies an over-parameterisation technique to describe the system as two sub-models and estimates the states by using the estimated parameters and vice versa. This algorithm allows the modeller to consider coloured noise. The latter algorithm makes use of a dual extended KF (DEKF) technique to divide the state and parameter estimation into different but dependent problems. Further, Wenzel et al. (2006) demonstrate that their DEKF is appropriate given a good model structure and stable states. The DEKF also allows for the algorithm to adapt focus on state-estimation if the parameter estimation is considered appropriate enough and vice versa. Xu et al. (2017) presented a multi-innovation gradient-based algorithm that included state-delays. Lastly, Pavelková and Kárný (2014) presented a joint parameter and state-estimation for linear state-space models which included Bayesian modelling (a generalised case of Kalman filtering) to bind the state-estimates to realistic values.

Appendix B

In Table 9 the three model structures with the lowest MSE are shown for each building and fold. MAE, MBE and correlation coefficient is shown alongside MSE.

The numbers written in columns $A(q)$, $B_2(q)$ and $B_3(q)$ indicate the time-delays included in the respective polynomial. For example, the lowest MSE model structure for building #1, fold 1, has polynomials

$$A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + a_{24} q^{-24} \tag{99}$$

and

$$B_2(q) = b_{2,0} + b_{2,1} q^{-1} + b_{2,6} q^{-6} + b_{2,7} q^{-7} + b_{2,8} q^{-8} + b_{2,20} q^{-20}. \tag{100}$$

Table 9. Quantitative measures from cross-validating simulation results of the three ARX-model structures which perform the lowest MSE for different buildings and folds. Indoor temperature is considered constant.

Building	Val. fold	$A(q)$	$B_2(q)$	$B_3(q)$	MAE	(Std.)	MBE	(Std.)	MSE	CC
#1	1	[1-2-24]	[0-1-6-7-8-20]	[0-10]	11.04	8.83	-0.36	14.13	199.77	0.61
		[1-2-24]	[0-1-6-7-8-20]	[0]	10.97	8.97	-1.12	14.13	200.86	0.61
		[1-2-24]	[0-1-6]	[0-10-20]	11.14	8.84	0.18	14.23	202.36	0.60
	2	[1-2-24]	[0-1-6-7-8-20]	[0]	18.03	13.41	6.64	21.48	505.03	0.87
		[1-2]	[0-1-6]	[0]	18.66	12.87	6.70	21.66	513.92	0.87
		[1-2]	[0-1-6-12]	[0]	18.77	13.01	7.22	21.66	521.23	0.87
	3	[1-2-24]	[0-1-6-7-8]	[0]	19.51	16.52	1.88	25.50	653.34	0.76
		[1-2-24]	[0-1]	[0]	19.63	16.48	2.54	25.50	656.63	0.76
		[1-2-24]	[0-1-6]	[0-10-20]	19.74	16.44	5.13	25.18	659.92	0.76
	4	[1-2-24]	[0-1-8-20]	[0-10-20]	20.42	14.21	-2.44	24.76	618.92	0.80
		[1-2-24]	[0-1-8-20]	[0-10-20]	20.42	14.21	-2.44	24.76	618.92	0.80
		[1-2-24]	[0-1-6-7-8-20]	[0]	20.42	14.24	-2.28	24.79	619.57	0.80
#2	1	[1-2-3-24]	[0-1]	[0-1-12]	3.94	1.72	-3.90	1.81	18.49	0.78

		[1-2-24]	[0-1-12]	[0-15]	3.96	1.74	-3.92	1.82	18.67	0.78
		[1-2 3-24]	[0-1-2-3-12-24]	[0-1-2-12-24]	3.98	1.69	-3.89	1.90	18.71	0.76
	2	[1-2-12-24]	[1-2]	[0]	3.36	2.55	-0.20	4.22	17.81	0.89
		[1-2-12-24]	[0-1-2-24]	[0]	3.41	2.52	-1.48	3.98	17.98	0.90
		[1-2-3-24]	[0-1-2-3-6]	[0-1-2-12]	3.50	2.40	-1.37	4.02	17.99	0.90
	3	[1-2-12-24]	[1-2]	[0]	3.52	2.13	1.51	3.82	16.88	0.87
		[1-2-3-24]	[0-1]	[0]	3.79	2.26	2.15	3.85	19.47	0.86
		[1-2-24]	[0-1-12]	[0-8]	3.80	2.24	2.17	3.85	19.49	0.86
	4	[1]	[0]	[0]	3.35	2.58	2.63	3.31	17.85	0.91
		[1-2]	[0]	[0]	3.52	2.58	2.90	3.26	19.04	0.91
		[2]	[0]	[0]	3.52	2.61	2.90	3.29	19.19	0.91
#3	1	[1-2-3-24]	[0-1-2-8-24]	[0-1-2-3]	2.09	2.16	0.02	3.01	9.05	0.87
		[1-2-3-24]	[0-1-2-8-24]	[0-1-2-15]	2.09	2.17	0.12	3.01	9.06	0.87
		[1-2-3-24]	[0-1-2-8-24]	[0-1-2]	2.09	2.16	0.03	3.01	9.06	0.87
	2	[1-2-3-24]	[0-1-2-8-24]	[0-1-15]	1.88	1.50	-0.64	2.32	5.78	0.64
		[1-2-3-24]	[0-1-2-8-24]	[0-15]	1.88	1.50	-0.65	2.32	5.79	0.64
		[1-2-3-24]	[0-1-2-8-24]	[0-1]	1.89	1.51	-0.72	2.31	5.86	0.64
	3	[1-2-3-24]	[0-1-2-8-24]	[0-1-2-8]	3.38	2.81	-0.10	4.40	19.36	0.88
		[1-2-3-24]	[0-1-2-8-24]	[0-1-2]	3.39	2.81	-0.13	4.40	19.37	0.88
		[1-2-3-24]	[0-1-2-8-24]	[0-1-2-15]	3.40	2.80	-0.31	4.39	19.39	0.88

	4	[1-2-3-24]	[0-1-2-8-24]	[0]	1.84	1.69	0.29	2.48	6.22	0.93
		[1-2-3-24]	[0-1-2-8-24]	[0-1-2]	1.85	1.67	0.25	2.49	6.24	0.93
		[1-2-3-24]	[0-1-2-8-24]	[0-1-2-3]	1.85	1.67	0.25	2.49	6.24	0.93
	5	[1-2-3-24]	[0-1-2-8-24]	[0]	1.73	1.43	-0.06	2.24	5.03	0.94
		[1-2-3-24]	[0-1-2-8-24]	[0-1-2-8]	1.78	1.41	0.41	2.24	5.16	0.94
		[1-2-3-24]	[0-1-2-8-24]	[0-1-2]	1.78	1.41	0.43	2.23	5.17	0.95
#4	1	[1-2-3]	[0-1-8]	[0-1-2-11]	7.43	5.59	-0.80	9.27	86.48	0.91
		[1-2-3]	[0-1-2-3]	[0-1-2-11]	7.47	5.57	-1.05	9.26	86.76	0.91
		[1-2-3]	[0-1-10]	[0-1-2-11]	7.45	5.60	-0.59	9.30	86.78	0.91
	2	[1-2-3]	[0-12-24]	[0-1-2-11]	5.16	5.87	-2.74	7.32	61.10	0.59
		[1-2-3]	[0-8-12]	[0-1-2-11]	5.16	5.94	-2.91	7.31	61.90	0.59
		[1-2-3]	[0-1-12-24]	[0-1-2-11]	5.20	5.93	-2.96	7.31	62.19	0.59
	3	[1-3]	[0-1-2-8]	[0-1-2-8-18]	9.23	7.00	1.65	11.47	134.11	0.86
		[1-2-8]	[0-1-12-24]	[0-1-2-11]	9.46	6.82	-2.11	11.47	135.90	0.85
		[1-3]	[0-1-2-3]	[0-1-2-8-18]	9.41	7.12	1.99	11.64	139.31	0.86
	4	[1-2-12-24]	[0-1-24]	[0-3-10]	7.39	5.55	0.10	9.24	85.38	0.83
		[1-2-12-24]	[0-1-8-24]	[0 -1-10]	7.50	5.54	-0.65	9.30	86.91	0.83
		[1]	[0-1-12-24]	[0-1-2-10]	7.13	6.04	-2.25	9.07	87.37	0.84
	5	[1-3]	[0-1-2]	[0-1-2-6-10]	12.70	10.96	8.65	14.38	281.48	0.66

		[1-3]	[0-1-2]	[0-1-2-12]	12.79	10.88	8.81	14.30	282.07	0.65
		[1-3]	[0-1-2]	[0-1-2]	12.86	10.90	8.90	14.32	284.24	0.65
#5	1	[1-2-3-24]	[0-1-2-3]	[0-10]	1.63	1.60	-1.36	1.84	5.22	0.87
		[1-2-3-24]	[0-1-2-3]	[0-8]	1.65	1.65	-1.33	1.91	5.43	0.87
		[1-2-3-24]	[0-1-2-3]	[0-1-2-3-8]	1.66	1.65	-1.33	1.93	5.48	0.87
	2	[1-2-5]	[0-1-2-3-15]	[0-8-24]	1.03	0.94	0.58	1.28	1.96	0.88
		[1-2-3]	[0]	[0]	1.04	1.00	0.48	1.35	2.07	0.87
		[1-2-3-24]	[0-1-2-3-12]	[0-8]	1.09	0.95	0.59	1.33	2.10	0.87
	3	-	[0-1-2]	[0]	1.40	2.07	0.00	2.50	6.26	0.36
		[1-2-3-24]	[0-1-2-3-12-24]	[0-8-24]	1.60	1.94	-0.44	2.47	6.31	0.36
		[1-2-3-8-24]	[0-1-2-3]	[0-8]	1.63	1.92	-0.63	2.44	6.34	0.41
	4	[1-2]	[0]	[0]	2.91	2.19	1.40	3.36	13.23	0.60
		[1-3]	[0]	[0]	2.97	2.24	1.51	3.40	13.87	0.59
		-	[0]	[0]	3.02	2.31	0.96	3.68	14.45	0.47
#6	1	[1-2-12-24]	[0-1-2-16-24]	[0-3-20]	4.91	2.55	-3.75	4.07	30.66	0.88
		[1-2-24]	[0-1-2-16-24]	[0-3-20]	4.94	2.55	-3.77	4.08	30.84	0.88
		[1-2-12-24]	[0-1-2-24]	[0-3]	4.90	2.63	-3.62	4.22	30.95	0.88
	2	[1-2-12-24]	[0-1-2-24]	[0-3-18-19-20]	4.29	4.28	-3.39	5.02	36.70	0.80
		[1-2-24]	[0-1-2-16-24]	[0-3-20]	4.34	4.32	-3.40	5.09	37.48	0.79
		[1-2-12-24]	[0-1-2-16-24]	[0-3-20]	4.33	4.33	-3.39	5.10	37.49	0.79
	3	-	[0]	[0]	4.18	2.57	3.25	3.67	24.06	0.79
		-	[0-1]	[0]	4.24	2.57	3.42	3.59	24.63	0.80

		-	[0-1-2]	[0]	4.30	2.58	3.52	3.57	25.16	0.80
	4	[1-2-24]	[0-1-2-24]	[0-3]	4.80	3.97	3.08	5.42	38.83	0.76
		[1-3]	[0-1]	[0-24]	4.98	4.10	3.35	5.51	41.56	0.75
		[1-2]	[0-1]	[0]	5.08	4.09	3.45	5.53	42.55	0.75
	5	[1-2-24]	[0-1-2-24]	[0-3-18-19-20]	2.85	2.33	-0.61	3.63	13.51	0.92
		[1-2-24]	[0-1-2-24]	[0-3-20]	2.86	2.34	-0.46	3.66	13.63	0.92
		[1-2-12-24]	[0-1-2-16-24]	[0-3-20]	2.95	2.38	-0.69	3.73	14.37	0.92
#7	1	[1-2-24]	[0-1-2]	[0]	9.46	6.54	6.68	9.37	132.30	0.87
		[1-2-24]	[0-8-24]	[0-1-2]	9.74	6.32	7.90	8.51	134.87	0.89
		[1-2-24]	[0-1]	[0]	9.74	6.66	7.28	9.29	139.22	0.87
	2	[1-2-24]	[0-8-24]	[0-12]	4.94	3.60	0.17	6.11	37.39	0.73
		[1-2-24]	[0-8-24]	[0]	4.89	3.74	-0.66	6.13	37.96	0.73
		[1-2-3-24]	[0-8-24]	[0-12]	5.06	3.56	0.95	6.12	38.32	0.72
	3	[1-2-3-24]	[0-1]	[0]	8.30	7.28	-1.71	10.91	121.86	0.88
		[1-2-24]	[0-8-24]	[0-8]	8.23	7.54	0.78	11.14	124.61	0.87
		[1-2-24]	[0-8-24]	[0-12]	8.73	7.06	-3.42	10.70	126.05	0.88
	4	[1-2-3-24]	[0-8-24]	[0-12]	11.11	7.17	-7.38	10.98	174.92	0.92
		[1-2-3-24]	[0-8-24]	[0-1-12]	11.21	7.21	-7.61	10.94	177.59	0.92
		[1-2-3-24]	[0-8-24]	[0-1-12]	11.21	7.21	-7.61	10.94	177.59	0.92
	5	[1-2-24]	[0-8-24]	[0-1-2-18]	7.95	6.87	-2.92	10.09	110.32	0.87
		[1-2-24]	[0-8-24]	[0-1-2]	8.00	6.88	-3.20	10.05	111.25	0.87
		[1-2-24]	[0-8-24]	[0-1-2-8]	8.02	6.91	-3.51	9.99	111.98	0.87