# Application of statistical models from simulated data to real data using the example of automated fault detection in building systems

Master Thesis
Department of Statistics
Ludwig–Maximilians–Universität
München



Written by:
Michael Parzinger

Supervisors:
Prof. Dr. Christian Heumann
Prof. Dr. Ulrich Wellisch

Munich 10.08.2021

I hereby declare that I have written this thesis independently and without outside assistance. Text passages, which are based literally or in the sense on publications or lectures of other authors, are marked as such.

The work has not yet been submitted to any other examination authority and has not yet been published.

Munich, 09.08.2021

Michael Parzinger

Faults in heating, ventilation and air conditioning (HVAC) systems affect the energy efficiency of buildings. Fault detection that is both economical and automated is of interest. Obtaining suitable data sets where the fault types and time points are known in real measurements to develop fault detection methods is very difficult, time-consuming, and costly.

This master thesis demonstrates the feasibility to use simulated data to train methods for detecting faults in buildings. The main objective of the thesis is to compare the simulated data with the measured data in detail using different methods to test the applicability. These methods distinguish between data-based approaches and model-based approaches. The focus of this work is on model-based approaches. Methods with standard errors, bootstrapping, and Bayesian approaches check if the simulated building behaves similarly to the real building. This work analyzes whether there is a relationship between the similarity of the data and the fault prediction accuracy. Different versions of data sets for both measured data sets and simulated data sets are available to this work and are studied for their influence on fault detection. Thus, artificial and real users were measured, and different building faults, periods of faults and climate data were considered. This thesis answers the question of what influence these factors have on similarity and fault detection and shows the difficulties in practice.

The study shows that it is possible to detect faults in real buildings using simulated data. The presented fault detection models have high specificity, whereas the sensitivity varies enormously. However, it is challenging to find the faults if the effects of the fault in simulation and measurement are too different. Overall the fault detection results fit well with the model-based similarity results.

In total, fault detection with simulation data and model-based comparisons are two feasible concepts with great potential for future improvements.

# Contents

# List of Figures

# List of Tables

# List of Algorithms

# Chapter 1

# Introduction

## 1.1 Problem description

Due to the climate crisis, it is also necessary to implement energy savings for buildings. Detecting faults in HVAC Systems (Heating, Ventilation and Air Conditioning) is one of many components to achieve the goal of having energy efficient buildings. Faults of HVAC Systems are faults in the building control system, such as a defective ventilation unit, a stuck thermostatic valve on the heating radiator, or a stuck summer bypass. For this reason, the Technical University of applied sciences in Rosenheim (THRo) carries about fault detection in the project In Situ Detection Methods [TH-17] funded by the BMWI [Bun15]. In this project, THRo deals with fault detection using statistical models. The statistical models need training data with faults already observed. A data set with observed faults is rarely available in practice. By simulating the building and its faults in the HVAC System, we can avoid this problem. The simulation includes fault and fault-free periods and is labeled accordingly. Thus, the simulation data train statistical models to detect actual building HVAC faults.

One question that arises is how accurately these models can detect the faults in actual buildings. The answer to this question depends on how similar the simulated and the measured building are. Therefore the focus of this thesis is how to evaluate the similarity between a building simulation and building measurements.

Figure 1.1 shows the fault detection process. Physical models generate the simulation data, and in an actual building, installed sensors provide the measurement data. Then different methods, including statistical models, bootstrapping, and credibility intervals, are used to check these data sets for similarity. The simulation data can be used as training data to obtain a prediction model for the faults. The measurement data serves as test data by using it as input for the prediction model. Finally, the fault detection method predicts the faults in the actual building.
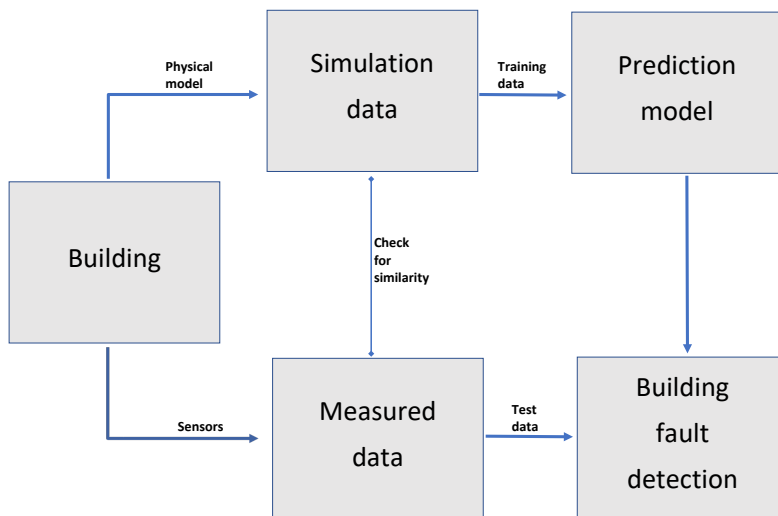
Figure 1.1: Fault detection scheme

## 1.2   Related work

In the past, the THRo studied another fault detection approach in cooperation with Annex 71 [Ann21]. In this approach, the measuring data of an actual building was split into two time periods. The first period was fault-free, and during the second period, some faults were included. The goal was to learn the behavior of the fault-free building by using the first period then to detect the faulty hours during the second phase. This fault detection method can be found in Parzinger et al. [PHS+20]. The disadvantage of this process is that it can only be used when there is a long fault-free period in the measured data set.

There are also other methodological approaches for fault detection. However, most papers about fault detection focus on the method itself and less on Data extraction. For example, Yan et al. [YSMA14] proposed a combination of SVM and ARX models or Luo et al. [LWL+19] who focused on deep learning. This work focuses on evaluating the simulation data and fault detection results using simulation data rather than on the fault detection methods themselves.

For the comparison between simulation and measurement data, Robert G. Sargent 2013 [Sar13] discussed four approaches. Other approaches can be found in Maile et al. [MBF12] he describes a method to compare simulated and measured data to assess building energy performance. Two other papers that specialize in building simulation evaluation are Figueiredo et al. [FKV+18], and Kleijnen 1995 [Kle95]. The interaction of simulated and real data is not only helpful for buildings. Under the keyword, *sim2real* works can be found, which deal with the solution of robot problems by simulation. A workshop summary can be found in [HBH+20].

## 1.3 Thesis structure

This thesis is structured as follows. One part of the thesis outlines all theoretical basics necessary for the comparison of simulation and measurement as well as fault detection (Chapters 2-7). The other part of this thesis describes the implemented concepts (Chapters 8-10).

The first chapter of the theory describes the data used to get a better idea of the challenges. Three theoretical chapters follow this. One chapter on time series since the simulated and measured data are time series - another on statistical models that uses the definitions. Moreover, a third chapter talks about evaluating binary classifiers because they are relevant for one model. Two following chapters describe the comparison of simulation and measurement data. The first one focuses on data-based comparison, while the other one introduces model-based methods.

For the implementation of the concept, there is an insert describing how missing climate data can be estimated and evaluated with the help of simulations. In the ninth chapter, the implementation of model-based comparison follows. This work reviewed two measured data and six simulations. The later chapter describes building fault detection. Finally, there is an outlook with other ideas that could lead to improved results and a summary of the findings.

# Chapter 2

# Data description

## 2.1 Experimental setup and measured data

The Laboratory for Intelligent Building Technology (LiGT) was used to generate data sets in this project. The abbreviation LiGT is derived from the German designation "Labor für intelligente Gebäudetechnik". Figure 2.1 shows a photo of the LiGT. The LiGT consists of a large, air conditionable container (service room) in which three rooms (A, B, and C) are located. The rooms are thermally decoupled from each other and from their surroundings. Each of the rooms has only one external facade. If the temperature in the rooms and in the service room is the same, heat exchange occurs only through the facade element. The energy brought into the room for heating and cooling corresponds to the heat flow through the outer facade. Rooms A and B have the same floor area of $9.85m^2$. Room C is slightly larger at $12.17m^2$, see figure 2.2.

The LiGT can be heated by an electric continuous heater type "Stiebel Eltron DHZ 15 E" with a maximum power of 15 kW. The power can be switched on in three stages. For ventilation, a central ventilation unit with supply and exhaust air of the "Wolf CGL" type was selected, with which all rooms, as well as the service room, can be simultaneously ventilated and deaerated.

In order to determine the possible influence of the user on the quality of the results, all three rooms were operated with the same boundary conditions. Only the type of user was determined differently for each room. The use of the rooms is determined in table 2.1. For the artificial user, a sedentary activity was assumed according to VDI 2078 [VDI]. At $20°C$ indoor temperature, the sensible heat emission of an adult is 90 W, and the moisture emission is 50 g/h/person. An adult person emits 22 l/h carbon dioxide ($CO_2$) [SSB19] during a light activity.

In addition to the heat emitted by the people, there were heat sources in the rooms such as building automation, lights, and computers. The heat sources of the rooms are listed in table 2.2.

A fixed schedule was set for the building services (ventilation and lighting), as well as for the users. Users were present on weekdays from 8 am to 12 am and from 1 pm to 5

Figure 2.1: LiGT (= Labor für intelligente Gebäudetechnik = Laboratory for intelligent building technology) at the Technical University of Applied Sciences in Rosenheim.

| Room | Description |
|---|---|
| Room A | Room with real user. |
| Room B | Room with artificial user. |
| Room C | Room without user. |

Table 2.1: Use of the rooms

pm. No users were present on weekends. Emissions from the artificial user were controlled according to the specified times. A maximum permissible deviation from the schedule of 30 minutes was specified for the real user. The ventilation and light were switched on simultaneously in all rooms at 6 a.m. and switched off at 8 p.m. on weekdays. On weekends, the ventilation and light were not switched on. According to DIN EN 13779 [DIN], $35 m^3/h$ per person is required in order not to exceed a $CO_2$ concentration of 1000 ppm above the ambient concentration. The supply and exhaust air volume flow for each room was set at $35 m^3/h$.
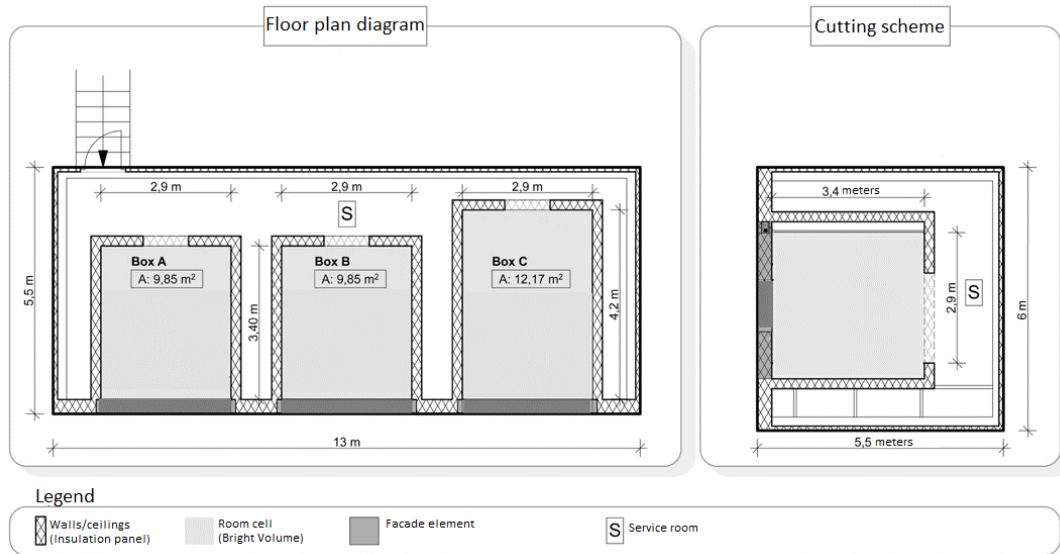
Figure 2.2: LiGT floor plan

| | Heat output in Watt | | | |
|---|---|---|---|---|
| | Light | Room controller | PC + Display | Person |
| Room A | 60 | 55 | 80 | 90 |
| Room B | 60 | 55 | 80 | 90 |
| Room C | 60 | 55 | 0 | 0 |

Table 2.2: Heat sources LiGT rooms

Three faults were incorporated for the series of experiments, which are described in table 2.3.

| Name | Fault | Impact |
|---|---|---|
| Bypass fault | Bypass ventilation system failure. | There is no heat recovery. |
| Ventilation fault | Ventilation system failure | There is no ventilation. |
| Heating fault | Heating radiator valve is stuck. | There is no heating. |

Table 2.3: Fault description

The bypass and ventilation faults quickly affect the supply air temperature or the $CO_2$ concentration in the room. The effects of the heating fault are expected to be significantly delayed due to thermal inertia, internal heat sources, and solar radiation. Based on these assumptions, bypass and ventilation faults were scheduled for a maximum of one day, but three times each during the experimental period. Due to the delayed effects in the heating trial, this was scheduled for three continuous days, but only once in the trial period. The temporal distribution of the faults in the test period is shown in table 2.4.

Many sensors were installed to record the condition of all building components and all

| Fault | Start | End |
|---|---|---|
| Bypass | 24.11.20 17:00 | 25.11.20 17:05 |
| Heating | 30.11.20 07:30 | 02.12.20 17:00 |
| Ventilation | 07.12.20 10:00 | 07.12.20 18:00 |
| Bypass | 09.12.20 17:00 | 10.12.20 15:00 |
| Bypass | 15.12.20 09:18 | 15.12.20 19:03 |
| Ventilation | 17.12.20 17:37 | 18.12.20 17:00 |
| Ventilation | 22.12.20 12:30 | 22.12.20 16:40 |

Table 2.4: Temporal distribution of the faults in the test series.

energy flows in the LiGT and climate data. From these, a selection was made to measure the most important variables for fault detection. These quantities are listed in table 2.5. The only measurements that were not measured within this setup are the diffuse and direct radiation. These were provided by Dr. Professor Schulze of the Technical University of Applied Sciences in Rosenheim. Further radiation data can be found, for example, in [Cop21].

| Room data | Air temperature, relative humidity, $CO_2$ and heating power |
|---|---|
| Climate data | Air temperature, relative humidity, direct and diffuse radiation |

Table 2.5: Measured variables for fault detection.

The period from 23/11/2020 to 23/12/2020 is used for the analysis. Measurements were taken at a frequency of 10 seconds. Only on 01/12/2020 are the six hours between 3:00 am and 7:59 am missing. For the analysis, the data are averaged to hours. This is because the simulated data is hourly, and the measurement frequency must match with these.

Note that the radiations, the people, and the computer affect the room temperature of the LiGT more than in ordinary buildings.

## 2.2 Simulation data

The simulation-based data is generated by a building performance simulation (BPS) tool IDA Indoor Climate and Energy (IDA-ICE). IDA ICE is an NMF-language-based multizone simulation tool that was validated in several studies like [MOd+21], [KZ03] and [Equ04]. A virtual physical model of the real-built LiGT was created to obtain artificial data which an acceptable accuracy. Figure 2.3 shows how the LiGT is represented in IDA-ICE.

For the simulation, input parameters are required, and these can be divided into three types. The first of these is the building model, which contains all the given information about the building. This includes the size of the rooms, the heating, and the fan. It also
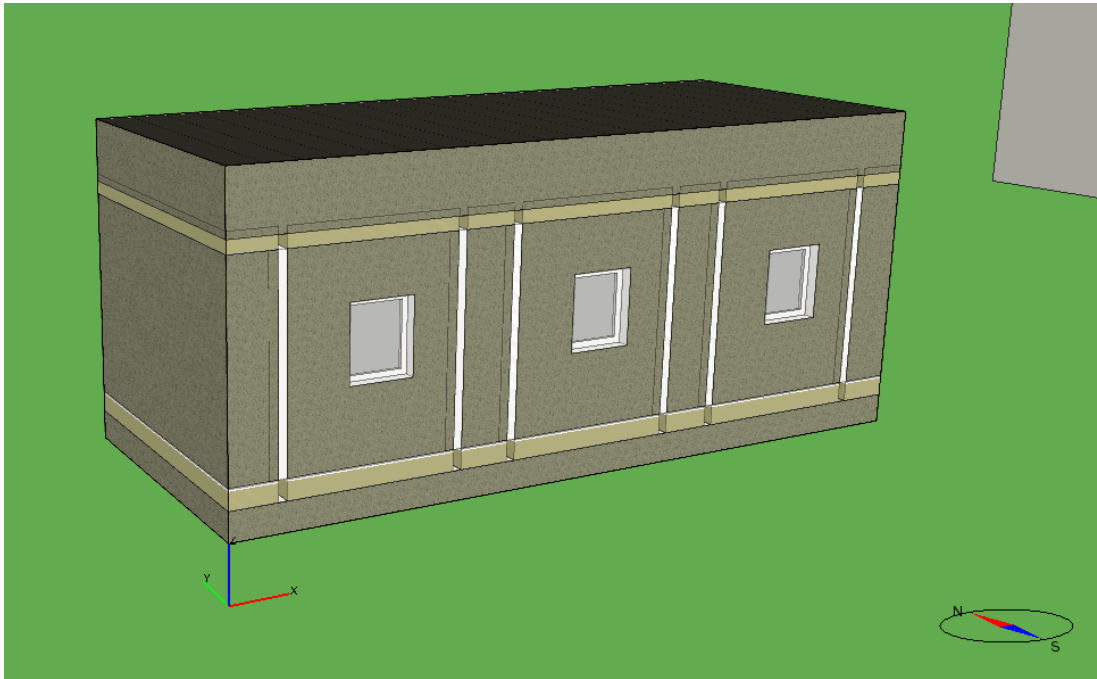
Figure 2.3: LiGT in IDA-ICE

contains information about the $CO_2$ emissions of a person and the energy consumption. For this purpose, the information from section 2.1 is transferred directly to IDA-ICE, where the room with the artificial user is taken as a reference.

Another parameter determines the event periods, which gives when which fault is active and when someone is present. The presence times are taken from section 2.1 for each simulation. For the fault periods, four different variants are simulated. One of these is based on the periods that were also used for the measurements. In the other three variants, the time periods differ from the original. The exact fault periods for the simulation can be found in appendix C. The faults were implemented by deactivating the heating, ventilation, or bypass.

The remaining input is the climate data. More detailed outdoor air temperature, diffuse radiation, direct radiation, relative humidity, wind force, and wind direction are required. The wind direction and speed are ignored in the following since they play an insignificant role. Therefore three different climate data sets are applied for the simulations. One of these is the climate data described in section 2.1. The other two climate data are example data sets that can be found in IDA-ICE. One data set is from Berlin [ASH01a], and the other is from Vienna [ASH01b].

The individual input parameters are combined to form the simulation model in which the physical relationships are calculated. An overview is shown in figure 2.4. In total, there are nine simulations, three of which are only relevant for chapter 8. Of the remaining six simulations, one uses the climate data measured at THRo as well as the real fault periods. Therefore, these simulation data will be referred to as *original* in the future. Three other

simulations use the measured THRo climate data but different fault periods. These will be referred to as *Shift1*, *Shift2*, and *Shift3* in the following. The remaining two simulations have the same fault periods as the *Shift3* simulation but additionally use different climate data. Due to these climate data, these simulations will be called *Berlin* and *Vienna* in the future.
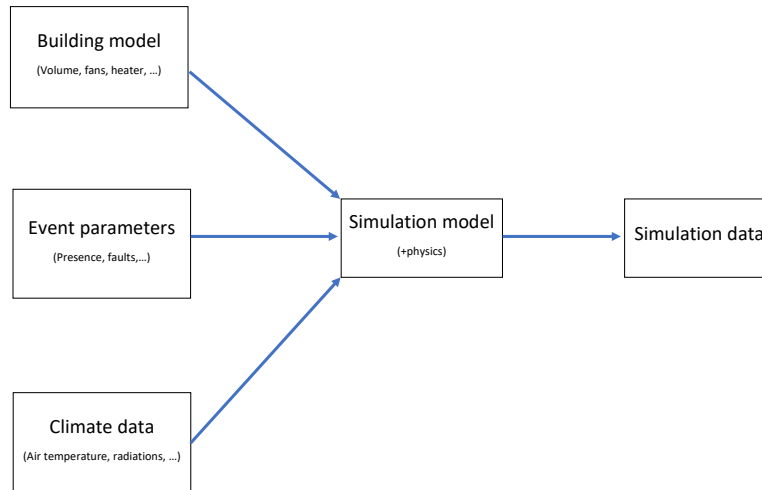


Figure 2.4: Simulation scheme

More information about IDA-ICE can be found in [EQU95].

## 2.3   Overview

While the other sections described how the data was generated, here is a summary of the essential information. For this purpose, table 2.6 shows an overview of crucial sizes and the shortcut used in the future. Below this, the units $°C$ for degree Celsius, ppm for parts per million, and $Wh/m^2$ for watt-hour per squared meter are used. Note that relative air humidity is a dimensionless measure. The outdoor relative humidity was only used to create the simulation data and not for the analysis. For that reason, it is not listed separately here.

| Shortcut | Description | Unit | Climate |
|:---:|:---:|:---:|:---:|
| $\text{RT}_t$ | Room temperature | $^\circ C$ | False |
| $\text{THP}_t$ | Total heating power | Watt | False |
| $\text{CO2}_t$ | Carbon dioxide | ppm | False |
| $\text{RH}_t$ | Relative humidity | | False |
| $\text{Pre}_t$ | Presence | | False |
| $\text{OT}_t$ | Outdoor air temperature | $^\circ C$ | True |
| $\text{difR}_t$ | Diffuse radiation | $Wh/m^2$ | True |
| $\text{dirR}_t$ | Direct radiation | $Wh/m^2$ | True |

Table 2.6: Overview of used measured variables

# Chapter 3

# Time Series

## 3.1 Introduction

A time series is a sequence of successive observations over time [BJRL15, Page 1]. Typically, these observations are made at regular time intervals. In this thesis, it is primarily hourly observations. From a statistical point of view, an observation at time $t$ is a realization of a random variable $y_t$ where $t$ represents the time point. Therefore a time series is defined in definition 3.1.1 as a realization of a stochastic process. Whereby the beginning of the process lies infinitely far in the past. Theoretically, the stochastic process exists at any time of the real number line, but it is sufficient to limit oneself to the integers. In this definition, the times between one and $n$ are those at which the observations are available.

---

**Definition 3.1.1** (Stochastic process, time series). Let $(y_t)$ a sequence of random variables and the index $t \in \mathbb{Z}$ represents the time point. Then $(y_t)$ is called a stochastic process.

The sequence of $n$ observations which are distributed like $y_1, \ldots, y_n$ is called time series [Sch20, Page 4].

---

Due to the relationship between stochastic processes, it is a task of time series analysis to determine whether a time series can be called a realization of a specific stochastic process. One example of a necessary process is the so-called white noise process which can be found in definition 3.1.2. For this process, the distribution $F$ can be chosen arbitrarily. In this thesis, principally, a normal distribution with expected value zero and variance $\sigma^2$ is assumed for each $\epsilon_t$. These processes are helpful because it is possible to check whether the residuals follow a white noise after successful modeling. If this is the case, then it can be assumed that all relevant components of the time series have been captured in the model [Sch20, Page 5].

**Definition 3.1.2** (White noise process). Let $(\epsilon_t)$ is a stochastic process and $F$ a distribution. If $\epsilon_t \sim F$ is independent identically distributed for each time point $t$, then $(\epsilon_t)$ is called a white noise process.

There are also processes that result in white noise after a transformation. One example is the random walk from definition 3.1.3. If $y_t$ is a random walk where $\text{Var}(\epsilon_t) = \sigma^2$ then $\text{Var}(y_t) = (t+1)\sigma^2 \neq t\sigma^2 = \text{Var}(y_{t-1})$. This shows that $y_t$ is not white noise process. But it holds that $y_t - y_{t-1} = \epsilon_t$ which is a white noise process by definition.

**Definition 3.1.3** (Random walk). If $(\epsilon_t)_{t \in \mathbf{N}_0}$ is a white noise process and $\mu \in \mathbb{R}$ then the stochastic process

$$y_t = \mu + \sum_{i=0}^{t} \epsilon_t$$

is called a random walk.

If $E(\epsilon_t) = 0$ then $y_t$ is a random walk without drift and for $E(\epsilon_t) \neq 0$ a random walk with drift.

In order to better characterize stochastic processes or time series, the concept of stationary is introduced in definition 3.1.4. This property means that the distribution of the process does not depend on the time $t$. This makes it easier to analyze the process. If the moments exist, then stationarity also means that the expected value and the variance are identical for each time point. Also the covariance between $y_t$ and $y_{t'}$ does only depend on the difference between $t$ and $t'$ and not on $t$ and $t'$ itself. In some books, the combination of these three properties is called stationary [Sch20, Page 16].

**Definition 3.1.4** (Stationary process). Let $(y_t)$ be a stochastic process. If $(y_t, \dots, y_{t+s})'$ has the same distribution as $(y_{t'}, \dots, y_{t'+s})'$. For each $t$, $s$ and $t'$ then $(y_t)$ is a stationary process.[HA18, 8.1]

Stationarity is an important property for analyzing time series. If a time series cannot be assumed to be stationary, then it is often necessary to decompose the time series to obtain a stationary part. This leads to the additive component model from definition 3.1.5.

**Definition 3.1.5** (Additive component model). Let $y_t$ a stochastic process. Then a additive component model can be written as

$$y_t = S_t + T_t + R_t,$$

where $S_t$ is the seasonal component, $T_t$ is the trend-cycle component and $R_t$ is the

remainder component each to the time point $t$ see [HA18, 6.1] and [Sch20, Page 26].

In this case, a time series or a stochastic process can be decomposed into a seasonal component, a trend component, and the remaining component. The goal is then to estimate trends and seasonal components to be able to analyze the remaining part. Note that there is no granite that the remaining part is stationary, but if season and trend existent, then the remainder analysis is usually more straightforward than the analysis of the complete time series. There are many ways to estimate the components in an additive component model like moving average, X11 decomposition, and STL decomposition, see [HA18, Chapter 6]. In this thesis, however, component decomposition does not play a major role, which is why it will not be discussed in detail.

## 3.2 ARIMA-Processes

For the following definitions $(y_t)$ is a stochastic process and $(\epsilon_t)$ is a white noise process for which $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$ hold for each $t$. The main goal is now to model the target variable $y_t$ by using past values of that variable. In an autoregressive process, $y_t$ is assumed to be composed as a linear combination of past observations [HA18]. In future this process is called AR(p) process see definition 3.2.1.

---

**Definition 3.2.1** (AR(p) process). Let $(y_t)$ be a stochstic process and $(\epsilon_t)$ a white noise process then a autoregressive process or AR(p) process is given by

$$y_t = \beta_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t,$$

where $\beta_0 \in \mathbb{R}$ and $\phi_i \in \mathbb{R}$ for each $i = 1, \ldots, p$.

---

This process can also be written by using the backshift notation. Therefore the backshift operator is defined in definition 3.2.2. Then a AR(p) process can be represented as

$$\Phi(B)y_t = \beta_0 + \epsilon_t,$$

if $\Phi(B) := 1 - \sum_{i=1}^{p} \phi_i B^i$.

---

**Definition 3.2.2** (Backshift operator). Let $y_t$ a observation or a random variable at time $t$. Then a backshift operator is defined as $B$ for which $By_t = y_{t-1}$ holds for each $t$.

(By this definition, it follows directly that $B^l y_t = y_{t-l}$ for each $l \in \mathbb{N}$.)

---

One important point about AR processes is the conditions for the parameters $\phi_1, \ldots, \phi_p$. Therefore theorem 3.2.1 specifies these conditions. This also shows that for an AR(1) process $|\phi_1| < 1$ must hold.

**Theorem 3.2.1** (Stationarity for AR-processes)**.** Let $\Phi(B) := 1 - \sum_{i=1}^{p} \phi_i B^i$ and $\Phi(B)y_t = \epsilon_t$ then $y_t$ is only stationary if the solutions of $\Phi(B) = 0$ are greater than one in absolute value [BJRL15, Page 9]. Where $\Phi(B)$ is considered a polynomial in $B$ of degree $p$.

In order to better understand AR processes, the simplified case p=1 and $\beta_0 = 0$ is considered in the following. This means that $y_t = \phi y_{t-1} + \epsilon_t$ holds for each $t \in \mathbb{Z}$. If $y_{t-1}$ is given then $y_t|y_{t-1} \sim \mathcal{N}(\phi y_{t-1}, \sigma^2)$ obviously holds. If $y_{t-2}$ is given instead of $y_{t-1}$ then the model equation can be used to replace $y_{t-1}$ by $\phi y_{t-2} + \epsilon_t$. As long as the stationarity condition of Theorem 3.2.1 is fulfilled, this procedure can theoretically be repeated infinitely often. In this case $y_t \sim \mathcal{N}(0, \frac{\sigma^2}{1-\phi^2})$ is valid which is shown in (3.1). This also shows the importance of theorem 3.2.1.

$$
\begin{aligned}
y_t &= \phi y_{t-1} + \epsilon_t \\
&= \phi(\phi y_{t-2} + \epsilon_{t-1}) + \epsilon_t = \phi^2 y_{t-2} + \phi \epsilon_{t-1} + \epsilon_t \\
&= \phi^2(\phi y_{t-3} + \epsilon_{t-2}) + \phi \epsilon_{t-1} + \epsilon_t \\
&\vdots \\
&= \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} \\
\Rightarrow y_t &\sim \mathcal{N}\left(0, \sigma^2 \sum_{i=0}^{\infty} \phi^{2i}\right) \overset{|\phi|<1}{=} \mathcal{N}\left(0, \frac{\sigma^2}{1-\phi^2}\right).
\end{aligned}
\tag{3.1}
$$

These distributions can now be used to set up the (log-)likelihood function and is shown in equation (3.2).

$$
\begin{aligned}
p(y_1, y_2, \ldots, y_n) &= p(y_n|y_1, \ldots, y_{n-1})p(y_1, \ldots, y_{n-1}) \\
&= p(y_n|y_{n-1}) \underbrace{p(y_{n-1}|y_1, \ldots, y_{n-2})}_{=p(y_{n-1}|y_{n-2})} p(y_1, \ldots, y_{n-3}) \\
&\vdots \\
&= p(y_1) \prod_{i=2}^{n} p(y_i|y_{i-1}) \\
&\overset{\log}{\to} \log p(y_1) + \sum_{i=2}^{n} \log p(y_i|y_{i-1})
\end{aligned}
\tag{3.2}
$$

In backshift notation there are two ways to denote a AR(1) process without intercept $y_t(1 - B\phi) = \epsilon_t$ and $y_t = \frac{\epsilon_t}{1-B\phi}$. If $\frac{1}{1-B\phi}$ is replaced by $\sum_{i=0}^{\infty} \phi^i B^i$, the representation from equation (3.1) is obtained [BJRL15][Page 9]. By that $y_t(1 - B\phi) = \epsilon_t$ can be seen as the

representation conditional to the past and $y_t = \frac{\epsilon_t}{1 - B\phi}$ is the marginal form. This can also be extended for general AR(p) processes see definition 3.2.3.

---

**Definition 3.2.3** (Backshift division). Let $\Phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$ be a backshift polynomial for which Theorem 3.2.1 holds. Then $\Phi(B)^{-1}$ is defined by

$$\frac{1}{\Phi(B)} := \sum_{j=0}^{\infty} \left( \sum_{i=1}^{p} \phi_i B^i \right)^j.$$

---

Instead of using past values of the target variable, it is also possible to use a linear combination of past errors. This process is then called moving average and is defined in definition 3.2.4.

---

**Definition 3.2.4** (MA(q) process). Let $(y_t)$ be a stochastic process and $(\epsilon_t)$ a white noise process then a moving average process or MA(q) process is given by

$$y_t = \epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{t-i} = \Theta(B)\epsilon_t$$

where $\theta_i \in \mathbb{R}$ for each $i = 1, \ldots, q$ and $\Theta(B) := 1 - \sum_{i=1}^{q} \theta_i B^i$.

---

The equation (3.1) shows that each AR(1) process can be written as MA($\infty$) process. On the other hand, a MA(1) process can be written as AR($\infty$) process, which is shown in equation (3.3). This only makes sense if $|\theta| < 1$ holds since this means the newest observations have the greatest influence on the current observation [HA18, Sec. 8.4]. Whereas $|\theta| > 1$ would mean that the further back in time the observations are, the more important, and $\theta = 1$ would mean that each observation has the same influence [HA18, Sec. 8.4]. For stationarity, on the other hand, it is true that finite MA(q) processes are always stationary while MA($\infty$) processes are only stationary if $\sum_{i=1}^{\infty} |\theta_i| < \infty$ [Sch20, Page 87].

$$
\begin{aligned}
y_t = \epsilon_t - \theta\epsilon_{t-1} &\Leftrightarrow \epsilon_t = y_t + \theta\epsilon_{t-1} \\
\Rightarrow \epsilon_t = y_t + \theta\epsilon_{t-1} &= y_t + \theta(y_{t-1} + \theta\epsilon_{t-2}) \\
&= y_t + \theta y_{t-1} + \theta^2 \epsilon_{t-2} \\
&= y_t + \theta y_{t-1} + \theta^2 (y_{t-2} + \theta\epsilon_{t-3}) \\
&= y_t + \theta y_{t-1} + \theta^2 y_{t-2} + \theta^3 \epsilon_{t-3} \\
&\vdots \\
&= \sum_{i=0}^{\infty} \theta^i y_{t-i}
\end{aligned}
\tag{3.3}
$$

It is also possible to combine AR(p) and MA(q) processes. This leads to autoregressive-moving average processes or, in short, ARMA(p,q) processes and are defined in definition 3.2.5. Note that ARMA(p,q) processes are here defined for $\phi_i \in \mathbb{R}$ and $\theta_j \in \mathbb{R}$ but for stationarity, both the AR(p) part and the MA(q) part must have the appropriate properties, such as those from theorem 3.2.1, see [Sch20, Page 92].

---

**Definition 3.2.5** (ARMA(p,q) process). Let $(y_t)$ be a stochastic process and $(\epsilon_t)$ a white noise process then a autoregressive-moving average processes or ARMA(p,q) process is given by

$$y_t - \sum_{i=1}^p \phi_i y_{t-i} = \beta_0 + \epsilon_t - \sum_{j=1}^q \theta_i \epsilon_{t-i} \quad \Leftrightarrow \quad \Phi(B)y_t = \beta_0 + \Theta(B)\epsilon_t$$

where $\beta_0 \in \mathbb{R}$, $\phi_i \in \mathbb{R}$ for each $i = 1, \ldots, p$, $\theta_i \in \mathbb{R}$ for each $j = 1, \ldots, q$, $\Theta(B) := 1 - \sum_{i=1}^q \theta_i B^i$ and $\Phi(B) := 1 - \sum_{i=1}^p \phi_i B^i$ [HA18, Chapter 8.5] or [Sch20, Page 92].

---

If a process itself does not follow a stationary ARMA(p,q) process, then it is still possible to obtain a stationary ARMA(p,q) after difference formation. This can occur, for example, when there is a trend component [Sch20, Page 100]. For this reason the autoregressive integrated moving average or ARIMA(p,d,q) model is defined in 3.2.6. This process is similar to the ARMA(p,q) process but $y_t$ gets replaced by $(1 - B)^d y_t$. Which means that for $d = 1$ the target $y_t$ gets replaced by $y_t - y_{t-1}$ and for $d = 2$ by $(y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$ and so on. Obviously a ARIMA(p,0,q) process is identical to a ARMA(p,q) process.

---

**Definition 3.2.6** (ARIMA(p,d,q) process). Let $(y_t)$ be a stochastic process and $(\epsilon_t)$ a white noise process then a autoregessive-moving average processes or ARMA(p,q) process is given by

$$\Phi(B)(1 - B)^d y_t = \beta_0 + \Theta(B)\epsilon_t$$

where $\beta_0 \in \mathbb{R}$, $\phi_i \in \mathbb{R}$ for each $i = 1, \ldots, p$, $\theta_i \in \mathbb{R}$ for each $j = 1, \ldots, q$, $\Theta(B) := 1 - \sum_{i=1}^q \theta_i B^i$ and $\Phi(B) := 1 - \sum_{i=1}^p \phi_i B^i$ [HA18, Chapter 8.5].

---

It is also possible to extend ARIMA(p,d,q) processes even further. This can be done, among other things, by adding seasonal components, see [BJRL15, Chapter 9]. This is often called a seasonal ARIMA or SARIMA model. Other extensions are the autoregressive fractionally integrated moving average (ARFIMA) model introduced by Granger [GJ08] or vector autoregression models like VARMA [HA18, Chapter 11].

# Chapter 4

# Statistical Models

This chapter describes the models used for the analysis. Among them are models that can be used to compare the simulation data and the measurement data and models that can be used for fault detection. The goal is to establish a link between a response variable $y_t$ and exogenous variables $(x_{t1}, \ldots, x_{tn})$ which are also called covariates for all $t = 1, \ldots, n$. In the following, the error term $\epsilon_t$ is mostly assumed to be normally distributed. This is not mandatory, but it facilitates the representation of the (log)-likelihood.

## 4.1 Linear regression

A classical approach to establish a relationship between $y_t$ and $(x_{t1}, \ldots, x_{tn})$ is given by (multiple) linear regression. This model is described among others in [RC12], [HA18, 5.1] and [Woo06].

---

**Definition 4.1.1** (Linear regression). Let $y_t \in \mathbb{R}$ be the response variable to the time point $t$ and $X_t = (1, x_{t1}, x_{t2}, \ldots, x_{tp}) \in \mathbb{R}^{1 \times (p+1)}$ the vector of exogenous variables to the time point $t$. Then the linear regression model is given by

$$y_t = X_t \beta + \epsilon_t,$$

where $\beta \in \mathbb{R}^{p+1}$ and $\epsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$.

---

Due to the normal distribution assumption, $y_t \overset{\text{iid}}{\sim} \mathcal{N}(X_t \beta, \sigma^2)$ holds for all $t = 1, \ldots, n$. Furthermore, due to the independence of the error term, the joint log-likelihood can be derived without major problems which is shown in equation (4.1).

$$\log p(y_1, y_2, \ldots, y_n) = \log \prod_{i=1}^{n} p(y_i)$$

$$= \sum_{i=1}^{n} \log p(y_i) \tag{4.1}$$

$$= -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - X_i\beta)^2$$

The independence of the errors simplifies the construction of the (log)-likelihood, but this assumption is unrealistic for time series data [FKL07, Page 67]. For this reason, ARMAX and dynamic regression models are introduced in the following. Linear regression models are also unsuitable for fault detection since the assumption of linearity is not meaningful for binary response variables. Therefore, logistic regression will be introduced later. Nevertheless, this model was defined here because the other models mentioned here can be considered extensions of linear regression.

## 4.2   ARMAX

In this section a way to combine ARMA models and linear regression is presented. Namely in the form of an autoregressive–moving-average with exogenous inputs model short ARMAX model.

---

**Definition 4.2.1** (ARMAX). Let $y_t$ be the response variable to the time point $t$ and $X_t = (1, x_{t1}, x_{t2}, \ldots, x_{tp}) \in \mathbb{R}^{1\times(p+1)}$ the vector of exogenous variables to the time point $t$. Then the ARMAX model is given by

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + X_t\beta + \sum_{j=1}^{q} \theta_j \epsilon_{t-i} + \epsilon_t,$$

where $\beta \in \mathbb{R}^{p+1}$, $\phi_i \in \mathbb{R}$, $\theta_j \in \mathbb{R}$ for all $i = 1, \ldots, p$ and $j = 1, \ldots, q$ and $\epsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$.

---

This model can be also written by using a backshift polynomials. Let $\Phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$ and $\Theta(B) = 1 - \sum_{i=1}^{q} \theta_i B^i$ then ARMAX is given by

$$y_t\Phi(B) = X_t\beta + \Theta(B)\epsilon_t \Leftrightarrow y_t = \frac{1}{\Phi(B)}(X_t\beta + \Theta(B)\epsilon_t) \tag{4.2}$$

An ARMAX model has a similar form as an ARMA process from definition 3.2.5 only that additional exogenous variables are added. Therefore, the conditions for stationarity described for the ARMA process also apply like theorem 3.2.1. If $y_t$ is replaced by $(1-B)^d y_t$,

then the model is also called an ARIMAX model. However, this will not be discussed further since the theory of ARMAX also covers the theory of ARIMAX models.

For now let $p = 1$, $q = 0$ and $|\phi| < 1$. Then a ARMAX model is given by

$$y_t = \phi y_{t-1} + X_t \beta + \epsilon_t,$$

which is also called a ARX Model. Then the conditional distribution of $y_t$ is given by

$$y_t | y_{t-1} \sim \mathcal{N}(\phi y_{t-1} + X_t \beta, \sigma^2). \tag{4.3}$$

For the marginal distribution it holds that

$$
\begin{aligned}
y_t &= \phi y_{t-1} + \underbrace{X_t \beta + \epsilon_t}_{=:\gamma_t} \\
&= \phi(\phi y_{t-2} + \gamma_{t-1}) + \gamma_t = \phi^2 y_{t-2} + \phi \gamma_{t-1} + \gamma_t \\
&= \phi^2(\phi y_{t-3} + \gamma_{t-2}) + \phi \gamma_{t-1} + \gamma_t \\
&\vdots \\
&= \sum_{i=0}^{\infty} \phi^i \gamma_{t-i} = \sum_{i=0}^{\infty} \phi^i (X_{t-i}\beta + \epsilon_{t-i}) \\
\Rightarrow y_t &\sim \mathcal{N}\left( \sum_{i=0}^{\infty} \phi^i X_{t-i}\beta, \frac{\sigma^2}{1-\phi^2} \right).
\end{aligned}
\tag{4.4}
$$

These distributions can be used to calculate the (log)-likelihood. For the derivation of the joint density, the steps from (3.2) can be used again. Substituting the distributions (4.3) and (4.4) into the equation (3.2) yields the following likelihood given by (4.5), (4.6) and (4.7). The term (4.7) can be written in the form of a linear regression. The term 4.6 is more complicated mainly because of two reasons. First it is not linear in $\phi$. Second the exogenous inputs $X_0, X_{-1}, X_{-2}, \ldots$ are not all available in practice. It is possible to approximate $\sum_{j=0}^{\infty} \phi^j X_{1-j}$ by using some past exogenous inputs. It can also be argued that for large sample size $n$ the influence of (4.6) is so small that it is negligible. In this case the log-likelihood $\log p(y_1, y_2, y_3, \ldots, y_n)$ is replaced by $\log p(y_2, y_3, \ldots, y_n | y_1)$ which is minimized instead.

$$\log p(y_1) + \sum_{i=2}^{n} \log p(y_i | y_{i-1}) \tag{4.5}$$

$$= \frac{1}{2} \log \frac{1-\phi^2}{2\pi\sigma^2} - \frac{1-\phi^2}{2\sigma^2} \left( y_1 - \sum_{j=0}^{\infty} \phi^j X_{1-j}\beta \right)^2 \tag{4.6}$$

$$- \frac{n}{2} \log 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_{i=2}^{n} (y_i - (\phi y_{i-1} + X_t\beta))^2 \tag{4.7}$$

## 4.3 Dynamic regression models

In addition to the ARMAX model, the dynamic regression model offers another possibility to combine ARMA or ARIMA models and regression. Dynamic regression models are often called regression with ARIMA errors [HA18]. Therefore, dynamic regression models are defined in definition 4.3.1 as regression models in which the error term follows an ARIMA process.

---

**Definition 4.3.1** (Dynamic regression models)**.** Let $y_t$ be the response variable to the time point $t$ and $X_t = (1, x_{t1}, x_{t2}, \ldots, x_{tp}) \in \mathbb{R}^{1 \times (p+1)}$ the vector of exogenous variables to the time point $t$. Then dynamic regression models are given by

$$y_t = X_t\beta + \nu_t \quad \wedge \quad \Phi(B)(1-B)^d\nu_t = \Theta(B)\epsilon_t,$$

where $\Phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$, $\Theta(B) = 1 - \sum_{i=1}^{q} \theta_i B^i$, $\beta \in \mathbb{R}^{p+1}$ and $\epsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$.

---

To estimate the coefficients, the sum of squared errors $\epsilon_t$ must be minimized. If, on the other hand, the sum of squared $\nu_t$ is minimized, then the autocorrelation in the errors is ignored [HA18, Sec. 9.1]. It is also possible to perform the estimations via maximum likelihood [HA18, Sec. 9.1]. An iterative procedure for estimating the coefficients is described in Schlittgen 2020 [Sch20, Pages 226-227]. First, the coefficients $\beta$ are estimated as in a linear regression model via minimizing the sum of the squared errors. In the next step, the coefficients $\phi_1, \ldots, \phi_p$ and $\theta_1, \ldots, \theta_q$ are estimated by maximum likelihood using the residuals. This gives the dependency structure, which will be considered in the next pass. These steps are then repeated until the results hardly change. If this procedure is continued, the standard errors of the coefficients are also obtained [Sch20, Page 227]. The proof that this procedure is suitable to estimate the coefficients is shown in [OK74].

According to [HA18, Sec. 9.1], an important condition for dynamic regression models is stationarity of all variables. This refers to $y_t$ as well as to all exogenous variables $(x_{t1}, \ldots, x_{tp}$. Otherwise, the estimates of the coefficients are not consistent. Except if there is a linear combination of non-stationary $y_t$ and exogenous variables, then the estimated coefficients are consistent even though not all variables are stationary, see [HA18, Sec. 9.1] or [RH03].

In the following it is assumed that $p = 1$, $q = d = 0$ and $|\phi| < 1$. In this case, it is a regression with AR(1) error. Then a dynamic regression model can be written as

$$y_t = X_t\beta + \nu_t \tag{4.8}$$
$$\nu_t = \phi\nu_{t-1} + \epsilon_t. \tag{4.9}$$

Under these assumptions the conditional distribution is given by

$$\nu_t | \nu_{t-1} \sim \mathcal{N}(\phi \nu_{t-1}, \sigma^2) \tag{4.10}$$

$$y_t | \nu_{t-1} \sim \mathcal{N}(X_t \beta + \phi \nu_{t-1}, \sigma^2) \tag{4.11}$$

$$y_t | y_{t-1} \sim \mathcal{N}(X_t \beta + \phi(y_{t-1} - X_{t-1}\beta), \sigma^2)$$
$$= \mathcal{N}(\phi y_{t-1} + (X_t - \phi X_{t-1})\beta, \sigma^2). \tag{4.12}$$

In (4.10) the fact that $\nu_t$ follows an AR process is used. For the equation (4.11) the regression part regression part was added. By replacing $t$ with $t-1$ the equation (4.8) can be written as $\nu_{t-1} = y_{t-1} - X_{t-1}\beta$. This $\nu_{t-1}$ is inserted into the distribution (4.11) to obtain the distribution in (4.12). This shows that $y_t | \nu_{t-1} = y_t | y_{t-1}$.

For the marginal distribution the fact that $\nu_t \sim \mathcal{N}(0, \frac{\sigma^2}{1-\phi^2})$ from equation (3.1) can be used. For the distribution of $y_t$ only the regression part has to be added to get equation 4.13.

$$y_t = X_t \beta + \nu_t \sim \mathcal{N}\left(X_t \beta, \frac{\sigma^2}{1-\phi^2}\right) \tag{4.13}$$

As with the ARMAX model, these results can be used together with equation (3.2) to establish the (log)-likelihood. The equations (4.14), (4.15) and (4.16) show the resulting log-likelihood. Unlike ARMAX, there are no problems for the term of observation at time $t = 1$, see (4.15), in terms of the exogenous variables needed. The equation (4.16) cannot simply be represented in the form of linear regression, so another numerical procedure is needed to maximize the log-likelihood - for example, the procedure described on page 20. This is mainly due to the fact that in $(X_t - \phi X_{t-1})\beta = X_t \beta - X_{t-1}\beta\phi$ two parameters $\phi$ and $\beta$ are multiplied together, which are both to be optimized.

$$\log p(y_1) + \sum_{i=2}^{n} \log p(y_i | y_{i-1}) \tag{4.14}$$

$$= \frac{1}{2} \log \frac{1-\phi^2}{2\pi\sigma^2} - \frac{1-\phi^2}{2\sigma^2}(y_1 - X_1\beta)^2 \tag{4.15}$$

$$- \frac{n}{2} \log 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_{i=2}^{n} (y_i - (\phi y_{i-1} + (X_t - \phi X_{t-1})\beta))^2 \tag{4.16}$$

When comparing ARMAX and Dynamic Regression models, the question arises: which models should be used and when. Rob J. Hyndman, the creator of the R package forecast [HAB+21] and one author of the fable [OWHW20] package, has a clear answer to this question. In his blog called "The ARIMAX model muddle" [Rob] he recommends to use the dynamic regression model. This is mainly since in ARMAX the coefficient $\beta$ does not have the same interpretation as in linear regression but can only be interpreted conditionally on the last time points. For the marginal form, the criticism is that when represented with backshift operators, the AR term is mixed with the error term and the covariates, see equation (4.2). These problems do not occur with the dynamic regression model. In addition, since no model is superior to the other in terms of predictive quality [Rob], the dynamic regression model is recommended.

## 4.4 Logistic regression

The previous models require a real valued response variable. However, these models are not suitable for binary questions such as whether a fault is present or not. For this reason, logistic regression is introduced in this section.

---

**Definition 4.4.1** (Logistic regression). Let $y_t \in \{0, 1\}$ be the response variable to the time point $t$ and $X_t = (1, x_{t1}, x_{t2}, \ldots, x_{tp}) \in \mathbb{R}^{1 \times (p+1)}$ the vector of covariates to the time point $t$. Then the logistic regression model is given by

$$\text{logit}(P(y_t = 1|X_t)) := \log \frac{P(y_t = 1|X_t)}{1 - P(y_t = 1|X_t)} = X_t \beta \Leftrightarrow P(y_t = 1|X_t) = \frac{\exp(X_t \beta)}{1 + \exp(X_t \beta)}$$

where $\beta \in \mathbb{R}^{p+1}$ and $P(y_t = 1|X_t)$ denotes the probability that $y_t = 1$ given the covariates $X_t$.

---

The estimation of the coefficients $\beta$ can be done by maximizing the log-likelihood. The log-likelihood can be derived via the distributional assumption $y_t \overset{\text{iid}}{\sim} \text{Ber}\left(\frac{\exp(X_t \beta)}{1 + \exp(X_t \beta)}\right)$ and is given in equation (4.17). For the estimation it is necessary to use a numerical procedure like the Newton–Raphson algorithm [HTF09, Page 120].

$$\log p(y_1, \ldots, y_n) = \sum_{i=1}^{n} (y_i X_i \beta - \log(1 + e^{X_t \beta})) \tag{4.17}$$

Notice that a logistic regression model estimates the probability for $y_t = 1$ given $X_t$. By setting a threshold between zero and one, these probabilities can be used to estimate whether $y_t = 1$ or $y_t = 0$ holds. There are different approaches for determining these thresholds which will be discussed in Chapter 5.

Even more details about this model can be found in [JWHT14, Sec. 4.3] and [HTF09, Sec. 4.4] among others.

## 4.5    Akaike information criterion (AIC)

Akaike information criterion or AIC is a method to estimate the quality for statistical models and is defined in definition 4.5.1. A lower AIC here indicates a better model. The AIC can also be applied step by step to find an optimal model. This procedure is then called stepwise AIC and works as follows. In the first step, a simple intercept model is assumed. In the next step, the covariate that reduces the AIC the most is added. This results in a model consisting of an intercept and a covariate. Next, the covariate that reduces the AIC the most is added again. The process is repeated until the addition of more variables would only increase the AIC. It is also possible to start with a model that uses all available covariates and reduce it to each step, then it is the Backward stepwise AIC. More information about the stepwise AIC is given in [FKL07, S.165], [BAH10] and [YYK07], among others.

---

**Definition 4.5.1** (AIC)**.** Let $L$ be the maximum likelihood and $\rho$ the number of independently adjusted parameters within the model then the AIC is given by,

$$\text{AIC} = -2\log(L) + 2\rho,$$

see [Aka74].

---

# Chapter 5

# Evaluation of binary classifiers

In fault detection, there are four possible outcomes at each time point. Either a fault is correctly suspected, correctly no fault is suspected, incorrectly a fault is suspected, or incorrectly no fault is suspected. Binary classification often distinguishes between positive and negative classifications. That is, in this thesis, time points that are classified as faults are called positive, and time points that are classified as fault-free are called negative. Also, correct classification is drawn as *true* and an incorrect one as *false*. Therefore the number of time points that are correctly classified as a fault is called *true positive* (TP) and the number of time points that are incorrectly classified as a fault is called *false positive* (FP). The same is applied with the fault-free time points, only that they are then called negative. These definitions are summarized in definition 5.0.1 and more general definitions can be found in [Pow08].

---

**Definition 5.0.1** (TP, FP, TN, FN)**.** Let time points that are classified as faults be positive and time points that are classified as fault-free be negative then:

- TP = True positive = Number of time points that are correctly classified as fault.

- FP = False positive = Number of time points that are incorrectly classified as fault.

- TN = True positive = Number of time points that are correctly classified as fault-free.

- FN = False positive = Number of time points that are incorrectly classified as fault-free.

---

The division into TP, FP, TN, and FN can derive other measures, making it easier to compare the classification results. An example of such a measure is the True Positive Rate, also called sensitivity. In this measure, only the points in time that were classified as positive are considered. From these, the proportion of correct classifications is calculated.

In contrast, specificity is the proportion of correct classifications within the time points classified as negative. Ideally, both sensitivity and specificity are close to one, but binary classification often requires a choice between high sensitivity and high specificity. Some measures consider both negative and positive classifications, such as accuracy, which describes the proportion of correct classifications. In the case of accuracy, the characteristic that occurred less frequently is also considered to a lesser extent. Other measures take both positive and negative values into account and avoid this problem, which will not be discussed further. These can be found in [Pow08], among others. Definition 5.0.2 lists the three definitions discussed.

---

**Definition 5.0.2** (Sensitivity, specificity, accuracy)**.** Given the definition 5.0.1 then:

- Sensitivity $:= \frac{TP}{TP+FP}$

- Specificity $:= \frac{TN}{TN+FN}$

- Accuracy $:= \frac{TP+TN}{TP+TN+FP+FN}$

---

Suppose the probability of a positive value is modeled for binary classification. In that case, it is possible to estimate the quality of the model using a Receiver Operating Characteristics (ROC) curve. Further information about the ROC curve can be found in [Pow08] and [HM82] among others. Simplified, we first assume a sensitivity of zero and a specificity of one. In the next step, the actual labels are sorted according to the modeled probabilities for positive. Now, these labels are gone through one by one, starting with the one that has the highest probability. For each positive observation, the sensitivity is then increased by $\frac{1}{TP+FN}$. For each negative observation, the specificity is decreased by $\frac{1}{FP+TN}$. The resulting curve is the ROC curve. The technique of sorting by probabilities can also be used to derive different thresholds. These can be applied to the training data, and finally, the threshold can be selected which is the most appropriate according to a given measure.

# Chapter 6

# Data-based comparison

This chapter discusses methods for data-based comparison. The term data-based is to be understood as the contrast to model-based, which is discussed in Chapter 7. The main difference between these two concepts is that in model-based comparison, models are created with the data, which are then compared. In data-based comparison, however, the data sets are compared directly or after a time shift. For this purpose, let $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ be two time series which are to be compared. As a simplification $x = (x_1, \ldots, x_n)' \in \mathbb{R}^n$ and $y = (y_1, \ldots, y_n)'$ is defined.

## 6.1 Methods

### Mean suqared error and mean absolute error

Simple examples for data-based comparisons are the mean squared error (MSE) and the mean absolute error (MAE), which are defined in definition 6.1.1. When the MSE is referred to in the following, the MAE is also meant. It is also possible to divide the time series $x$ and $y$ into equal blocks. Then two blocks can be compared with each other using the MSE. In this way, it is possible to examine when the distance between $x$ and $y$ is high and when it is low. This can also be done between two observations $x_i$ and $y_i$ using $(x_i - y_i)^2$.

---

**Definition 6.1.1** (MSE, MAE). Let $x = (x_1, \ldots, x_n)' \in \mathbb{R}^n$ and $y = (y_1, \ldots, y_n)' \in \mathbb{R}^n$ then the mean squared error and the mean absolute error are given by,

$$\mathrm{MSE}(x, y) = \frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2$$

$$\mathrm{MAE}(x, y) = \frac{1}{n} \sum_{i=1}^{n} |x_i - y_i|$$

A disadvantage of the MSE is that time shifts are not taken into account. For example, if $x_i = y_{i+1}$ holds, then the MSE would not detect that $x$ and $y$ are identical except for time shift. It is of course possible to try different time delays when using the MSE. Means the MSE is replaced by,

$$\text{MSE}_l(x, y) = \frac{1}{n - l} \sum_{i=1}^{n-l} (x_i - y_{i+l})^2, \tag{6.1}$$

where $l \in \mathbb{N}$. Then the $l$ for which (6.1) gets minimized can be used.

## Dynamic Time Warping

Instead, the method Dynamic Time Warping (DTW) can be used in such a case. DTW was introduced by Vintsyuk 1968 [Vin68]. It was used for speech recognition, but this approach is also applicable for time series in general. An advantage of DTW is that this algorithm is practical to compare two time series of different lengths and can be helpful for reel valued data and discrete data. In simple terms, a distance measure and possible steps are required in addition to the two time series. These steps define which points in time may be compared with each other. These steps are then applied to find the path where the distance becomes minimized. This distance is then the result of the algorithm. For the implementation, the R package *dtw* [Gio09] can be used, for example. However, since DTW is not applied in this thesis, it will not be discussed in detail in the following. More information on Dynamic Time Warping is discoverable in [TGQS08] and [MÖ7], among others.

## Cross-correlation

Another option is to determine the similarity by cross-correlation. Cross-correlation is identical to the basic correlation with the difference that an additional time delay is built-in see the definition 6.1.2. The advantage of this method is that the results are always between -1 and 1. On the other hand, the disadvantage is that only the linear relationship is tested and not the distance between two time series. For example, if $y_i = 100x_i$ for all $i = 1, \ldots, n$, then the result of the (cross-)correlation would be one. In this case, it is easy to establish the similarity between $y_i$ and $x_i$. However, the cross-correlation does not give any information about whether the data have to be normalized or not.

**Definition 6.1.2** (Cross-correlation and cross-covariance)**.** Let $x = (x_1, \ldots, x_n)' \in \mathbb{R}^n$ and
$y = (y_1, \ldots, y_n)' \in \mathbb{R}^n$ then the cross-covariance to time lag and $l \in \mathbb{Z}$ is given by

$$C_{xy}(l) = \sum_{i=1}^{n-l} (x_i - \bar{x})(y_{i+l} - \bar{y}) \quad \text{and}$$

$$C_{xy}(-l) = \sum_{i=1}^{n-l} (x_{i+l} - \bar{x})(y_i - \bar{y}).$$

This can be used to define the cross-correlation by

$$\hat{\rho}_{xy}(l) = \frac{C_{xy}(l)}{\sqrt{C_{xy}(0)C_{xy}(0)}}.$$

See [Mad07, Page 150].

## Further methods

Prof. Dr. Kriegel provided an overview of other methods for similarity search of time series see [Kri07]. These include Dynamic Time Warping, Whole-Matching, Subsequence Matching, Singular Value Decomposition, and many more.

## 6.2 Discussion

The disadvantage of data-based comparisons is that they cannot represent the interrelationships within a building. For example, assume that climate data is used for the simulation, which differs significantly from measured climate data. In this case, the difference in climate data already explains the difference between the simulated and the measured heating power. This means that data-based comparison is only helpful if there are identical conditions between simulation and measurement, such as similar climatic conditions and similar attendance and fault periods. In order to be able to compare simulation and measurement, even if the preconditions are not identical, the following chapter introduces model-based comparison.

# Chapter 7

# Model-based comparison

In model-based comparisons, the data are not compared with each other directly. Instead, separate models are trained on the simulation data and the measurement data. These models should describe the relationships within a building or any system. Subsequently, the coefficients estimated with the simulated data are compared with those estimated with the measured data. If these coefficients are similar, then the simulation data appears to be similar to the measured data. Overall the time points themselves are not relevant in model-based comparisons - much more important are the relationships within a system.

For example, the carbon dioxide content of a room depends on whether there are people in the room or not. If the room is long enough without users, then the carbon dioxide content is relatively constant. As long as a person is in the room, the carbon dioxide content increases, and when the person leaves the room, the carbon dioxide content should decrease. In such a case, a data-based comparison between the simulated and the measured carbon dioxide only makes sense if the periods for the presence are identical. In contrast, the model-based comparison first models the relationships between presence and carbon dioxide change in the room. Thereby, the model equations are identical for the simulation and the measurement data. After the estimation of the coefficients, they can be compared with each other. In this way, differences between simulation and measurement can be detected even if the presence periods differ.

Another example would be the heating capacity of a building. The lower the outside air temperature, the higher the required heating capacity. This means that a data-based comparison only makes sense if the temperature in the simulation and in the measurement period are identical. On the other hand, only the relationship between power and outside air temperature is essential with a model-based comparison.

This principle can also used to compare two stochastic processes with each other. Let for example be $y_t = \phi_y y_{t-1} + \epsilon_t$ and $x_t = \phi_x x_{t-1} + \tilde{\epsilon}_t$ where $(\epsilon_t)$ and $(\tilde{\epsilon}_t)$ are two white noise processes. In data-based comparison pairs $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ would be compared with each other. In model-based comparison, it would be checked whether $\phi_x \approx \phi_y$ holds.

## 7.1   A note to model-based comparisons

Let $S \in \mathbb{R}^{n \times p}$ be a design matrix created with simulation data and $M \in \mathbb{R}^{n \times p}$ a design matrix created with measurement data. In addition let $Y_S = (Y_{1S}, \ldots, Y_{nS})' \in \{0,1\}^n$ be the vector of the simulated binary target variables and $Y_M = (Y_{1M}, \ldots, Y_{nM})' \in \{0,1\}^n$ the vector of the measured target variables. The question is whether the relationship between $S$ and $Y_S$ is similar to the relationship between $M$ and $Y_M$. To analyze this, the two models $(\text{logit}(\pi_{1S}), \ldots, \text{logit}(\pi_{nS}))' = S\beta_S$ and $(\text{logit}(\pi_{1M}), \ldots, \text{logit}(\pi_{nM}))' = M\beta_M$ are created where $\pi_{tS} := P(Y_{tS} = 1 | S\beta_S) \in (0,1)$ and $\pi_{tM} := P(Y_{tM} = 1 | M\beta_M) \in (0,1)$. For this purpose, let $\hat{\beta}_S$ be the estimator of $\beta_S$ and $\hat{\beta}_M$ the estimator of $\beta_M$. If $\hat{\beta}_S$ and $\hat{\beta}_M$ are similar then the simulation of $Y_S$ seems to reflect the reality well. The problem is, however, that this approach is only possible if all information is present. Therefore, the question arises how to apply this method when $Y_M$ is not given.

One idea would be to estimate $Y_M$ with the model trained by simulation data. This means at first $\hat{\beta}_S$ is estimated. Then $(\text{logit}(\hat{\pi}_{1M}), \ldots, \text{logit}(\hat{\pi}_{nM}))' := M\hat{\beta}_S$ is calculated to estimate $\hat{Y}_M$. Where $\hat{Y}_M$ is one in components where $\hat{\pi}_{tM} \in (0,1)$ is bigger than some threshold value and zero otherwise. Now that an estimator for $Y_M$ is available, the remaining steps can be performed as before to compare $\hat{\beta}_S$ and $\hat{\beta}_M$. The question that remains is how useful this approach actually is.

There is a clear answer to this question, and it is that **this approach is not meaningful at all**. For one $\hat{\beta}_S$ is used to estimated $\hat{\beta}_M$ and $\hat{\beta}_S$ depends on the design matrix $S$. Therefore $\hat{\beta}_M$ also depends on $S$, which means that this approach does not check how similar $S$ and $M$ are. That this procedure is not useful can be shown especially within the linear regression. For this, $Y_S \in \mathbb{R}^n$ and $Y_M \in \mathbb{R}^n$ are redefined for a regression problem. The equations (7.1) up to (7.4) proof that $\hat{\beta}_M = \hat{\beta}_S$ if this procedure is done in a linear regression model. Where (7.1) shows the model equation and the estimation of $\beta_S$ using the simulation. In the next step (7.2) $\hat{\beta}_S$ is used to estimate $\hat{Y}_M$. In equation (7.3) $\hat{Y}_M$ is used to build a model and estimate $\beta_M$ using the measured data. The equation (7.4) is identical to the equation (7.3) but $\hat{Y}_M$ is replaced by the $\hat{Y}_M$ from the equation (7.2). Then $(M^T M)^{-1}$ and $(M^T M)$ cancel each other out, which means that $\hat{\beta}_M$ does not depend on the measured data.

$$Y_S = S\beta_S \Rightarrow \hat{\beta}_S = (S^T S)^{-1} S^T Y_S \tag{7.1}$$

$$\hat{Y}_M = M\hat{\beta}_S \tag{7.2}$$

$$\hat{Y}_M = M\beta_M \Rightarrow \hat{\beta}_M = (M^T M)^{-1} M^T \hat{Y}_M \tag{7.3}$$

$$\Rightarrow \hat{\beta}_M \stackrel{(7.2)\underline{)(7.3)}}{=} (M^T M)^{-1} M^T M\hat{\beta}_S = \hat{\beta}_S \quad \blacksquare \tag{7.4}$$

In summary, this shows that mixing simulation and measured data or models should be avoided in model-based comparisons. If it is nevertheless considered necessary, then the extent to which the results influence each other should be examined.

## 7.2   Comparison by standard error

Most statistical models require coefficients. These coefficients will be estimated by using data. For this purpose, let $\hat{\beta}_S \in \mathbb{R}$ be an estimated coefficient, which was estimated by simulation data. While $\hat{\beta}_M \in \mathbb{R}$ represents the estimation using real measurement data. It is almost impossible that $\hat{\beta}_S = \hat{\beta}_M$ holds even if the simulation is nearly perfect. For this reason, it is necessary to replace the point estimators $\hat{\beta}_S$ and $\hat{\beta}_S$ with interval estimators. Therefore let $\widehat{SE}_S := \sqrt{\mathrm{Var}(\hat{\beta}_S)}$ be the standard error of $\hat{\beta}_S$ and $\widehat{SE}_M := \sqrt{\mathrm{Var}(\hat{\beta}_M)}$ the standard error of $\hat{\beta}_M$. These standard errors can now be used to construct interval estimators by $I_S := (\hat{\beta}_S - \widehat{SE}_S, \hat{\beta}_S + \widehat{SE}_S)$ and $I_M := (\hat{\beta}_M - \widehat{SE}_M, \hat{\beta}_M + \widehat{SE}_M)$. The coefficients $\hat{\beta}_S$ and $\hat{\beta}_M$ can now be considered similar if $I_S \cap I_M \neq \emptyset$. All steps are summarized in algorithm 7.2.1.

---

**Input:** $\mathrm{Data}_S$, $\mathrm{Data}_M$, Model-Formula
**Input:** R
1  Estimate $\hat{\beta}_S$ with $\mathrm{Data}_S$ and the Model-Formula
2  Estimate $\hat{\beta}_M$ with $\mathrm{Data}_M$ and the Model-Formula
3  $\widehat{SE}_S \leftarrow \sqrt{\mathrm{Var}(\hat{\beta}_S)}$
4  $\widehat{SE}_M \leftarrow \sqrt{\mathrm{Var}(\hat{\beta}_M)}$
5  **if** $(\hat{\beta}_S - \widehat{SE}_S, \hat{\beta}_S + \widehat{SE}_S) \cap (\hat{\beta}_M - \widehat{SE}_M, \hat{\beta}_M + \widehat{SE}_M) \neq \emptyset$ **then**
6  $\quad\mid\quad R \leftarrow True$
7  **else**
8  $\quad\mid\quad R \leftarrow False$
9  **end**

**Algorithm 7.2.1:** Comparison with standard error

---

This method can also be extended by replacing $\widehat{SE}_S$ and $\widehat{SE}_M$ with $\lambda \cdot \widehat{SE}_S$ and $\lambda \cdot \widehat{SE}_M$ where $\lambda \in \mathbb{R}_{\geq 0}$. Through this, the strictness for the similarity can be determined. In the one extreme case, $\lambda = 0$, the coefficients are similar only if they are identical. On the other hand, for $\lambda \to \infty$, all coefficients are similar.

# 7.3 Parametric bootstrapping

Another way to generate intervals for the coefficients of a model is by bootstrapping. The idea of bootstrapping can be found, for example, in Efron 1994 [ET94] or in Hall 2013 [Hal13]. Let $x = (x_1, \ldots, x_n)'$ be a random sample with $x \sim F$. The approach is to draw $B$ times new samples of length $n$ where each should be distributed like $F$. Then, a statistic $T$ can be applied to each of these samples. This gives $B$ suggestions for the statistic $T$ on data that are distributed like $F$. These suggestions can now be used to generate standard deviations or intervals for $T(x)$. The difficulty is that the true distribution $F$ is mainly unknown in practice. For this reason, there are two main approaches to implement this procedure. These are the parametric and the non-parametric bootstrapping. This section describes a model-based comparison method using a variant of parametric bootstrapping. A non-parametric variant is presented in the section 7.4.

Therefore let $Y = X\beta + \nu$ be a regression model. Where $\beta = (\beta_0, \beta_1, \ldots, \beta_p)' \in \mathbb{R}^{p+1}$, $Y$ is the response vector, $X$ is a $n \times (p+1)$ design matrix and $\nu \sim F(\vartheta)$ is the error term. The distribution of that error $F$ depends on a parameter $\vartheta \in \Theta \subseteq \mathbb{R}^k$ where $k \in \mathbb{N}$. In a linear regression model, $F$ follows a multivariate normal distribution where all components are independent in a dynamic regression model $F$ is a ARMA process.

After the parameters of the model have been estimated using training data, $\hat{\beta}$ and $\hat{\vartheta}$ can be used to perform parametric bootstrapping. For this reason, let $\hat{Y} := X\hat{\beta}$ be an estimation of $Y$. Now new noises $\nu^*$ can be drawn from $F(\hat{\vartheta})$ to add up to $\hat{Y}$. This generates a new response vector $Y^* := \hat{Y} + \nu^*$. Now the coefficients $\beta$ can be estimated again, but $Y$ from the train data will be replaced by $Y^*$. This process can be repeated many times, and after each pass, a new estimation for $\beta$ and $\vartheta$ is obtained. These estimations can be used to calculate intervals for each $\beta_j$ and each component in $\vartheta$ where $j = 0, 1, \ldots, p$. The algorithm 7.3.1 shows all steps the parametric bootstrapping steps for regression models. In this algorithm $\hat{\beta}(Y, X)$ and $\hat{\vartheta}(Y, X)$ denotes the estimation function for $\beta$ and $\vartheta$ using $Y$ and $X$ as train data.

There are several ways to create intervals using bootstrapping. One way lead to the percentile intervals suggested in Efron 1994 [ET94, Page 170]. Therefore let $\hat{\beta}_j^* := (\hat{\beta}_{j1}^*, \ldots, \hat{\beta}_{jB}^*)'$ be the vector of the bootstrap estimations for $\beta_j$ after $B$ iterations. Also let $\hat{\beta}_j^{*(\alpha)}$ be the $100 \cdot \alpha$th percentile of $\hat{\beta}_j^*$. Then the $1 - \alpha$ percentile interval can be written as $[\hat{\beta}_j^{*(\frac{\alpha}{2})}; \hat{\beta}_j^{*(1-\frac{\alpha}{2})}]$.

Obviously, this procedure can be done for every $\beta_j$ and each component in $\vartheta$.

These bootstrap intervals must be performed once with the simulation data and once with the real data. Then, as in section 7.2, these intervals can be compared with each other. If the intersection between the interval generated with simulation data and the interval generated with real data is equal to the empty set, then the influence to $Y$ differs. Suppose the intersection is not equal to the empty set. In that case, the influence of that exogenous variable on the response variable $Y$ can be assumed to be similar in the simulation and the real data.

```
    Input: Y ∈ ℝⁿ, X ∈ ℝⁿ×⁽ᵖ⁺¹⁾, B ∈ ℕ
    Output: (β̂*, ϑ̂*)
  1  β̂ ← β̂(Y, X)
  2  ϑ̂ ← ϑ̂(Y, X)
  3  Ŷ ← Xβ̂
  4  for b = 1, ..., B do
  5  │   Sample ν*_b ∈ ℝⁿ from F(ϑ̂)
  6  │   Y*_b ← Ŷ + ν*_b
  7  │   β̂*_b ← β̂(Y*_b, X)
  8  │   ϑ̂*_b ← ϑ̂(Y*_b, X)
  9  end
 10  β̂* ← (β̂*_1, ..., β̂*_B)'
 11  ϑ̂* ← (ϑ̂*_1, ..., ϑ̂*_B)'
```

**Algorithm 7.3.1:** Parametric bootstrapping for a regression model

# 7.4 Non-parametric bootstrapping

Besides parametric bootstrapping for regression there also exists also a non-parametric version. This time the focus will be directly on dynamic regression models. Let $Y_t = X_t\beta + \nu_t$ be a dynamic regression model where $Y_t \in \mathbb{R}$, $X_t$ is the $t$-th row of a $n \times (p+1)$ design matrix, $\beta \in \mathbb{R}^{p+1}$ and the noise $\nu_t = \sum_{i=1}^{p} \phi_i \nu_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t$ follows an ARMA(p,q) process. Non-parametric bootstrapping is supposed to draw from the data. In the case of a linear regression model, which means that $\phi_1 = \cdots = \phi_p = \theta_1 = \cdots = \theta_q = 0$ holds, it is possible to randomly draw pairs $(Y_i, X_i)$ or draw residuals $\hat{\nu}_i := Y_i - X_i\hat{\beta}$, see [ET94, Page 113]. In [ET94, Page 113] it is also suggested to bootstrap from the pairs because it is less sensitive to the model assumptions then bootstrapping the residuals. But in a dynamic regression model the noise to a time point $t$ may depend on the past noise and this structure would be ignored by bootstrapping the pairs $(Y_i, X_i)$. For this reason a method for bootstrapping from a AR(1) process is suggested in Effron 1994 [ET94, Pages 95-99]. This proposal is in principle drawing from the residuals and is expandable for dynamic regression models.

Overall this means that at first the parameters $\hat{\beta}$, $\hat{\phi} = (\hat{\phi}_1, \ldots, \hat{\phi}_p)$ and $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_q)$ will be estimated. Afterwards $\hat{Y} := X\hat{\beta}$ and the residuals $\hat{\nu} = (\hat{\nu}_1, \ldots, \hat{\nu}_n)' = Y_i - X_i\hat{\beta}$ can be calculated. Since the residuals depend on each other it is not really appropriate to draw directly from the residuals. Instead independent residuals will be calculated by $\hat{\epsilon}_t = \hat{\nu}_t - \sum_{i=1}^{\min(t-1,p)} \hat{\phi}_i \hat{\nu}_{t-i} - \sum_{j=1}^{\min(t-1,q)} \hat{\theta}_j \hat{\epsilon}_{t-j}$ for $t = 1, \ldots, n$. Note that $\min(t-1, p)$ and $\min(t-1, q)$ in the end of the sums are only relevant for the time points for which $t + 1 < \max(p, q)$ holds for the time points after that $\min(t-1, p) = p$ and $\min(t-1, q) = q$ is true. Now it is possible to draw $n$ times from these independent residuals $\hat{\epsilon}_t$ with replacement to get new bootstrapped independent residuals $\hat{\epsilon}^* = (\hat{\epsilon}^*_1, \ldots, \hat{\epsilon}^*_n)'$. These can be used to

calculated bootstrapped ARMA errors by $\hat{\nu}_t^* = \sum_{i=1}^{\min(t-1,p)} \hat{\phi}_i \hat{\nu}_{t-i}^* + \sum_{j=1}^{\min(t-1,q)} \hat{\theta}_j \hat{\epsilon}_{t-j}^* + \hat{\epsilon}_t^*$. The remaining steps are similar to the non-parametric version. The bootstrapped response vector $Y^* := \hat{Y} + \hat{\nu}^*$ will be calculated where $(\hat{\nu}_1^*, \ldots, \hat{\nu}_n^*)'$. Concluding the estimations of $\beta$, $\phi$ and $\theta$ can be repeated but with $Y^*$ instead of $Y$. These steps can be repeated as often as desired to subsequently generate intervals for the coefficients, which can be used for model-based comparisons. Algorithm 7.4.1 shows all bootstrapping steps where $\hat{\beta}(Y, X)$, $\hat{\phi}(Y, X)$ and $\hat{\theta}(Y, X)$ denotes the estimation functions for $\beta$, $\phi$ and $\theta$.

---

**Input:** $Y \in \mathbb{R}^n$, $X \in \mathbb{R}^{n \times (p+1)}$, $B \in \mathbb{N}$
**Output:** $(\hat{\beta}^*, \hat{\phi}^*, \hat{\theta}^*)$

1   $\hat{\beta} \leftarrow \hat{\beta}(Y, X)$
2   $(\hat{\phi}_1, \ldots, \hat{\phi}_p) \leftarrow \hat{\phi}(Y, X)$
3   $(\hat{\theta}_1, \ldots, \hat{\theta}_q) \leftarrow \hat{\theta}(Y, X)$
4   $\hat{Y} \leftarrow X\hat{\beta}$
5   $(\hat{\nu}_1, \ldots, \hat{\nu}_n) \leftarrow Y - X\hat{\beta}$
6   **for** $t = 1, \ldots, n$ **do**
7      $\hat{\epsilon}_t \leftarrow \hat{\nu}_t - \sum_{i=1}^{\min(t-1,p)} \hat{\phi}_i \hat{\nu}_{t-i} - \sum_{j=1}^{\min(t-1,q)} \hat{\theta}_j \hat{\epsilon}_{t-j}$
8   **end**
9   $\hat{\epsilon} \leftarrow (\hat{\epsilon}_1, \ldots, \hat{\epsilon}_n)$
10   **for** $b = 1, \ldots, B$ **do**
11      Draw $n$ times with replacement from $\hat{\epsilon}$ to get $(\hat{\epsilon}_{1b}^*, \ldots, \hat{\epsilon}_{nb}^*)$
12      **for** $t = 1, \ldots, n$ **do**
13          $\hat{\nu}_{tb}^* \leftarrow \sum_{i=1}^{\min(t-1,p)} \hat{\phi}_i \hat{\nu}_{t-i,b}^* + \sum_{j=1}^{\min(t-1,q)} \hat{\theta}_j \hat{\epsilon}_{t-j,b}^* + \hat{\epsilon}_{t,b}^*$
14      **end**
15      $\hat{\nu}_b^* \leftarrow (\hat{\nu}_{1b}^*, \ldots, \hat{\nu}_{nb}^*)$
16      $Y_b^* \leftarrow \hat{Y} + \hat{\nu}_b^*$
17      $\hat{\beta}_b^* \leftarrow \hat{\beta}(Y_b^*, X)$
18      $\hat{\phi}_b^* = (\hat{\phi}_{1b}^*, \ldots, \hat{\phi}_{nb}^*) \leftarrow \hat{\phi}(Y_b^*, X)$
19      $\hat{\theta}_b^* = (\hat{\theta}_{1b}^*, \ldots, \hat{\theta}_{nb}^*) \leftarrow \hat{\theta}(Y_b^*, X)$
20   **end**
21   $\hat{\beta}^* \leftarrow (\hat{\beta}_1^*, \ldots, \hat{\beta}_B^*)$
22   $\hat{\phi}^* \leftarrow (\hat{\phi}_1^*, \ldots, \hat{\phi}_B^*)$
23   $\hat{\theta}^* \leftarrow (\hat{\theta}_1^*, \ldots, \hat{\theta}_B^*)$

**Algorithm 7.4.1:** Non-parametric bootstrapping for a dynamic regression model

---

## 7.5   Posterior cut area comparison

In statistics, there are two main philosophical approaches [BC16, Page 5]. One of them is often called the frequentest approach. In this approach, parameters are assumed to be fixed,

but unknown constants and probabilities are interpreted as relative frequency [BC16, Pages 5-6]. The other approach is the Bayesian approach, in which random variables express the uncertainty about the true parameter. The probabilities in the Bayesian statistics are interpreted as a degree of belief [BC16, Pages 6-7]. These different interpretation concepts lead to different methodological approaches. Usually, a Bayesian approach starts with a subjective prior distribution of the parameter, which is to be estimated. In the second step, data is used to update the belief by using Bayes' theorem to obtain a posterior distribution for that parameter. The posterior distribution can now be used for the analysis. In the case of model-based comparison, the simulation data and the real data can be used to derive a posterior distribution. These posterior distributions can now be compared with each other.

This leads to the question of how to compare the distributions. One way is by using the density function of the distribution. Therefore let $g_1(x)$ and $g_2(x)$ be two continuous univariate density functions. Each density function is non-negative everywhere, and it holds that the integral is equal to one [Sti99, Page 170]. Therefore

$$\int_{\mathbb{R}} \min(g_1(x), g_2(x))dx \tag{7.5}$$

is a value between zero and one. If $g_1(x) = g_2(x)$ for each $x \in \mathbb{R}$ then the integral in (7.5) is equal to one. If however $g_1(x) > 0 \Leftrightarrow g_2(x) = 0$ then (7.5) is equal to zero. Graphically, the integral in (7.5) is the cut area between two density functions. If the densities are replaced by the posterior distributions obtained from the simulation and the measured data, then the posterior cut area from definition 7.5.1 is obtained. If the posterior cut area is bigger than some threshold between zero and one, then the posteriors can be assumed to be similar.

---

**Definition 7.5.1** (Posterior cut area)**.** Let $p(\vartheta)$ be a univariate and continuous prior distribution, $p(\vartheta|S)$ the posterior distribution given by the simulation data and $p(\vartheta|M)$ the posterior given by the measured data. Then the posterior cut area is given by

$$\int_{\mathbb{R}} \min(p(\vartheta|S), p(\vartheta|M))d\vartheta.$$

---

The problem, however, is the fact that the true posteriors are analytically mostly hard to derive. For that reason, numerical Markov Chain Monte Carlo Methods, or short MCMC Methods, are used to sample from these posteriors. Examples of these methods are the EM algorithm, Gibbs-sampler, and the Metropolis-Hastings algorithm. Explanations of these methods can be found in Lee 2012 [Lee12, Ch. 9]. The exact method is irrelevant for the further procedure. It is only important that (pseudo-)random numbers can be drawn from the desired posteriors. These samples will now be used to estimate densities.

Therefore, a kernel function is needed. Examples of kernel functions are trackable in Table 7.1. These are usable to define the kernel density estimator in Definition 7.5.2.

| Name | $K(u)$ |
|---|---|
| Uniform | $\frac{1}{2}I(|u| < 1)$ |
| Triangle | $(1 - |u|)I(|u| < 1)$ |
| Epanechnikov | $\frac{3}{4}(1 - u^2)I(|u| < 1)$ |
| Gaussian | $\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{u^2}{2}\right)$ |

Table 7.1: Kernel functions
See [HMS06].

---

**Definition 7.5.2** (Kernel density estimator). Let $x_1, x_2, \ldots, x_n$ be independent samples from the same continuous distribution and let $K$ be a kernel function. Then

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

is a kernel denstiy estimator to the bandwidth $h$ [Par62].

---

Now by replacing the posteriors in definition 7.5.1 with the kernel density estimator a estimator for the posterior cut area in definition 7.5.3 can be derived.

---

**Definition 7.5.3** (Estimated posterior cut area). Let $\hat{f}_S(\vartheta)$ and $\hat{f}_M(\vartheta)$ be arbitrary kernel density estimator generated with samples drawn by a posterior distribution $p(\vartheta|S)$ and $p(\vartheta|M)$ where $S$ are simulation data and $M$ are measurement data. Then the estimated posterior cut area is given by

$$\int_{\mathbb{R}} \min(\hat{f}_S(\vartheta), \hat{f}_M(\vartheta|M))d\vartheta.$$

---

All steps for the posterior cut area comparison are summarized in the algorithm 7.5.1.

A major drawback to this approach is that it can cause numerical problems. Since first numerically samples are drawn from posterior densities to subsequently generate the kernel density estimators to numerically integrate over them. These are three possible sources of potential numerical problems that build on each other. In practice, it has been shown that especially kernel density estimation can lead to problems. Especially if the standard deviation of the posterior samples is too small, for this reason, a Bayesian method is proposed in the following section 7.6, which is numerically more stable.

---

**Input:** $S$,$M$, $p(\vartheta)$, $\alpha \in [0,1)$, $I \in \mathbb{N}$
**Output:** R
**1** Sample $\vartheta_1|S, \ldots, \vartheta_I|S$ from $p(\vartheta|S)$ via MCMC
**2** Sample $\vartheta_1|M, \ldots, \vartheta_I|M$ from $p(\vartheta|M)$ via MCMC
**3** Calculate the kernel density estimator $\hat{f}_S(\vartheta)$ with $\vartheta_1|S, \ldots, \vartheta_I|S$
**4** Calculate the kernel density estimator $\hat{f}_M(\vartheta)$ with $\vartheta_1|M, \ldots, \vartheta_I|M$
**5** $C \leftarrow \int_{\mathbb{R}} \min(\hat{f}_S(\vartheta), \hat{f}_M(\vartheta|M))d\vartheta$
**6 if** $C > \alpha$ **then**
**7**  $\mid$  $R \leftarrow True$
**8 else**
**9**  $\mid$  $R \leftarrow False$
**10 end**

**Algorithm 7.5.1:** Posterior cut area comparison

---

## 7.6   Credible interval comparison

In addition to the method presented in section 7.5, posteriors can also be compared using credibility intervals. These intervals are defined in definition 7.6.1.

---

**Definition 7.6.1** (Credible Interval)**.** Let $p(\vartheta|x)$ be a posterior distribution. Then for a fixed $\alpha \in (0,1)$, a $\alpha \cdot 100\%$ credible Interval is defined through two real numbers $L$ and $U$, that fulfill

$$\int_L^U p(\vartheta|x)d\vartheta = \alpha.$$

The quantity $\alpha$ is called the credible level of the credible interval $[L, U]$. See also [HB13, Page 172].

---

There are several ways to set the lower and upper bound for the confidence intervals. A simple way to specify $L$ and $U$ would be via quantiles. If the $\frac{\alpha}{2}$-quantile is used for $L$ and the $(1 - \frac{\alpha}{2})$-quantile of posterior densities for $U$, it gives a possible credibility interval to the credibility level $\alpha$. These quantiles can now be used to compare posterior densities of the simulation and the measured data. Therefore two credibility intervals will be created. One of these is generated with the simulation data posterior and the other one with measured data posterior. For the lower and upper bound, the $\frac{\alpha}{2}$- and the $1 - \frac{\alpha}{2}$-quantile is used in each case. If the intersection of these two intervals is not equal to the empty set, then the parameter can be assumed to be similar. See also definition 7.6.2.

---

**Definition 7.6.2** (Credible interval comparison)**.** Let $q_\alpha(\vartheta|S)$ be a $\alpha$-quantile of the posterior distribution $p(\vartheta|S)$ and $q_\alpha(\vartheta|M)$ be a $\alpha$-quantile of the posterior distribution

$p(\vartheta|M)$. Where $p(\vartheta)$ is a univariate prior distribution for $\vartheta$. Then $p(\vartheta|S)$ and $p(\vartheta|M)$ are similar to the credible level $\alpha \in (0,1)$ if

$$\left(q_{\frac{\alpha}{2}}(\vartheta|S), q_{1-\frac{\alpha}{2}}(\vartheta|S)\right) \cap \left(q_{\frac{\alpha}{2}}(\vartheta|M), q_{1-\frac{\alpha}{2}}(\vartheta|M)\right) \neq \emptyset.$$

As described in section 7.5, it is mainly necessary to use MCMC algorithms to sample from the posteriors. In this case, the quantiles can be replaced by the empirical quantiles. From a numerical point of view, empirical quantiles are not problematic. That is the primary advantage of this method compared to the method described in section 7.5. All steps of this method are summarized in algorithm 7.6.1.

---

**Input:** $S$,$M$, $p(\vartheta)$, $\alpha \in (0,1)$, $I \in \mathbb{N}$
**Output:** R
1 Sample $\vartheta_1|S, \ldots, \vartheta_I|S$ from $p(\vartheta|S)$ via MCMC
2 Sample $\vartheta_1|M, \ldots, \vartheta_I|M$ from $p(\vartheta|M)$ via MCMC
3 Define $q_{\frac{\alpha}{2}}(\vartheta|S)$ and $q_{1-\frac{\alpha}{2}}(\vartheta|S)$ as the empirical $\frac{\alpha}{2}$- and $(1-\frac{\alpha}{2})$-quantile of the
    samples $\vartheta_1|S, \ldots, \vartheta_I|S$
4 Define $q_{\frac{\alpha}{2}}(\vartheta|M)$ and $q_{1-\frac{\alpha}{2}}(\vartheta|M)$ as the empirical $\frac{\alpha}{2}$- and $(1-\frac{\alpha}{2})$-quantile of the
    samples $\vartheta_1|M, \ldots, \vartheta_I|M$
5 $C \leftarrow (q_{\frac{\alpha}{2}}(\vartheta|S), q_{1-\frac{\alpha}{2}}(\vartheta|S)) \cap (q_{\frac{\alpha}{2}}(\vartheta|M), q_{1-\frac{\alpha}{2}}(\vartheta|M))$
6 **if** $C \neq \emptyset$ **then**
7   |  $R \leftarrow True$
8 **else**
9   |  $R \leftarrow False$
10 **end**

**Algorithm 7.6.1:** Credible Interval comparison

## 7.7 Discussion

Note that neither concept is superior when comparing model-based and data-based comparisons. In model-based comparison, information about individual points in time is lost, which can be an advantage or disadvantage, depending on the context. However, the main problem with model-based comparison is the modeling itself. If important influencing variables are not taken into account, the result will not be meaningful. Especially with time series, it can be unclear how many hours of time lags should be considered to get a conclusive impact. Furthermore, it is unclear which covariates must match for global similarity to occur. However, there is the possibility to derive global metrics like the number of matches or a weighting of them, for example. Nevertheless, it is not clear how meaningful weights can be derived.

# Chapter 8

# Check unmeasured climate data

This chapter is an insertion and shows that building simulations and measured data are helpful to evaluate missing climate data. Here it is assumed that the simulation is suitable to represent the building. Now it is investigated which radiation data best fit the measurement.

## 8.1   Estimate Radiations

Assuming only the global radiation was measured, but for the simulation, diffuse and direct radiations are needed. Then the global radiation can be used to estimate the required radiations. These estimates are derivable by using the radiations from another location. Such radiation data can be found, for example, at the homepage of Deutscher Wetterdienst (DWD) [Deu]. The Rosenheim radiation data is not available on their site. Therefore the data from the nearest location with available radiation data is used. This location is Weihenstepan. The measurements available at the DWD are the global radiation, the diffuse radiation, and the zenith angle in degrees. The following goal is to estimate the diffuse and direct radiation from Rosenheim with the given Weihenstephan values and the global radiation from Rosenheim.

In the first step, it is necessary to adjust the units. The Rosenheimer radiation was measured in Watt-hour per square meter. In contrast, the DWD uses the hourly sum in Joule per square centimeter. For this reason, the DWD radiations are multiplied with the value $\frac{25}{9}$. Equation (8.1) shows the conversion of the units.

$$\frac{Wh}{m^2} = \frac{J}{s}\frac{h}{10000\text{cm}^2} = \frac{J}{\frac{h}{3600}}\frac{h}{10000\text{cm}^2} = \frac{36J}{100\text{cm}^2} = \frac{9J}{25\text{cm}^2} \tag{8.1}$$

Figure 8.1 shows the comparison between the Rosenheimer global radiation and the Weihenstephan global radiation.

In the case of the Weihenstepan radiation, the division into diffuse and direct is known. The division for the Rosenheimer radiation is unknown but needed. For this reason, three
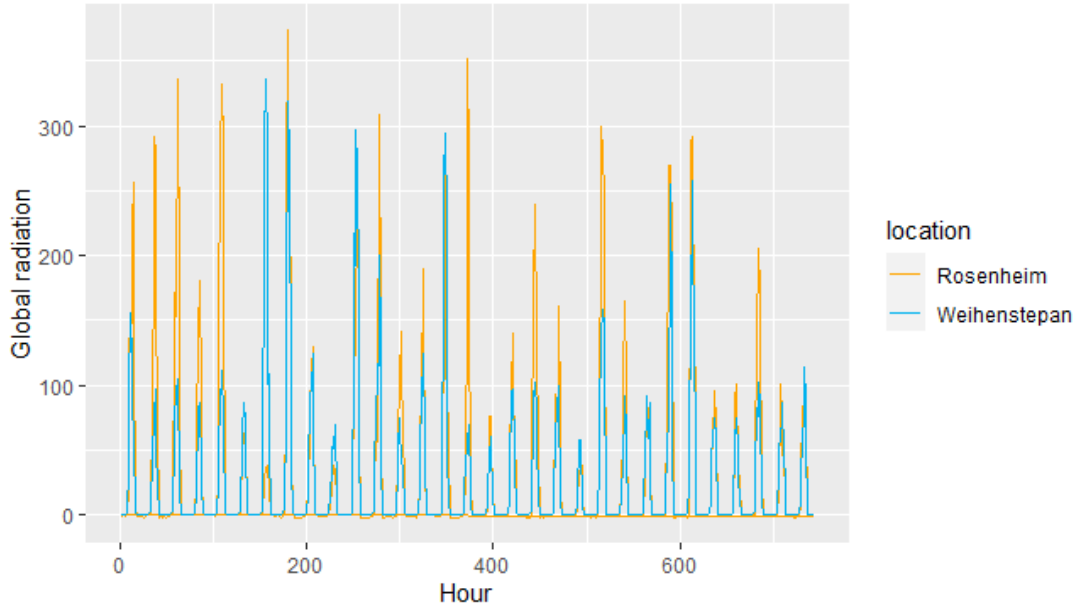
Figure 8.1: Comparison of global radiation

approaches are presented for splitting the Rosenheimer radiation using the Weihenstepan radiations. Therefore the radiations are defined by in definition 8.1.1.

---

**Definition 8.1.1** (Diffuse, direct and global radiation). Let $\text{difR}_{l,t}$ be the diffuse radiation, $\text{dirR}_{l,t}$ the direct radiation and $\text{gloR}_{l,t}$ the gloabal radiation, each to the hour $t$ at the location $l$.

Where $t \in \{1, \ldots, 744\}$ and $l \in \{\text{Ro}, \text{Ws}\} = \{\text{Rosenheim}, \text{Weihenstephan}\}$

---

Also the solar zenith angle at mid of interval are needed.

---

**Definition 8.1.2** (Zenith angle). Let $\text{ZW}_{\text{Ws},t} \in [0; 180]$ be the zenith angle at mid of interval of Weihenstepan to the time $t$. Where $t \in \{1, \ldots, 744\}$.

This value is related to the solar height by $\text{Zenith} = 90° - \text{solar height}$ [Deu].

---

Now the tools to estimate the radiations in Rosenheim are defined. For the first estimator, the Weihenstephan values are used for Rosenheim. For this reason, this estimator is called the Weihenstephan-Estimator, see definition 8.1.3.

**Definition 8.1.3** (Weihenstepan-Estimation)**.** Let the terms from definition 8.1.1 be given and let sin be the sine function in angular measure. Then the Weihenstephan-Estimator is given by,

- $\widehat{\text{difR}}_{\text{Ro},t} = \text{difR}_{\text{Ws},t}$,

- $\widehat{\text{dirR}}_{\text{Ro},t} = \frac{\text{gloR}_{\text{Ws},t} - \text{difR}_{\text{Ws},t}}{\sin(90° - \text{ZW}_{\text{Ws},t})}$.

For the second estimator, the main idea is to split the Rosenheimer global radiation like the Weihenstepan radiation for each hour. This means that the relative proportion of diffuse- and direct radiation from Weihenstephan is calculated for each point in time. This proportion is then used for the Rosenheimer radiations.

**Definition 8.1.4** (Time-Proportional-Estimation)**.** Let the terms from definition 8.1.1 be given and let sin be the sine function in angular measure. Then the Time-Proportional-Estimation is given by,

- $\widehat{\text{difR}}_{\text{Ro},t} = \left( \frac{\text{difR}_{\text{Ws},t}}{\text{gloR}_{\text{Ws},t}} \right) \cdot \text{gloR}_{\text{Ro},t}$,

- $\widehat{\text{dirR}}_{\text{Ro},t} = \left( 1 - \frac{\text{difR}_{\text{Ws},t}}{\text{gloR}_{\text{Ws},t}} \right) \cdot \frac{\text{gloR}_{\text{Ro},t}}{\sin(90° - \text{ZW}_{\text{Ws},t})}$.

For each $t$ where $\text{gloR}_{\text{Ro},t} > 0$. For the hours $t$ where $\text{gloR}_{\text{Ro},t} = 0$ the medians of $\frac{\text{difR}_{\text{Ws},t}}{\text{gloR}_{\text{Ws},t}}$ during the remaining hours is used instead.

One disadvantage of this estimator is that if the global radiation differs significantly at one point, it is probably not meant to use the same proportion. For this reason, the following estimator sorts both global radiations from smallest to biggest. Then for each position, the proportion of Weihenstephan is used for Rosenheim. Then finally, the values are sorted again according to time.

> **Definition 8.1.5** (Rank-Proportional-Estimation)**.** Let the terms from definition 8.1.1 be given and let sin be the sine function in angular measure. Then the Rank-Proportional-Estimation is given by,
>
> - $\widehat{\text{difR}}_{\text{Ro},r_t} = \left( \frac{\text{difR}_{\text{Ws},w_t}}{\text{gloR}_{\text{Ws},w_t}} \right) \cdot \text{gloR}_{\text{Ro},r_t}$
>
> - $\widehat{\text{dirR}}_{\text{Ro},r_t} = \left( 1 - \frac{\text{difR}_{\text{Ws},w_t}}{\text{gloR}_{\text{Ws},w_t}} \right) \cdot \frac{\text{gloR}_{\text{Ro},r_t}}{\sin(90° - \text{ZW}_{\text{Ws},r_t})}$
>
> Where $r_1, \ldots, r_{744}$ and $w_1, \ldots, w_{744}$ are time points for which $\text{gloR}_{\text{Ro},r_1} \leq \cdots \leq \text{gloR}_{\text{Ro},r_{744}}$ and $\text{gloR}_{\text{Ws},w_1} \leq \cdots \leq \text{gloR}_{\text{Ws},w_{744}}$. For hours where $\text{gloR}_{\text{Ws},w_t} = 0$ the median of the remaining hours, the median of the remaining proportion is used.

## 8.2 Evaluation

The simulated room temperatures are compared with the measured one to evaluate the estimations. Since the fault periods in the measurement and simulations are identical, the mean squared error (MSE) from section 6.1 is used as the measure. For a better comparison, the simulation with the original climate data is also included in the evaluation. The results are presented in the table 8.1, showing both the MSE between the room temperature and the MSE between the actual and estimated radiations. It can be seen that the better the radiations were estimated, the better the room temperature was simulated.

| Simulation | MSE-RT | MSE-Diffuse | MSE-Direct |
|---|---|---|---|
| Weihenstepan | 1.32 | 474.29 | 16114.89 |
| Time-prop | 1.04 | 969.25 | 12133.41 |
| Rank-prop | 0.95 | 451.38 | 7468.55 |
| Original | 0.62 | 0 | 0 |

Table 8.1: Evaluation of the estimators for the radiations by MSE

# Chapter 9

# Comparison between simulation and real data

In this chapter, the simulation data and the real measured data are compared. Six simulations are tested. In one simulation, the fault periods and the climate data are identical to those in the measured data. In the other five simulations, the periods of the faults are shifted. In two of these, external climate data was also used.

Because of this, model-based comparison is used here since the shift of the fault periods should not play a decisive role. To be more precise, the four methods *comparison by standard error* (section 7.2), *parametric bootstrapping* (section 7.3), *non parametric bootstrapping* (section 7.4) and *credible interval comparison* (section 7.6) are applied. The method *posterior cut area comparison* was omitted due to numerical problems, which were also described in section 7.5. In each case, the change in the target variable at the time before is modeled. This is because it makes it easier to justify stationarity. The "d" at the beginning of the model designations stands for difference.

In addition to the data, further inputs are required for these methods - these are described in the following. The two bootstrap algorithms 7.3.1 and 7.4.1 need the number of bootstrap iterations $B$. These are set to 1000 in each run. For the comparison, a $\alpha$ of 0.05 was selected. The same $\alpha$ is also chosen for the credible interval comparison from algorithm 7.6.1. For the number of draws from the a-posterior density, $I = 10000$ is used. For the prior densities, distributions with high variance are chosen to express low prior knowledge. More precisely, the following a priori distributions are chosen: $\phi \sim \mathcal{N}(0, 1000)$, $\sigma^2 \sim \text{InvGamma}(\frac{1}{100}, \frac{1}{100})$ and $\beta_i \sim \mathcal{N}(0, 1000)$ for each $i = 0, 1, \ldots, p$.

The results are presented in tables. The four columns *St.E*, *P.BS*, *BS* and *Bayes* represent the four methods used for the comparison. The rows, on the other hand, represent the parameters. Each cell is either *TRUE* if there is a similarity between parameters by corresponding method or *FALSE* if not. The column *Sum* indicates how many methods assume a similarity between simulation and measurement.

In this chapter, periods with faults have been removed if they affect the corresponding target variable. For the change in carbon dioxide, two weeks without the ventilation fault were used in each data set. In contrast, the bypass and heating fault periods were removed

for the room temperature and heating power. This means that a comparison between simulation and measurement in the fault-free state is investigated because it is closer to practice.

Note that if several simulations provide identical results tables, only one of them is shown for the sake of clarity.

## 9.1 Carbon dioxide

> **Definition 9.1.1** (dCO2-Model). The dCO2-Model is given by,
>
> $$CO2_t - CO2_{t-1} = \beta_0 + \beta_1 Pre_t + \beta_2 Pre_{t-1} \beta_3 Pre_{t-2} + \nu_t$$
> $$\nu_t = \phi \nu_{t-1} + \epsilon_t,$$
>
> where $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$, $\sigma^2 \in \mathbb{R}_{>0}$, $\phi \in [0, 1)$ and $\beta_i \in \mathbb{R}$ for all $i = 0, 1, 2, 3$.

Definition 9.1.1 shows the model used to compare the simulated change of carbon dioxide and the measured. The results of the model-based comparison are shown in the tables 9.1, 9.2 and 9.3. The tables 9.1 and 9.2 show the results of Room B and the table 9.3 show the results for Room A. The graphical representation of the results are shown in Figure 9.2 for Room B and in Figure 9.2 for Room A.

### Simulation vs. Room B

The results for Room B are identical for the simulation with the same climate data. Here, all four methods assume similarity for the intercept and the influence of the presence from two hours ago. For the effect of the autoregressive error, the two bootstrapping methods bear similarity, while the other two methods do not. For the two data sets with different climate data, there is a difference with the method standard error, in the influence of the presence of two hours ago. The method using standard errors does not assume similarity here.

|           | St.E  | P.BS  | BS    | Bayes | Sum |
|----------:|-------|-------|-------|-------|-----|
| ar1       | FALSE | TRUE  | TRUE  | FALSE | 2   |
| Pre       | FALSE | FALSE | FALSE | FALSE | 0   |
| L1_Pre    | FALSE | FALSE | FALSE | FALSE | 0   |
| L2_Pre    | TRUE  | TRUE  | TRUE  | TRUE  | 4   |
| intercept | TRUE  | TRUE  | TRUE  | TRUE  | 4   |

Table 9.1: dCO2 comparison results: Original, Shift1, Shift2, Shift3 vs. Room B

|           | St.E  | P.BS  | BS    | Bayes | Sum |
|----------:|-------|-------|-------|-------|-----|
| ar1       | FALSE | TRUE  | TRUE  | FALSE | 2   |
| Pre       | FALSE | FALSE | FALSE | FALSE | 0   |
| L1_Pre    | FALSE | FALSE | FALSE | FALSE | 0   |
| L2_Pre    | FALSE | TRUE  | TRUE  | TRUE  | 3   |
| intercept | TRUE  | TRUE  | TRUE  | TRUE  | 4   |

Table 9.2: dCO2 comparison results: Vienna, Berlin vs. Room B
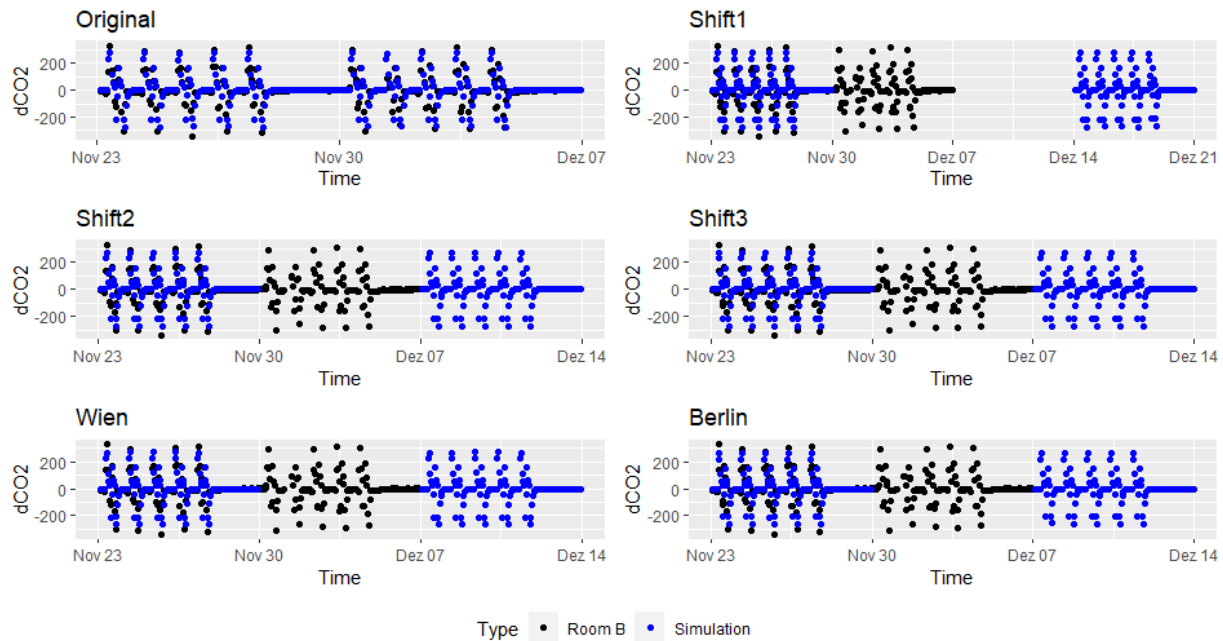
Figure 9.1: Change in carbon dioxide: Simulation vs. Room B

## Simulation vs. Room A

In Room A, the results are identical for each simulation. According to the method via standard error, there is no similarity between simulation and measurement for any parameter. According to the two bootstrap methods, on the other hand, there is a similarity in the autoregressive error in the intercept and the presence of two hours ago. When comparing by credibility intervals, the results are similar to bootstrapping with the difference that no similarity is assumed for the autoregressive error.

|           | St.E  | P.BS  | BS    | Bayes | Sum |
|-----------|-------|-------|-------|-------|-----|
| ar1       | FALSE | TRUE  | TRUE  | FALSE | 2   |
| Pre       | FALSE | FALSE | FALSE | FALSE | 0   |
| L1_Pre    | FALSE | FALSE | FALSE | FALSE | 0   |
| L2_Pre    | FALSE | TRUE  | TRUE  | TRUE  | 3   |
| intercept | FALSE | TRUE  | TRUE  | TRUE  | 3   |

Table 9.3: dCO2 comparison results: Simulation vs. Room A
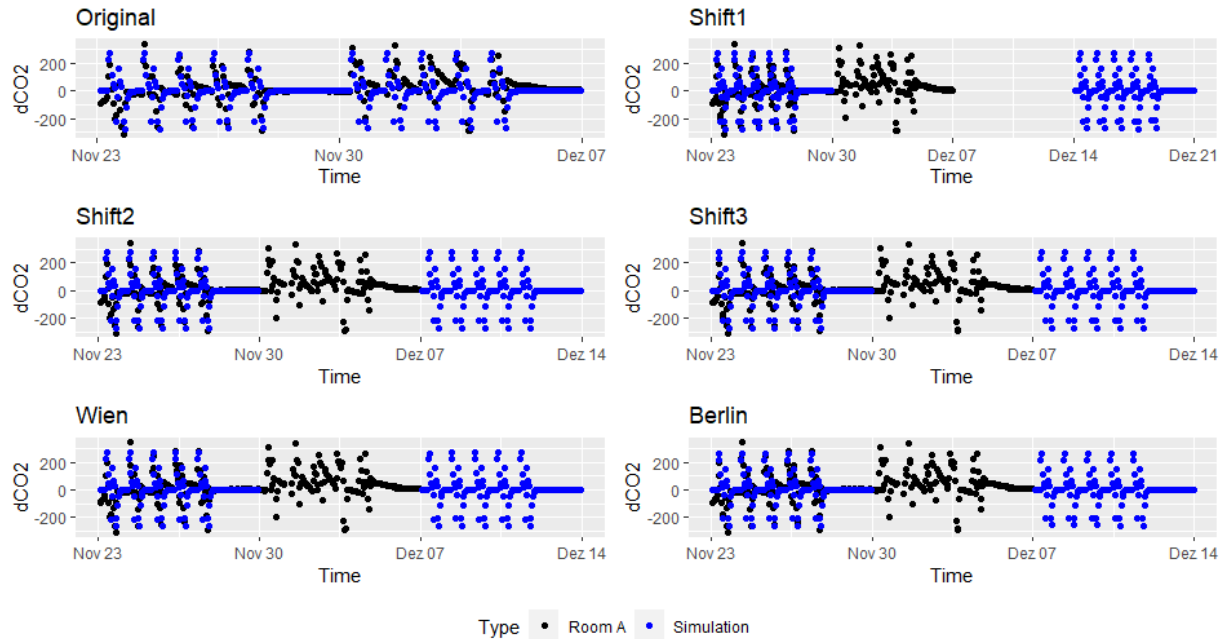
Figure 9.2: Change in carbon dioxide: Simulation vs. Room A

## 9.2    Room temperature

---

**Definition 9.2.1** (dRT-Model). The dRT-Model is given by,

$$\text{RT}_t - \text{RT}_{t-1} = \beta_0 + \beta_1\text{OT}_t + \beta_2\text{dirR}_t + \beta_3\text{difR}_t + \beta_4\text{Pre}_t + \nu_t$$
$$\nu_t = \phi\nu_{t-1} + \epsilon_t,$$

where $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$, $\sigma^2 \in \mathbb{R}_{>0}$, $\phi \in [0,1)$ and $\beta_i \in \mathbb{R}$ for all $i = 0, 1, 2, 3, 4$.

---

Definition 9.2.1 shows the model used to compare the simulated change of room temperature and the measured. The tables 9.4 to 9.8 show the results comparing the simulation with Room B and the tables 9.9 to 9.11 the comparisons with Room A. The graphical representation of the results are shown in Figure 9.3 for Room B and in Figure 9.4 for Room A.

### Simulation vs. Room B

When comparing Room B and the simulations, there are many different results, even if these differences are not particularly large. For the influence of direct and diffuse radiation, all methods assume similarity for each data set. With the outdoor air temperature, three to four methods assume similarity, depending on the data set. The presence influence is assumed to be similar only for the Berlin data set in the non-parametric bootstrap. For

the autoregressive term, the similarity depends on whether the original climate dataset was used. For the intercept, all possible results are provided.
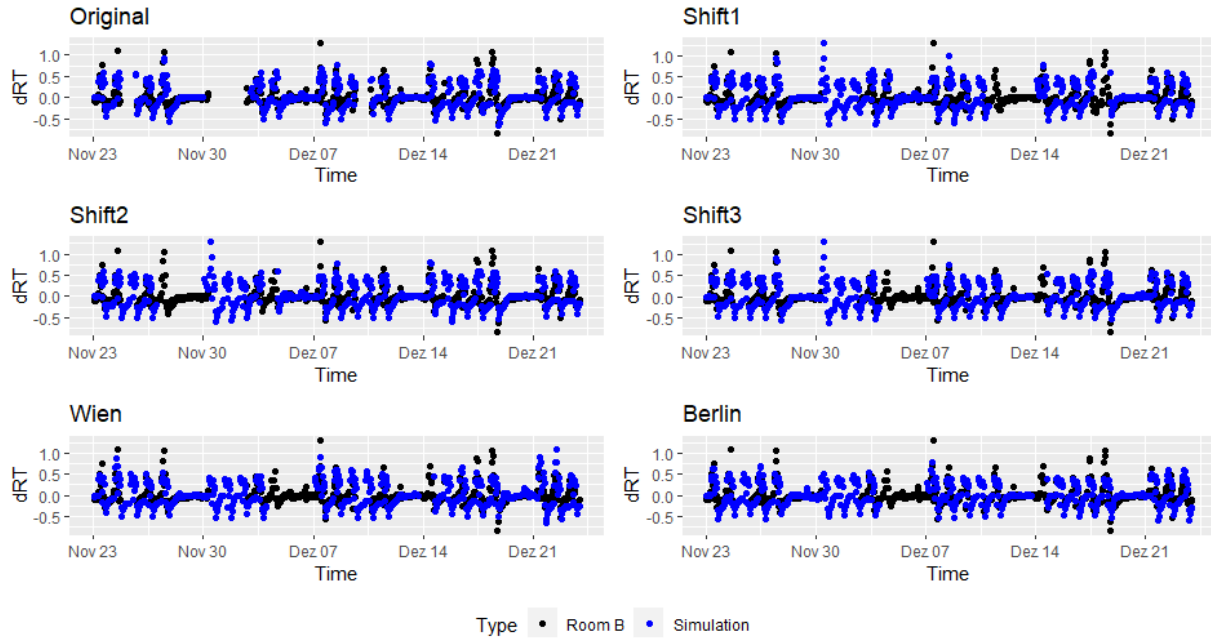


Figure 9.3: Room temperature change: Simulation vs. Room B

|  | St.E | P.BS | BS | Bayes | Sum |
|---|---|---|---|---|---|
| ar1 | FALSE | TRUE | TRUE | TRUE | 3 |
| OT | FALSE | TRUE | TRUE | TRUE | 3 |
| DirRad | TRUE | TRUE | TRUE | TRUE | 4 |
| DiffRad | TRUE | TRUE | TRUE | TRUE | 4 |
| Pre | FALSE | FALSE | FALSE | FALSE | 0 |
| intercept | FALSE | FALSE | TRUE | TRUE | 2 |

Table 9.4: dRT comparison results: Original vs. Room B

|          | St.E   | P.BS   | BS     | Bayes  | Sum |
|---------:|--------|--------|--------|--------|-----|
| ar1      | FALSE  | TRUE   | TRUE   | TRUE   | 3   |
| OT       | FALSE  | TRUE   | TRUE   | TRUE   | 3   |
| DirRad   | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DiffRad  | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| Pre      | FALSE  | FALSE  | FALSE  | FALSE  | 0   |
| intercept| FALSE  | FALSE  | TRUE   | FALSE  | 1   |

Table 9.5: dRT comparison results: Shift1, Shift2 vs. Room B

|          | St.E   | P.BS   | BS     | Bayes  | Sum |
|---------:|--------|--------|--------|--------|-----|
| ar1      | FALSE  | TRUE   | TRUE   | TRUE   | 3   |
| OT       | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DirRad   | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DiffRad  | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| Pre      | FALSE  | FALSE  | FALSE  | FALSE  | 0   |
| intercept| FALSE  | FALSE  | TRUE   | FALSE  | 1   |

Table 9.6: dRT comparison results: Shift3 vs. Room B

|          | St.E   | P.BS   | BS     | Bayes  | Sum |
|---------:|--------|--------|--------|--------|-----|
| ar1      | FALSE  | FALSE  | FALSE  | TRUE   | 1   |
| OT       | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DirRad   | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DiffRad  | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| Pre      | FALSE  | FALSE  | FALSE  | FALSE  | 0   |
| intercept| FALSE  | FALSE  | TRUE   | FALSE  | 1   |

Table 9.7: dRT comparison results: Vienna vs. Room B

|          | St.E   | P.BS   | BS     | Bayes  | Sum |
|---------:|--------|--------|--------|--------|-----|
| ar1      | FALSE  | FALSE  | FALSE  | FALSE  | 0   |
| OT       | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DirRad   | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DiffRad  | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| Pre      | FALSE  | FALSE  | TRUE   | FALSE  | 1   |
| intercept| FALSE  | TRUE   | TRUE   | TRUE   | 3   |

Table 9.8: dRT comparison results: Berlin vs. Room B

## Simulation vs. Room A

Compared with Room A, the similarity seems lower than with Room B, especially since there is no similarity for direct radiation, presence, and autoregressive error for any method. Interestingly, the Berlin dataset has the most similarities with the simulation, even if the differences between the datasets are minor.
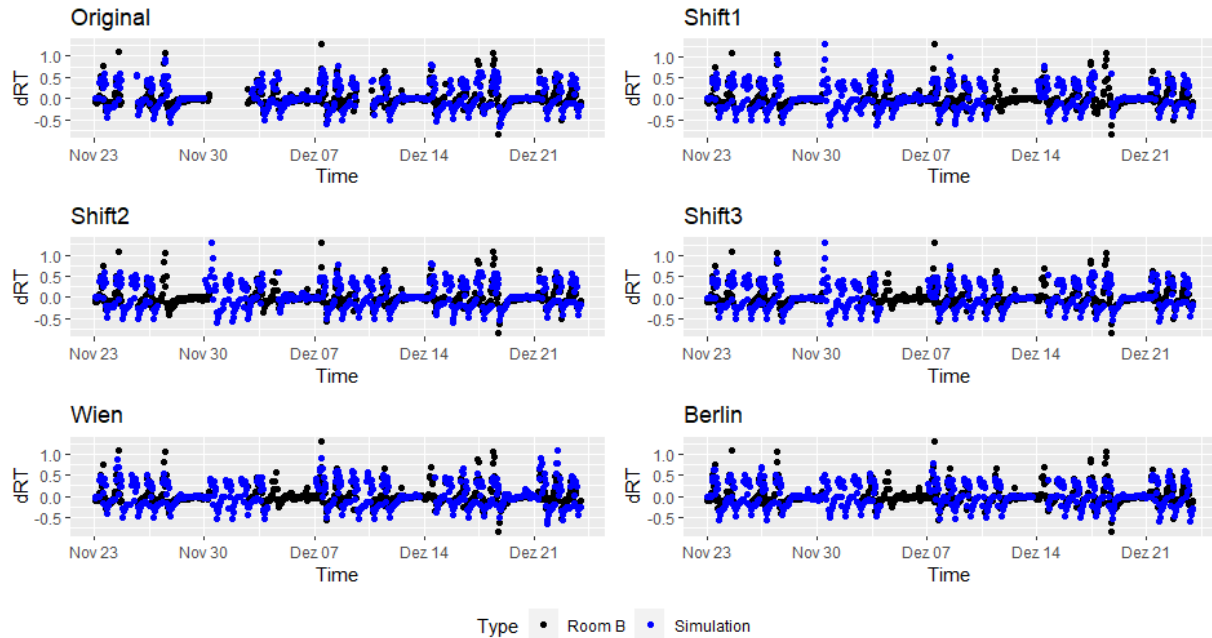


Figure 9.4: Room temperature change: Simulation vs. Room A

|         | St.E  | P.BS  | BS    | Bayes | Sum |
|---------|-------|-------|-------|-------|-----|
| ar1     | FALSE | FALSE | FALSE | FALSE | 0   |
| OT      | FALSE | TRUE  | TRUE  | TRUE  | 3   |
| DirRad  | TRUE  | TRUE  | TRUE  | TRUE  | 4   |
| DiffRad | FALSE | FALSE | FALSE | FALSE | 0   |
| Pre     | FALSE | FALSE | FALSE | FALSE | 0   |
| intercept | FALSE | TRUE | TRUE  | TRUE  | 3   |

Table 9.9: dRT comparison results: Original vs. Room A

|          | St.E   | P.BS   | BS     | Bayes  | Sum |
|---------:|--------|--------|--------|--------|-----|
| ar1      | FALSE  | FALSE  | FALSE  | FALSE  | 0   |
| OT       | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DirRad   | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DiffRad  | FALSE  | FALSE  | FALSE  | FALSE  | 0   |
| Pre      | FALSE  | FALSE  | FALSE  | FALSE  | 0   |
| intercept | FALSE | TRUE   | TRUE   | TRUE   | 3   |

Table 9.10: dRT comparison results: Shift1, Shift2, Shift3, Vienna vs. Room A

|          | St.E   | P.BS   | BS     | Bayes  | Sum |
|---------:|--------|--------|--------|--------|-----|
| ar1      | FALSE  | FALSE  | FALSE  | FALSE  | 0   |
| OT       | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DirRad   | TRUE   | TRUE   | TRUE   | TRUE   | 4   |
| DiffRad  | FALSE  | FALSE  | FALSE  | FALSE  | 0   |
| Pre      | FALSE  | FALSE  | FALSE  | FALSE  | 0   |
| intercept | TRUE  | TRUE   | TRUE   | TRUE   | 4   |

Table 9.11: dRT comparison results: Berlin vs. Room A

## 9.3 Total heating power

> **Definition 9.3.1** (dTHP-Model). The dTHP-Model is given by,
>
> $$\text{THP}_t - \text{THP}_{t-1} = \beta_0 + \beta_1 \text{OT}_t + \beta_2 \text{dirR}_t + \beta_3 \text{difR}_t + \beta_4 \text{Pre}_t + \nu_t$$
> $$\nu_t = \phi \nu_{t-1} + \epsilon_t,$$
>
> where $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$, $\sigma^2 \in \mathbb{R}_{>0}$, $\phi \in [0,1)$ and $\beta_i \in \mathbb{R}$ for all $i = 0, 1, 2, 3, 4$.

Definition 9.3.1 shows the model used to compare the simulated change of room temperature and the measured. The tables 9.12 to 9.14 show the results comparing the simulation with Room B and the tables 9.15 to 9.17 the comparisons with Room A. The graphical representation of the results are shown in Figure 9.5 for Room B and in Figure 9.6 for Room A.

### Simulation vs. Room B

When looking at the results, it is noticeable that the influence of the autoregressive error varies greatly depending on the data set. There is no similarity in the autoregressive error for the data sets with external climate data and Shift2. Whereas for the original data set, all methods assume a similarity at this point. The data sets Shift1 and Shift3 are in-between. For the influence of the two radiations and the presence, the similarity is not

assumed. Unlike the intercept here, there seem to be no significant differences between simulation and measurement.
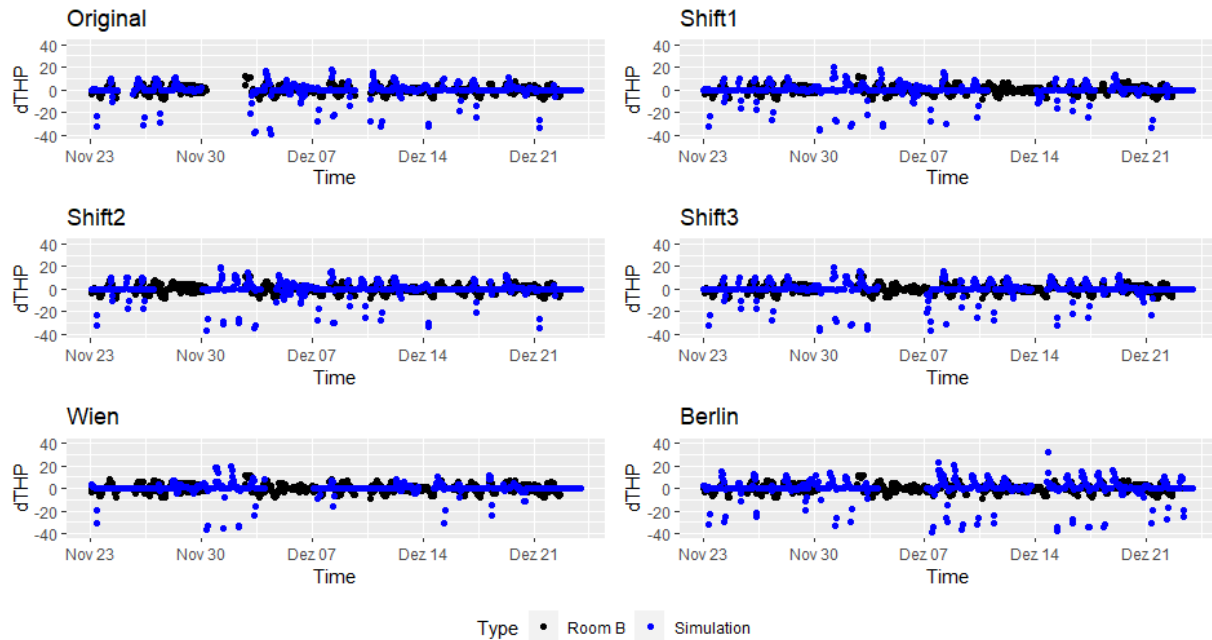


Figure 9.5: Change in total heating power: Simulation vs. Room B

|        | St.E  | P.BS  | BS    | Bayes | Sum |
|-------:|-------|-------|-------|-------|-----|
| ar1    | TRUE  | TRUE  | TRUE  | TRUE  | 4   |
| OT     | FALSE | TRUE  | TRUE  | TRUE  | 3   |
| DirRad | FALSE | FALSE | FALSE | FALSE | 0   |
| DiffRad| FALSE | FALSE | FALSE | FALSE | 0   |
| Pre    | FALSE | FALSE | FALSE | FALSE | 0   |
| intercept | TRUE | TRUE | TRUE | TRUE | 4 |

Table 9.12: dTHP comparison results: Original vs. Room B

|          | St.E  | P.BS  | BS    | Bayes | Sum |
|----------|-------|-------|-------|-------|-----|
| ar1      | FALSE | TRUE  | TRUE  | TRUE  | 3   |
| OT       | FALSE | FALSE | TRUE  | TRUE  | 2   |
| DirRad   | FALSE | FALSE | FALSE | FALSE | 0   |
| DiffRad  | FALSE | FALSE | FALSE | FALSE | 0   |
| Pre      | FALSE | FALSE | FALSE | FALSE | 0   |
| intercept| TRUE  | TRUE  | TRUE  | TRUE  | 4   |

Table 9.13: dTHP comparison results: Shift1, Shift3 vs. Room B

|          | St.E  | P.BS  | BS    | Bayes | Sum |
|----------|-------|-------|-------|-------|-----|
| ar1      | FALSE | FALSE | FALSE | FALSE | 0   |
| OT       | FALSE | TRUE  | TRUE  | TRUE  | 3   |
| DirRad   | FALSE | FALSE | FALSE | FALSE | 0   |
| DiffRad  | FALSE | FALSE | FALSE | FALSE | 0   |
| Pre      | FALSE | FALSE | FALSE | FALSE | 0   |
| intercept| TRUE  | TRUE  | TRUE  | TRUE  | 4   |

Table 9.14: dTHP comparison results: Shift2, Vienna, Berlin vs. Room B

## Simulation vs. Room A

The Results from Room A are comparable to those from Room B. With the difference that the simulations with the original climate data have no agreement in the autoregressive part, the simulations with external climate data do.

|          | St.E  | P.BS  | BS    | Bayes | Sum |
|----------|-------|-------|-------|-------|-----|
| ar1      | FALSE | FALSE | FALSE | FALSE | 0   |
| OT       | FALSE | TRUE  | TRUE  | TRUE  | 3   |
| DirRad   | FALSE | FALSE | FALSE | FALSE | 0   |
| DiffRad  | FALSE | FALSE | FALSE | FALSE | 0   |
| Pre      | FALSE | FALSE | FALSE | FALSE | 0   |
| intercept| TRUE  | TRUE  | TRUE  | TRUE  | 4   |

Table 9.15: dTHP comparison results: Original, Shift1, Shift2, Shift3 vs. Room A
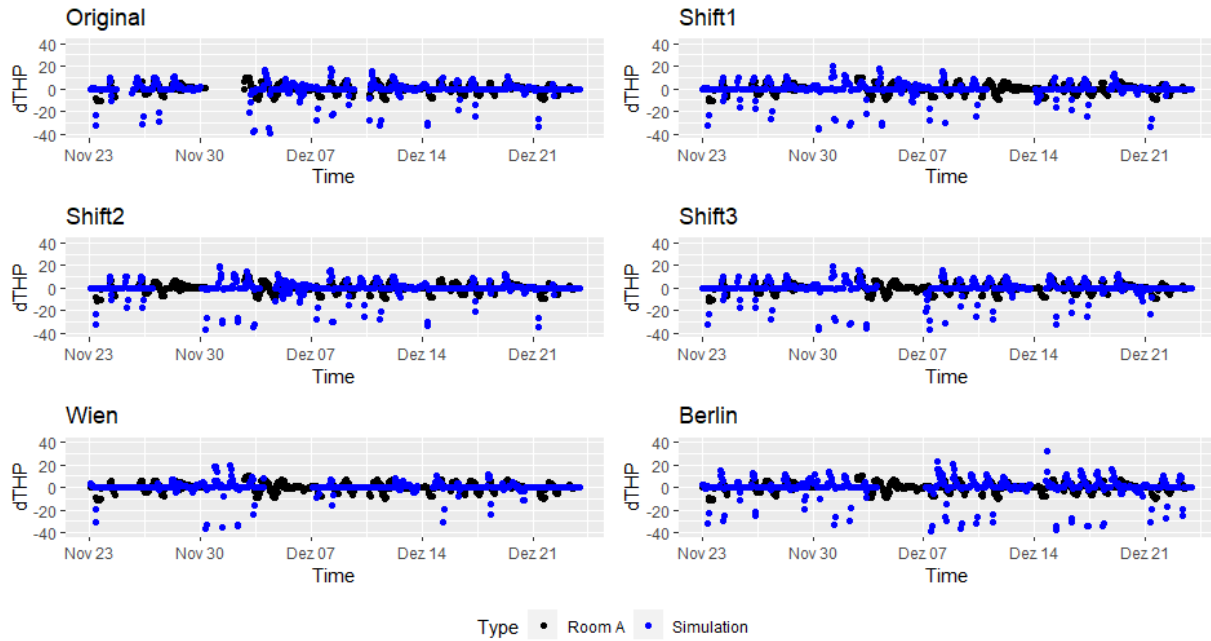
Figure 9.6: Change in total heating power: Simulation vs. Room A

|  | St.E | P.BS | BS | Bayes | Sum |
|---|---|---|---|---|---|
| ar1 | FALSE | TRUE | TRUE | TRUE | 3 |
| OT | FALSE | TRUE | TRUE | TRUE | 3 |
| DirRad | FALSE | FALSE | FALSE | FALSE | 0 |
| DiffRad | FALSE | FALSE | FALSE | FALSE | 0 |
| Pre | FALSE | FALSE | FALSE | FALSE | 0 |
| intercept | TRUE | TRUE | TRUE | TRUE | 4 |

Table 9.16: dTHP comparison results: Vienna vs. Room A

|  | St.E | P.BS | BS | Bayes | Sum |
|---|---|---|---|---|---|
| ar1 | FALSE | TRUE | TRUE | FALSE | 2 |
| OT | FALSE | TRUE | TRUE | TRUE | 3 |
| DirRad | FALSE | FALSE | FALSE | FALSE | 0 |
| DiffRad | FALSE | FALSE | FALSE | FALSE | 0 |
| Pre | FALSE | FALSE | FALSE | FALSE | 0 |
| intercept | TRUE | TRUE | TRUE | TRUE | 4 |

Table 9.17: dTHP comparison results: Berlin vs. Room A

## 9.4   Overview

Finally, the table 9.18 shows an overview for Room B, and 9.19 shows a summary for Room A. The values in the table are the number of matches for each data set and per measurement. The number of matches for Room B is almost always lower than those for Room A. The only exception is the change in total heat power for the Vienna and Berlin data sets. It seems like there is little difference between the simulations. This is not a bad sign since the same building was simulated each time. The usefulness of model-based comparisons is discussed at the end of the following chapter.

Furthermore, when looking at the individual results, it is noticeable that if three out of four methods find a similarity, the method uses standard error whose result does not agree with the others. This means this method seems to be more stringent than the others. That is probably because the other methods compare an interval at the 95% level, while the standard error method uses an interval at the 68% level, assuming a normal distribution. Theoretically, this could be adjusted, but it is not problematic to combine different strict methods.

| Simulation | dCO2 | dRT | dTHP |
|---|---|---|---|
| Original | 10 | 16 | 11 |
| Shift1 | 10 | 15 | 9 |
| Shift2 | 10 | 15 | 7 |
| Shift3 | 9 | 16 | 9 |
| Vienna | 9 | 14 | 7 |
| Berlin | 9 | 16 | 7 |

Table 9.18: Number of matches with Room B

| Simulation | dCO2 | dRT | dTHP |
|---|---|---|---|
| Original | 8 | 10 | 7 |
| Shift1 | 8 | 11 | 7 |
| Shift2 | 8 | 11 | 7 |
| Shift3 | 8 | 11 | 7 |
| Vienna | 8 | 11 | 10 |
| Berlin | 8 | 12 | 9 |

Table 9.19: Number of matches with Room A

# Chapter 10

# Fault Detection

In the following chapter, models for fault detection are presented one by one. These are first trained with the *original* simulation data. In a separate section, a comparison is made with the other simulations. The goal is to detect the faults in Room B and Room A. Quick reminder Room B is the room with the artificial user, and Room A is the room with the real user.

The three faults are, firstly, the fault on the ventilation unit. Models that start with $V$ for ventilation are to detect this fault. Models to detect the fault on the heating unit start with $H$ for heating, and accordingly, models for the bypass fault begin with $B$.

The results are first presented in figures in these true predictions are colored green and wrong orange. In addition, the positive predictions (the model does assume a fault) are colored light, and the negative predictions (the model does not assume a fault) are colored dark. This means that a light green point means that the decision for the fault is correct, and a dark orange point means that a fault is overlooked at this point.

Each of the following models is a logistic regression see section 4.4. For each model, the threshold is chosen by maximizing the sum of sensitivity and specificity in the training data.

## 10.1 Ventilation Fault

The ventilation fault strongly affects the carbon dioxide ($CO_2$) of a room. Therefore the Figure 10.1 shows the $CO_2$ curve in the simulation compared to the $CO_2$ curve in Room B, while Figure 10.2 shows Room A (real user) compared with Room B (artificial user). In the simulation, it can be seen that the $CO_2$ content rises to approximately 1200 ppm five days in a row and then falls again. Afterward, the $CO_2$ is relatively constant for two days at about 500 ppm. This means that in the fault-free state, the presence areas are recognizable. As soon as a fault occurs, which can be seen in Figure 10.1 and Figure 10.2 by the orange color gradient, the $CO_2$ content increases significantly higher than in the usual case. After fault correction, it takes a few hours for the $CO_2$ to normalize again. This delay must be taken into account for the modeling.
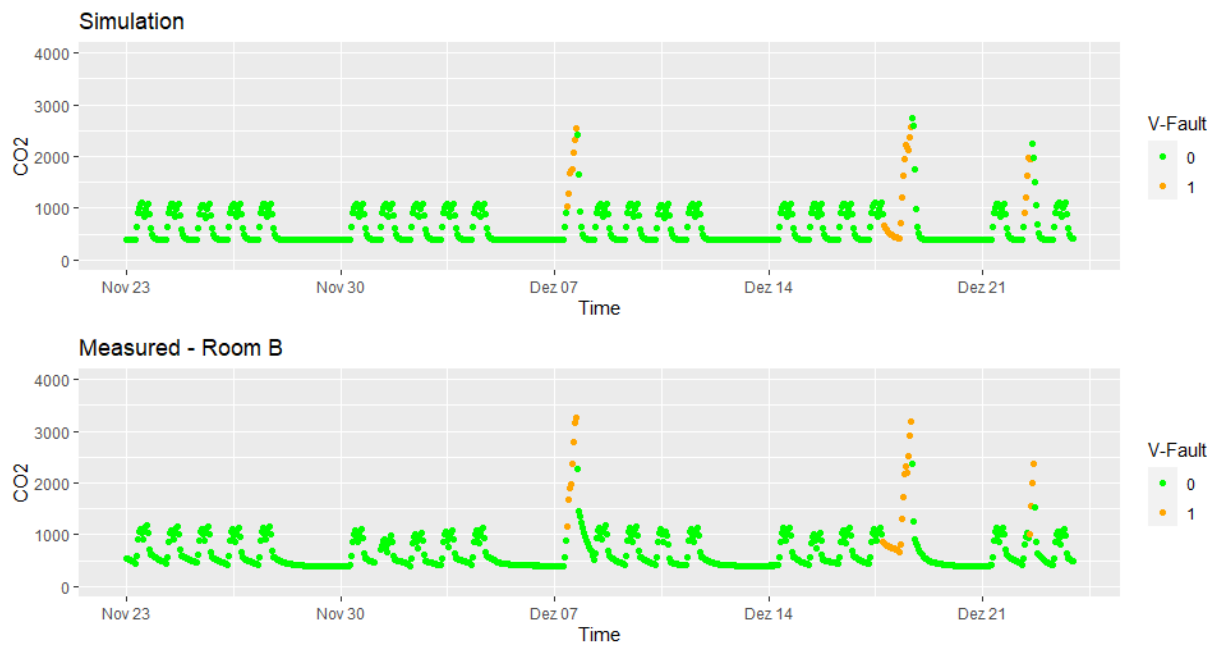
Figure 10.1: $CO_2$ simulation and $CO_2$ Room B



Figure 10.2: $CO_2$ Room A and $CO_2$ Room B

## V-CO2

The first approach tries to detect the fault with the help of the $CO_2$ concentration and the presence. This model is defined in definition 10.1.1.

---

**Definition 10.1.1** (V-CO2). Let $\pi_t \in (0,1)$ be the probability for the ventilation fault to the time point $t$. Then the V-CO2 model is given by,

$$\text{logit}(\pi_t) = \beta_0 + \beta_1 CO2_t + \beta_2 CO2_{t-1} + \beta_3 \text{Pre}_{t-1},$$

where $\beta_i \in \mathbb{R}$ for all $i = 0,1,2,3$.

---

The coefficients were trained with the simulated data. Then, the trained model is used to detect the faults on the same simulated data in Room B and Room A. The graphical results of this model are shown in Figure 10.3, 10.4 and 10.5. The relevant results are the results for Room A and Room B, but the result on the simulation data is helpful to identify what the model considers as a fault.
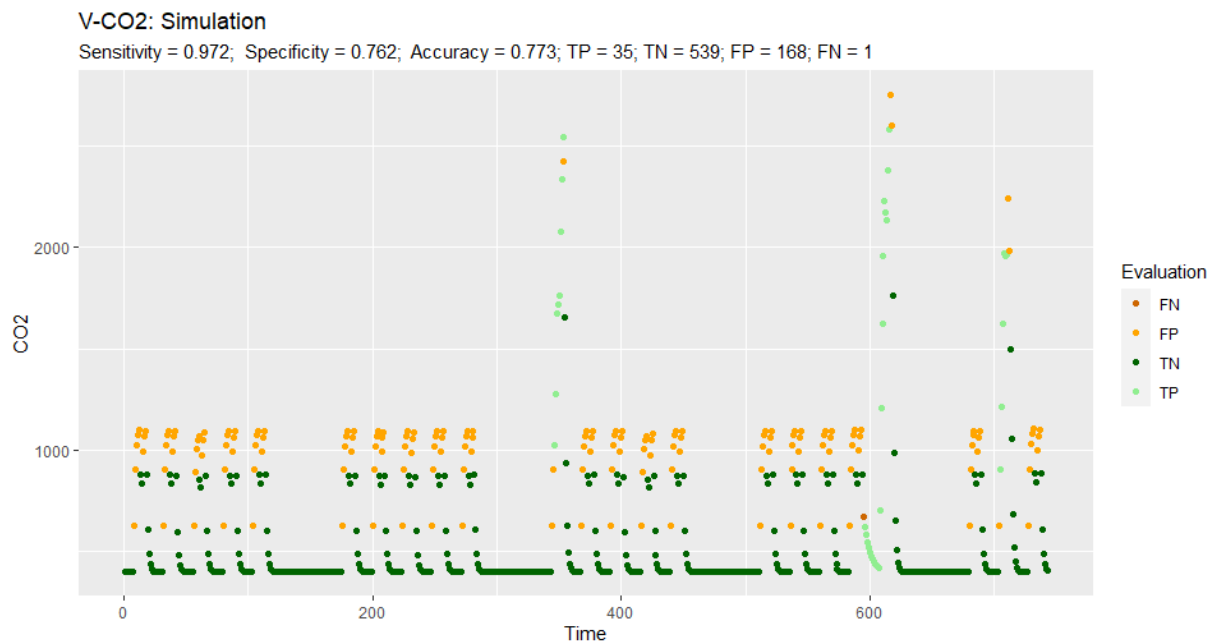


Figure 10.3: Results V-CO2: Simulation

The V-CO2 model has a high sensitivity. Therefore the periods with faults are also detected as a fault with high accuracy. Unfortunately, this is at the expense of specificity. When looking at the graphs, it is noticeable that many presence periods are detected falsely as faults. That is why this model is not problematic for practical use. Finally, table 10.1 shows the estimated coefficients.
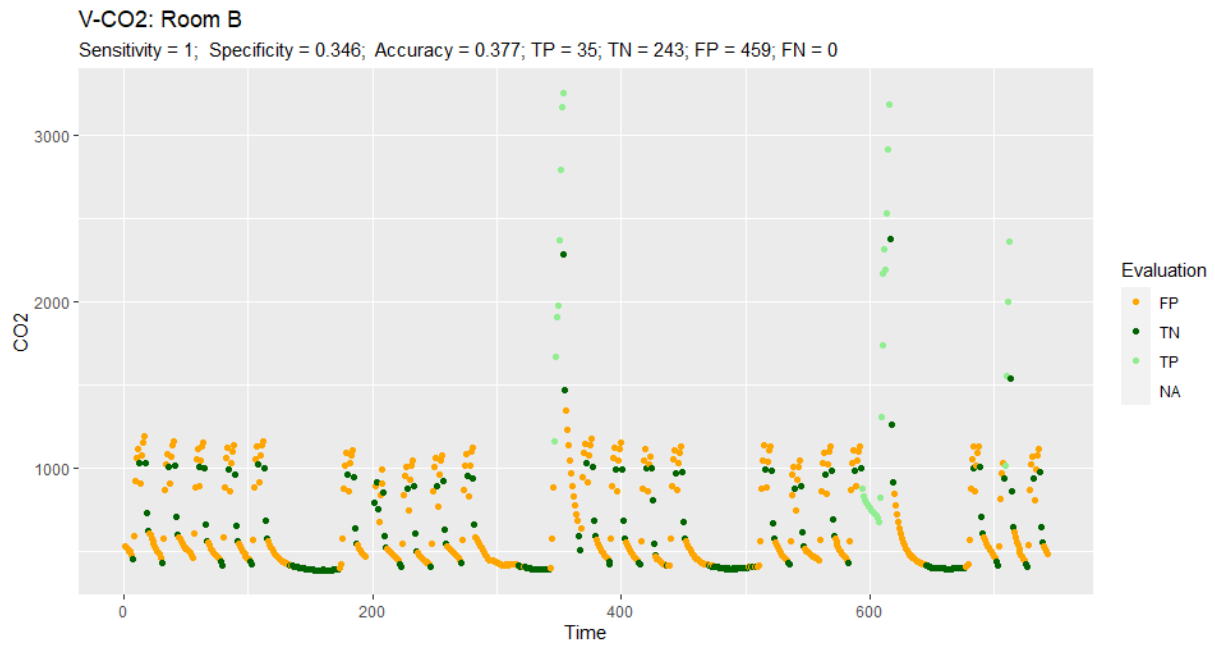
Figure 10.4: Results V-CO2: Room B
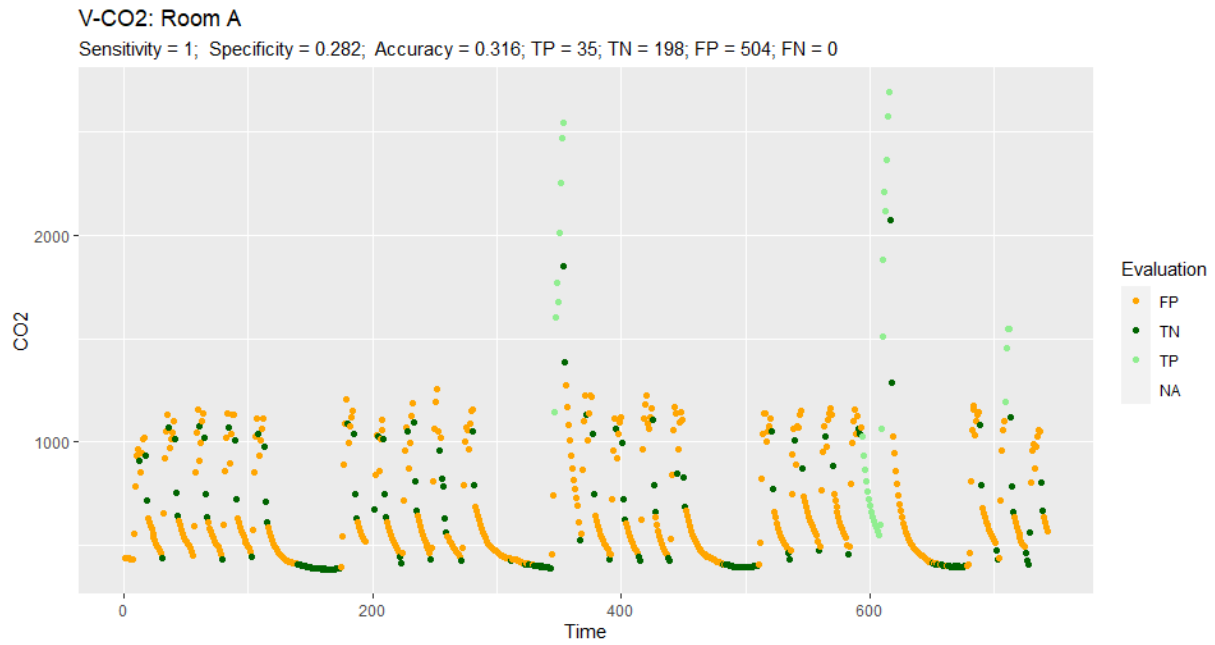


Figure 10.5: Results V-CO2: Room A

| | Estimate | Std. Error | z value | $\Pr(> |z|)$ |
|---|---|---|---|---|
| (Intercept) | -5.1604 | 0.3956 | -13.04 | 0.0000 |
| CO2 | 0.0099 | 0.0016 | 6.02 | 0.0000 |
| L1_CO2 | -0.0068 | 0.0016 | -4.34 | 0.0000 |
| L1_Pre1 | -1.8789 | 0.5804 | -3.24 | 0.0012 |

Table 10.1: Estimated coefficients for the V-CO2 model
L1 is short for time lag of one hour.

## V-Categorical

The second approach attempts to explain the ventilation fault using categorical variables. One category indicates whether the $CO_2$ content is above a certain threshold. Since the ventilation fault is not present in the first two weeks, this period is helpful to define a threshold value. An appropriate value would be 1200 ppm because $CO_2$ is constantly below this value within this period. In addition, a second variable is needed, which indicates whether the $CO_2$ content in the room has decreased compared to the last time. The full model is defined in definition 10.1.2.

---

**Definition 10.1.2** (V-Categorical). Let $\pi_t \in (0, 1)$ be the probability for the ventilation fault to the time point $t$. Then the V-Categorical model is given by,

$$\text{logit}(\pi_t) = \beta_0 + \beta_1 \text{BC}_t + \beta_2 \text{DC}_t,$$

where $\beta_i \in \mathbb{R}$ for all $i = 0, 1, 2$ and

$$\text{BC}_t := \begin{cases} 1, & \text{CO2}_t > 1200 \\ 0, & \text{CO2}_t \leq 1200 \end{cases},$$

$$\text{DC}_t := \begin{cases} 1, & \text{CO2}_t > \text{CO2}_{t-1} \\ 0, & \text{CO2}_t \leq \text{CO2}_{t-1} \end{cases}.$$

---

Figure 10.6 shows the results by using this model to detect the faults on the simulation data, whereas Figure 10.7 and 10.8 show the results in Room B and Room A, respectively. In contrast to the V-CO2 model, the specificity of this model is very high, whereas the sensitivity is fragile. To put it simply, the V-Categorical model recognizes the periods without faults very well, but many faults are not recognized. A problematic region is the period around 600 hours. There the ventilation fault is present while no person is in the room. Therefore, the threshold value of 1200 ppm is not exceeded, and the fault is not detected.
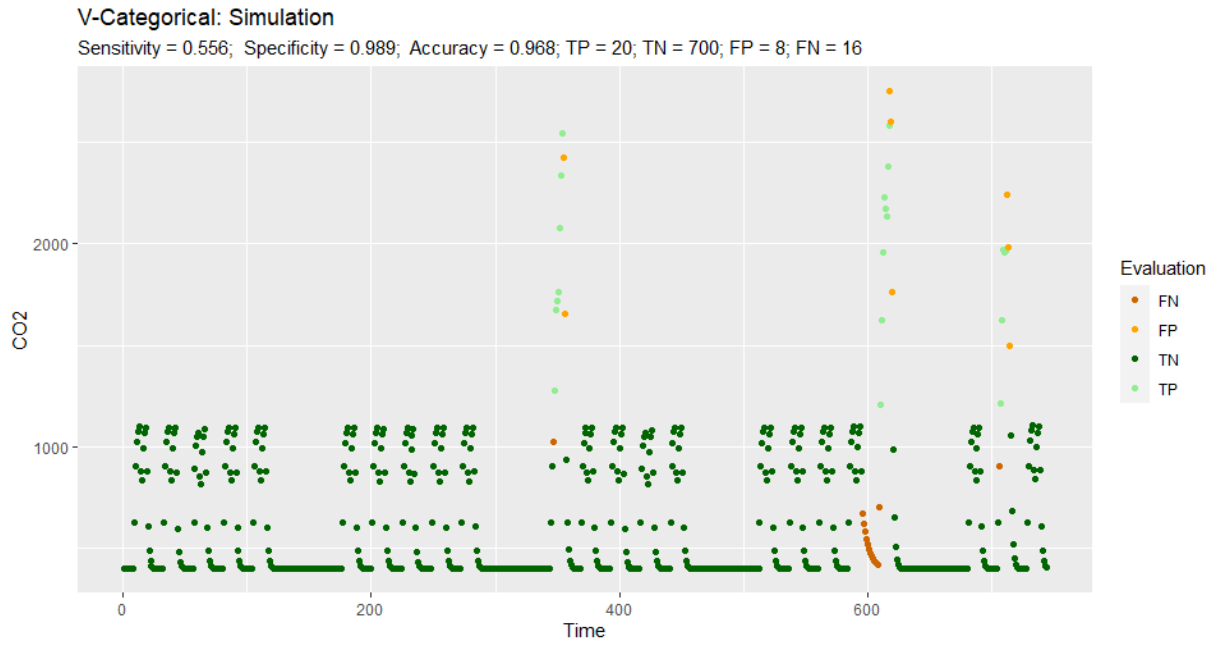
Figure 10.6: Results V-Categorical: Simulation
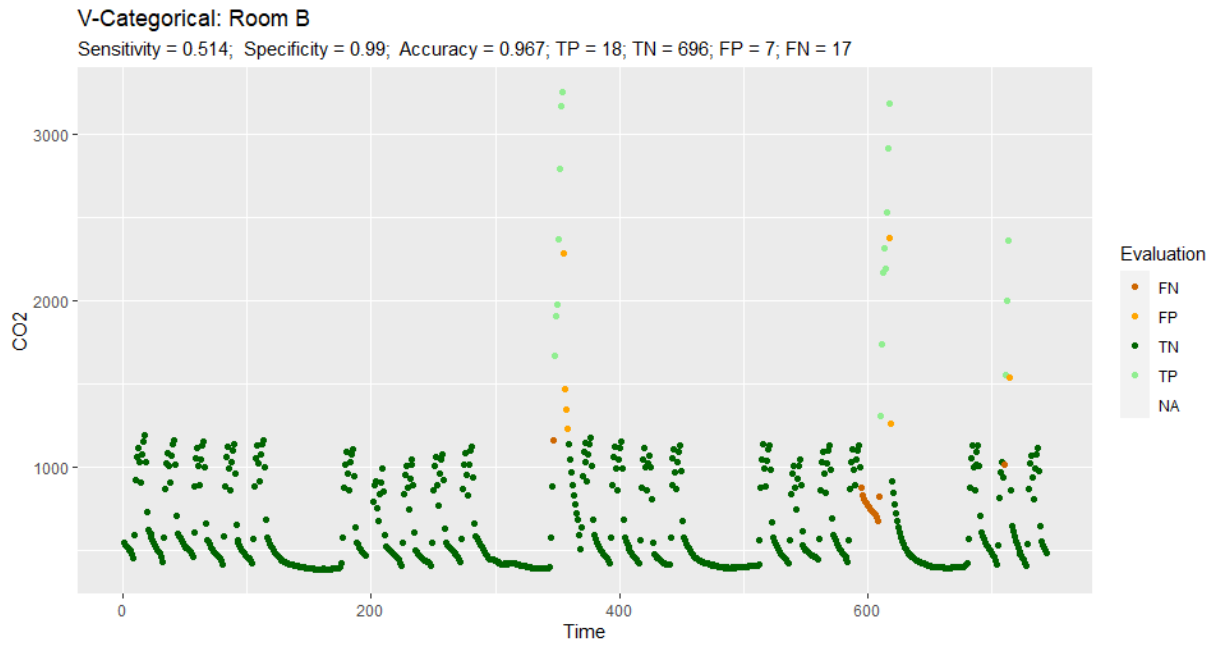


Figure 10.7: Results V-Categorical: Room B
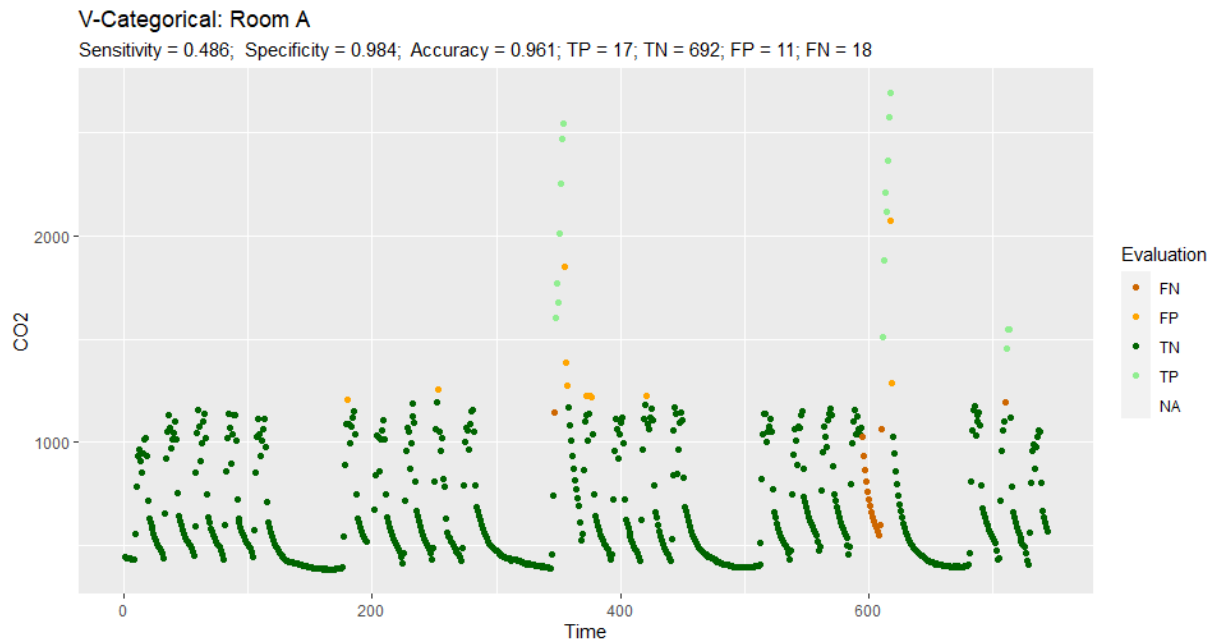
Figure 10.8: Results V-Categorical: Room A

Unlike the V-CO2 model, however, the V-Categorical model is applicable in practice. If no one is in the room, it is questionable if the detection of the ventilation fault is relevant. The other periods, however, are detected well even if the characteristics do not reflect this directly. Nevertheless, an attempt is made to find a model with similarly high specificity and good sensitivity in the following. Table 10.2 shows the estimated coefficients.

|             | Estimate | Std. Error | z value | Pr($> |z|$) |
|-------------|----------|------------|---------|-------------|
| (Intercept) | -3.9525  | 0.2986     | -13.23  | 0.0000      |
| BC1         | 4.4770   | 0.5077     | 8.82    | 0.0000      |
| DC1         | 0.6012   | 0.4533     | 1.33    | 0.1847      |

Table 10.2: Estimated coefficients for the V-Categorical model

## Carbon dioxide presence category (CPC)

Since the categorical variables have been shown to be helpful in V-Categorical, this idea is extended here. This time, categories are created that reflect the relationship between presence and the change in $CO_2$. For this, five categories are defined. The first category indicates whether the $CO_2$ content has remained unchanged since the last time. The other categories indicate whether the $CO_2$ content increased or decreased while no person or one person was present see definition 10.1.3. Since this model considers the interaction between the presence and $CO_2$ change, it will be called the carbon dioxide presence category model or short CPC model. In addition, the $CO_2$ at the current time and from an hour ago is also included.

---

**Definition 10.1.3** (V-CPC). Let $\pi_t \in (0,1)$ be the probability for the ventilation fault to the time point $t$. Then the V-CPC model is given by,

$$\log \frac{\pi_t}{1 - \pi_t} = \beta_0 + \beta_1 CO2_t + \beta_2 CO2_{t-1} + \sum_{k=2}^{5} \beta_{1+k} \mathcal{I}(\mathrm{CPC}_t = C_k)$$

where $\mathcal{I}$ is the indicator function, $\beta_i \in \mathbb{R}$ for all $i = 0, 1, 2, 3, 4, 5$ and

$$\mathrm{CPC}_t = \begin{cases} C_1, & CO2_t = CO2_{t-1} \\ C_2, & CO2_t > CO2_{t-1} \wedge \mathrm{Pre}_t = 1 \\ C_3, & CO2_t > CO2_{t-1} \wedge \mathrm{Pre}_t = 0 \\ C_4, & CO2_t < CO2_{t-1} \wedge \mathrm{Pre}_t = 1 \\ C_5, & CO2_t < CO2_{t-1} \wedge \mathrm{Pre}_t = 0 \end{cases}.$$

(Note that $C_1$ is part of the intercept.)

---

Figure 10.9 shows the results for the simulation data. It can be seen that both the fault periods and the fault-free periods are learned very well. This is reflected in the high sensitivity of 94.4% and specificity of 96.5%. So by looking at the simulation data only, it seems like this is an excellent model.

However, looking at the figures 10.10 and 10.11, this model seems to be very weak. Therefore, the question arises why the results are so different. To examined this question, the coefficients in table 10.3 are used. Noticeable are the high standard errors of the intercept and the CPC categories, which lead to p-values close to one. Thus, none of the CPC categories is significant. The question arises, what is the problem with the CPC categories? For this purpose, the table 10.4 can be considered. This table indicates how many times with fault and how many times without fault the respective category contains. Most categories rarely occur at fault-free periods. This creates numerical problems, which is why the standard error becomes so large. Furthermore, the difference between the simulation and measurement data indicates overfitting.
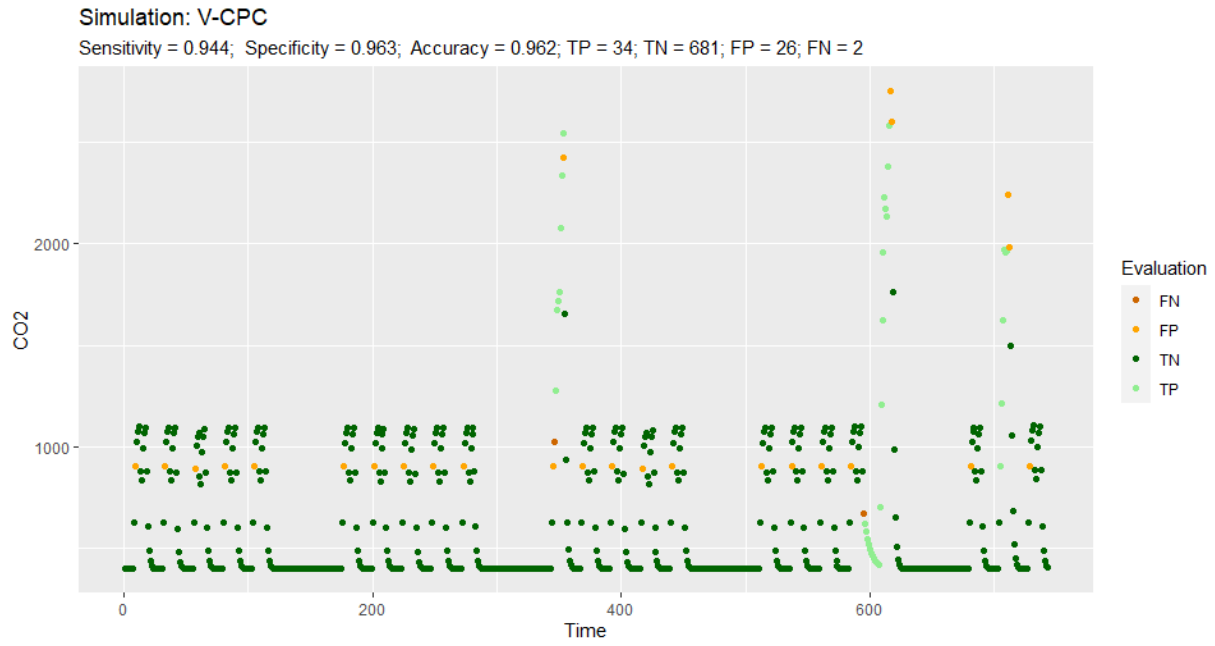
Figure 10.9: Results V-CPC: Simulation
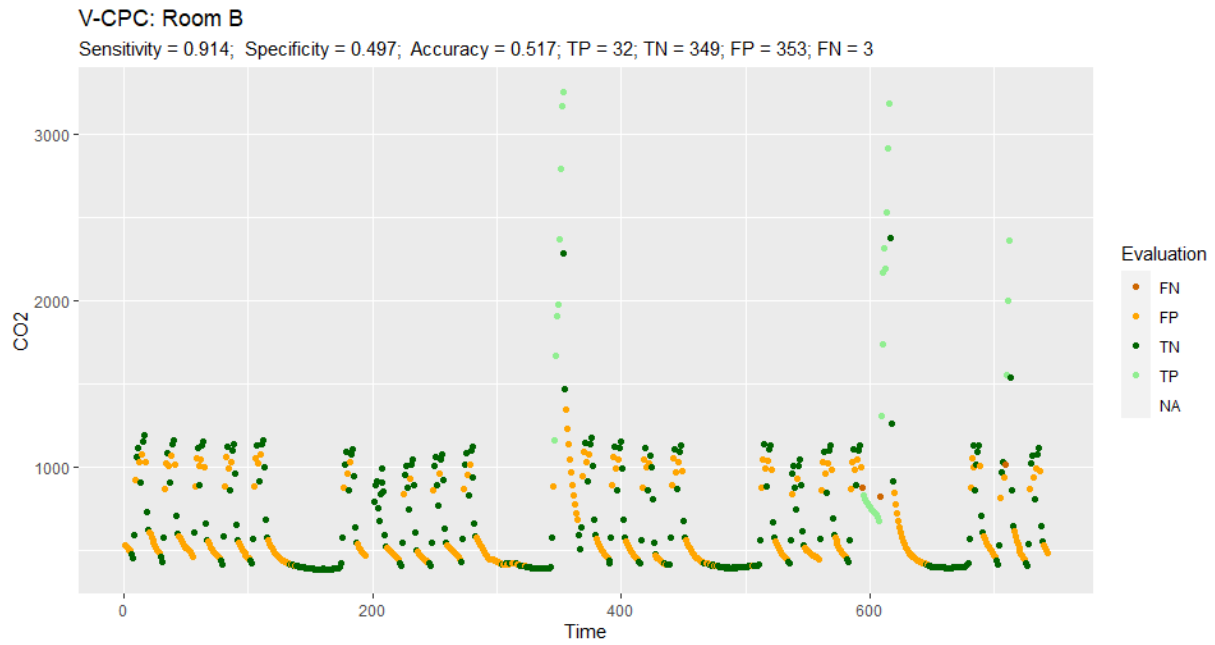


Figure 10.10: Results V-CPC: Room B
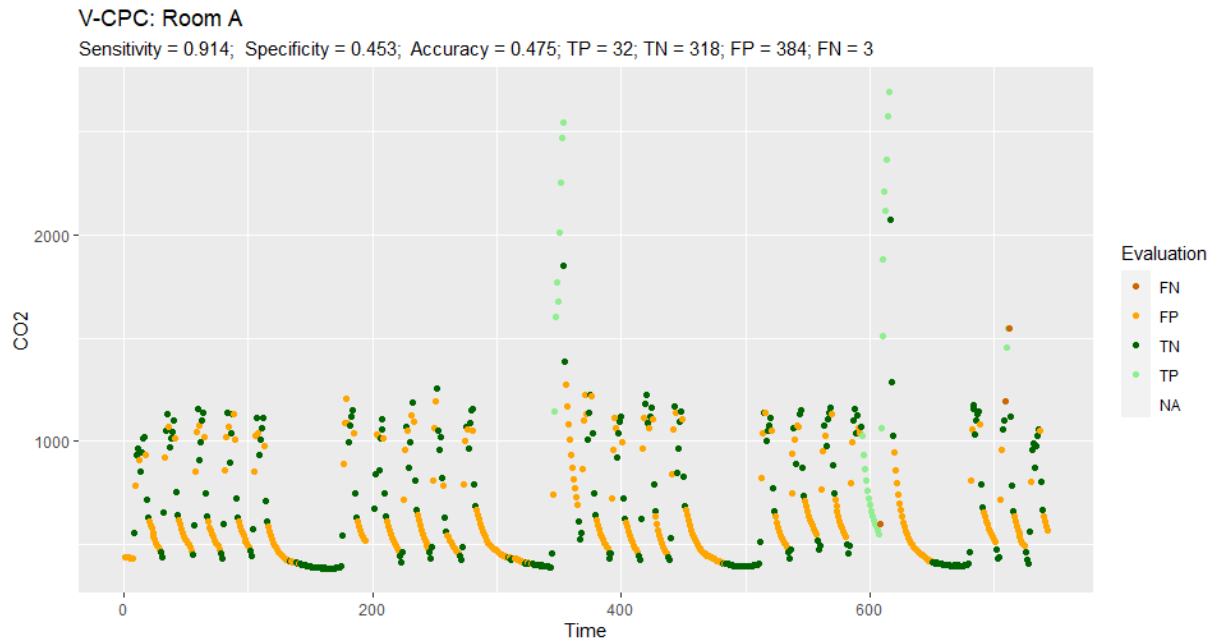
Figure 10.11: Results V-CPC: Room A

|              | Estimate | Std. Error | z value | $\Pr(> |z|)$ |
|---|---|---|---|---|
| (Intercept) | -21.8323 | 1154.1488 | -0.02 | 0.9849 |
| CO2 | 0.0151 | 0.0030 | 5.09 | 0.0000 |
| L1_CO2 | -0.0119 | 0.0028 | -4.23 | 0.0000 |
| cpcC2 | 13.9067 | 1154.1491 | 0.01 | 0.9904 |
| cpcC3 | 12.6448 | 1154.1499 | 0.01 | 0.9913 |
| cpcC4 | 16.4350 | 1154.1492 | 0.01 | 0.9886 |
| cpcC5 | 17.7070 | 1154.1488 | 0.02 | 0.9878 |

Table 10.3: Estimated coefficients for the V-CPC model

| CPC-Category | hours without faults | hours with faults |
|---|---|---|
| $C_1$ | 236 | 0 |
| $C_2$ | 124 | 16 |
| $C_3$ | 41 | 4 |
| $C_4$ | 41 | 3 |
| $C_5$ | 265 | 13 |

Table 10.4: Simulation data CPC-Categories during fault and fault free time periods.

## V-RH

The next attempt tries to use relative humidity (RH) instead of carbon dioxide. Definition 10.1.4 shows the model. The presence periods have worsened the AIC, which is why it is not considered here. The figures 10.12, 10.13 and 10.14 show the results on the simulation the Room B and Room A. The fault-free period is detected with an accuracy of over 97% each. Nevertheless, the accuracy of the faulty period is not even 50% for Room A and B. These results show that relative humidity can be helpful to detect the ventilation fault. However, it seems complicated to detect these faults using relative humidity alone.

---

**Definition 10.1.4** (V-RH). Let $\pi_t \in (0,1)$ be the probability for the ventilation fault to the time point $t$. Then the V-RH model is given by,

$$\text{logit}(\pi_t) = \beta_0 + \beta_1 \text{RH}_t + \beta_2 \text{RH}_{t-1},$$

where $\beta_i \in \mathbb{R}$ for all $i = 0, 1, 2$.

---
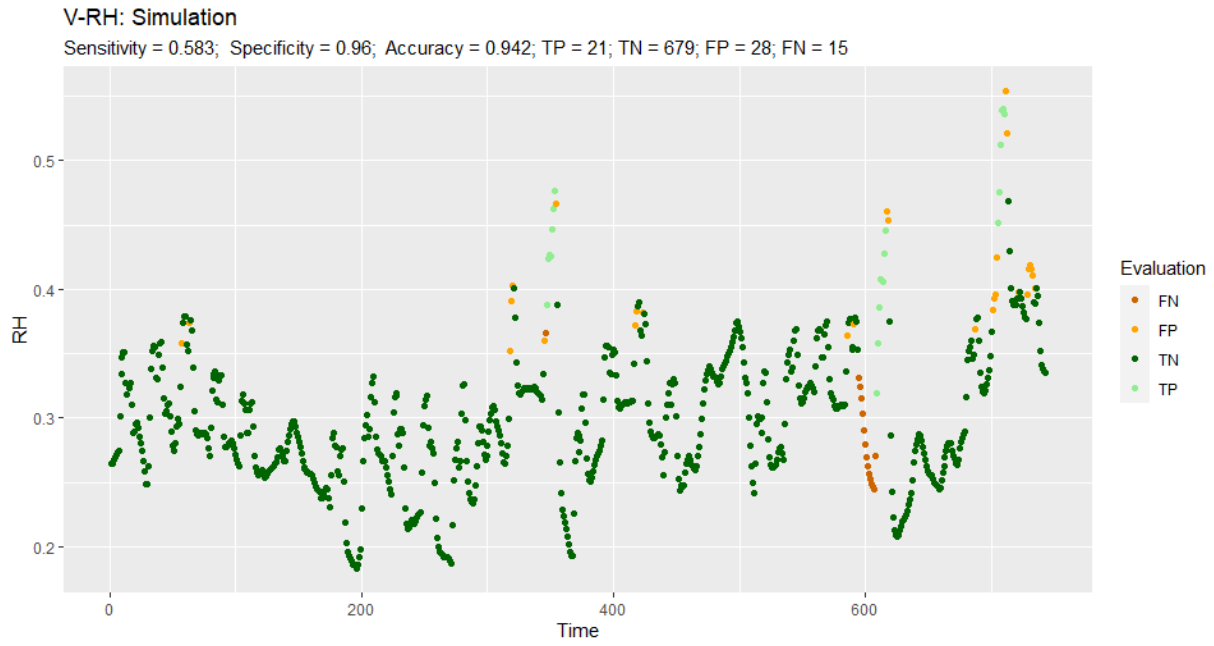
Figure 10.12: Results V-RH: Simulation
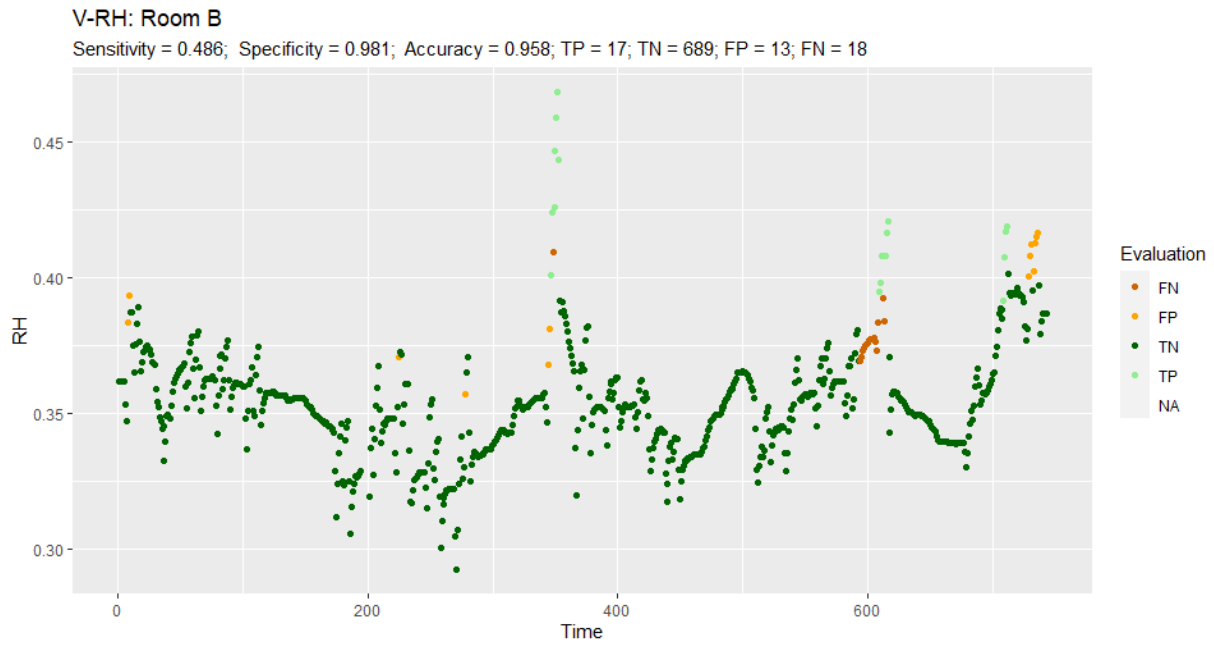


Figure 10.13: Results V-RH: Room B
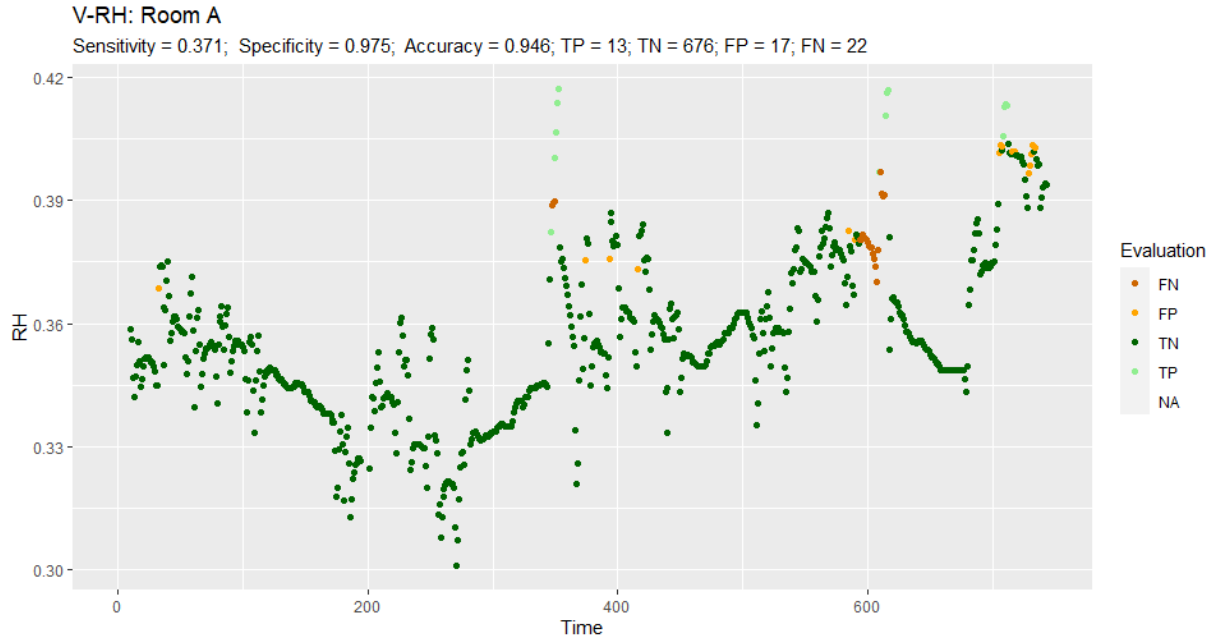
Figure 10.14: Results V-RH: Room A

## Stepwise AIC

For the last models for the ventilation fault, the forwarded stepwise AIC is used. It starts with the intercept model, and after each step, it searches for the variable that minimizes the AIC the most. If the AIC cannot be further minimized by adding more covariates, the algorithm stops. So far, only time lags of one hour have been used. Now, time lags of several hours are also to be tested. Therefore two models are defined in definition 10.1.5 and definition 10.1.6. These are not the models used but the maximum possible for each maximum time lag $l$ using the stepwise AIC. V-AIC can only use the $CO_2$ content at different times and the presence, while V-AIC-RH can also use relative humidity. So overall, 20 models are tried in this section.

---

**Definition 10.1.5** (V-AIC-RH). Let $\pi_t \in (0, 1)$ be the probability for the ventilation fault to the time point $t$ and $l \in \{1, \ldots, 10\}$. Then the V-AIC-RH models are given by,

$$\text{logit}(\pi_t) = \alpha + \sum_{i=0}^{l} x_{t-i}\beta_i$$

where $x_t = (\text{CO2}_t, \text{Pre}_t, \text{RH}_t)$, $\alpha \in \mathbb{R}$, $\beta_i \in \mathbb{R}^3$ for all $i = 0, 1, \ldots, l$.

---

**Definition 10.1.6** (V-AIC)**.** Let $\pi_t \in (0,1)$ be the probability for the ventilation fault to the time point $t$ and $l \in \{1, \ldots, 10\}$. Then the V-AIC models are given by,

$$\text{logit}(\pi_t) = \alpha + \sum_{i=0}^{l} x_{t-i}\beta_i$$

where $x_t = (\text{CO2}_t, \text{Pre}_t)$, $\alpha \in \mathbb{R}$ $\beta_i \in \mathbb{R}^2$ for all $i = 0, 1, \ldots, l$.

The tables 10.5 and 10.6 show the results of the V-AIC-RH models in Room B and Room A. On the measured data, one must again choose between high sensitivity and high specificity for each model. The results for the V-AIC model can be seen in the tables 10.7 and 10.8. The question of the best model can only be answered subjectively. Since sensitivity is considered more important than specificity in our case, the models with a maximum time lag between seven and nine without relative humidity seem like a good choice.

| | Sim_Sensitivity | Sim_Specificity | RoomB_Sensitivity | RoomB_Specificity |
|---|---|---|---|---|
| Lag_1 | 0.9722 | 0.7624 | 1.0000 | 0.3462 |
| Lag_2 | 0.9722 | 0.7309 | 0.9714 | 0.7729 |
| Lag_3 | 0.9444 | 0.7319 | 0.9143 | 0.7851 |
| Lag_4 | 0.9444 | 0.7315 | 0.9143 | 0.7848 |
| Lag_5 | 0.9722 | 0.7525 | 0.8571 | 0.7824 |
| Lag_6 | 0.9444 | 0.8120 | 0.7143 | 0.7861 |
| Lag_7 | 0.8889 | 0.8417 | 0.6857 | 0.8232 |
| Lag_8 | 1.0000 | 0.7829 | 0.8571 | 0.7471 |
| Lag_9 | 0.9167 | 0.8498 | 0.7143 | 0.7609 |
| Lag_10 | 0.8611 | 0.8668 | 0.7429 | 0.7997 |

Table 10.5: Results: V-AIC-RH - Room B

|         | Sim_Sensitivity | Sim_Specificity | RoomA_Sensitivity | RoomA_Specificity |
|---------|-----------------|-----------------|-------------------|-------------------|
| Lag_1   | 0.9722          | 0.7624          | 1.0000            | 0.2821            |
| Lag_2   | 0.9722          | 0.7309          | 0.9143            | 0.6903            |
| Lag_3   | 0.9444          | 0.7319          | 0.7714            | 0.7783            |
| Lag_4   | 0.9444          | 0.7315          | 0.7714            | 0.7783            |
| Lag_5   | 0.9722          | 0.7525          | 0.6571            | 0.7985            |
| Lag_6   | 0.9444          | 0.8120          | 0.6857            | 0.8053            |
| Lag_7   | 0.8889          | 0.8417          | 0.6571            | 0.8370            |
| Lag_8   | 1.0000          | 0.7829          | 0.6857            | 0.7452            |
| Lag_9   | 0.9167          | 0.8498          | 0.6857            | 0.7947            |
| Lag_10  | 0.8611          | 0.8668          | 0.6571            | 0.8281            |

Table 10.6: Results: V-AIC-RH - Room A

|         | Sim_Sensitivity | Sim_Specificity | RoomB_Sensitivity | RoomB_Specificity |
|---------|-----------------|-----------------|-------------------|-------------------|
| Lag_1   | 0.9722          | 0.7624          | 1.0000            | 0.3462            |
| Lag_2   | 0.9722          | 0.8683          | 0.9714            | 0.4151            |
| Lag_3   | 0.9167          | 0.8723          | 0.9714            | 0.4341            |
| Lag_4   | 0.8889          | 0.8722          | 0.9714            | 0.4612            |
| Lag_5   | 0.6389          | 0.9303          | 0.9714            | 0.8689            |
| Lag_6   | 0.7500          | 0.9017          | 0.7143            | 0.8268            |
| Lag_7   | 0.6944          | 0.9558          | 0.6286            | 0.9565            |
| Lag_8   | 0.6944          | 0.9257          | 0.6571            | 0.9549            |
| Lag_9   | 0.6667          | 0.9542          | 0.7429            | 0.9300            |
| Lag_10  | 0.7778          | 0.8481          | 0.7714            | 0.7368            |

Table 10.7: Results: Stepwise AIC - V-Fault - Room B

|         | Sim_Sensitivity | Sim_Specificity | RoomA_Sensitivity | RoomA_Specificity |
|---------|-----------------|-----------------|-------------------|-------------------|
| Lag_1   | 0.9722          | 0.7624          | 1.0000            | 0.2821            |
| Lag_2   | 0.9722          | 0.8683          | 1.0000            | 0.3338            |
| Lag_3   | 0.9167          | 0.8723          | 1.0000            | 0.3854            |
| Lag_4   | 0.8889          | 0.8722          | 1.0000            | 0.3951            |
| Lag_5   | 0.6389          | 0.9303          | 0.8286            | 0.8372            |
| Lag_6   | 0.7500          | 0.9017          | 0.6857            | 0.8384            |
| Lag_7   | 0.6944          | 0.9558          | 0.6857            | 0.9304            |
| Lag_8   | 0.6944          | 0.9257          | 0.6857            | 0.9172            |
| Lag_9   | 0.6667          | 0.9542          | 0.6857            | 0.9475            |
| Lag_10  | 0.7778          | 0.8481          | 0.6857            | 0.7632            |

Table 10.8: Results: Stepwise AIC - V-Fault - Room A

### V-Fault overview

Table 10.9 and 10.10 show another overview of the models for the ventilation fan. The models by the Stepwise AIC are not represented for reasons of clarity. Overall, it seems that a choice must be made between high sensitivity and high specificity. As mentioned before, for this problem, the sensitivity is more relevant for practical use because a system that decides too often on faults is no longer taken seriously by the user. Therefore the V-Categorical model seems like a suitable choice.

|      | Sim_Sensitivity | Sim_Specificity | RoomB_Sensitivity | RoomB_Specificity |
|------|-----------------|-----------------|-------------------|-------------------|
| CO2  | 0.97            | 0.76            | 1.00              | 0.35              |
| RH   | 0.58            | 0.96            | 0.49              | 0.98              |
| Cat. | 0.56            | 0.99            | 0.51              | 0.99              |
| CPC  | 0.94            | 0.96            | 0.91              | 0.50              |

Table 10.9: Ventilation fault results overview: Room B

|      | Sim_Sensitivity | Sim_Specificity | RoomA_Sensitivity | RoomA_Specificity |
|------|-----------------|-----------------|-------------------|-------------------|
| CO2  | 0.97            | 0.76            | 1.00              | 0.28              |
| RH   | 0.58            | 0.96            | 0.37              | 0.98              |
| Cat. | 0.56            | 0.99            | 0.49              | 0.98              |
| CPC  | 0.94            | 0.96            | 0.91              | 0.45              |

Table 10.10: Ventilation fault results overview: Room A

## 10.2 Heating fault

In this section, the fault on the heating radiator valve, or in short the heating fault, is investigated. As the name suggests, this fault primarily affects the total heating power (THP) and thus also the room temperature (RT). Therefore figure 10.15 shows the room temperature during the simulation and in Room B, while figure 10.16 shows the room temperature of Room A and B. As before, the time points with fault are orange, and the time points without being green. Unlike the ventilation fault, a suitable model is derived step by step on the simulation data, and this model is then applied to the measured data.
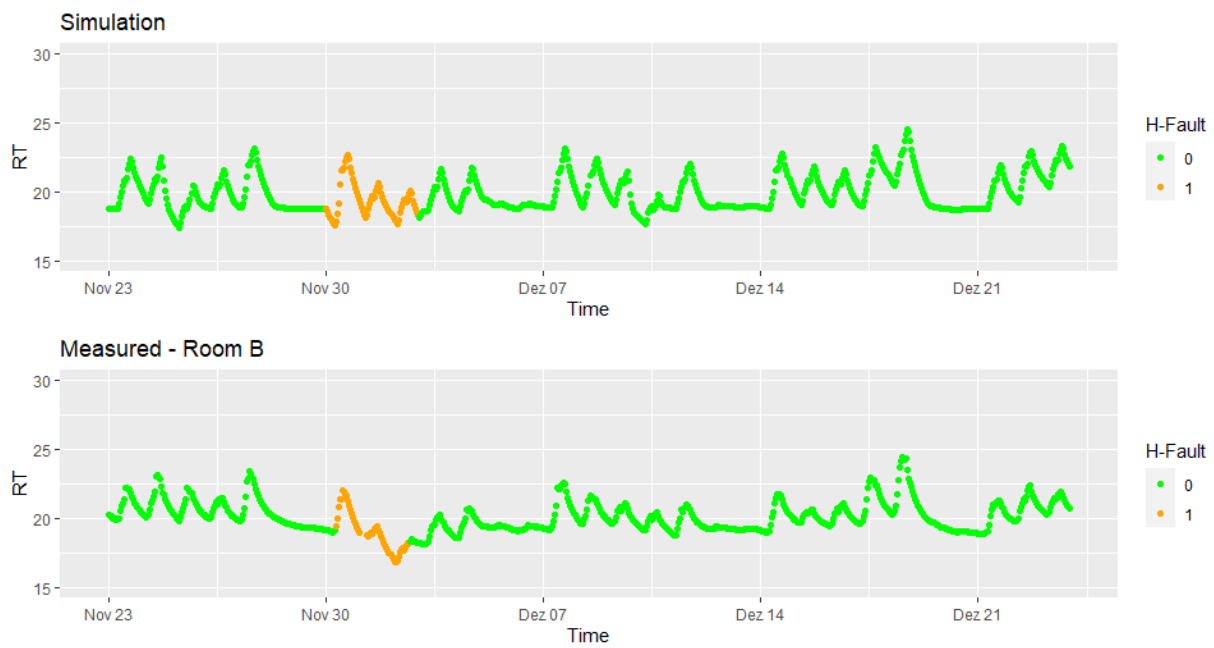
Figure 10.15: Heating fault: Room temperature simulation and room temperature in Room B
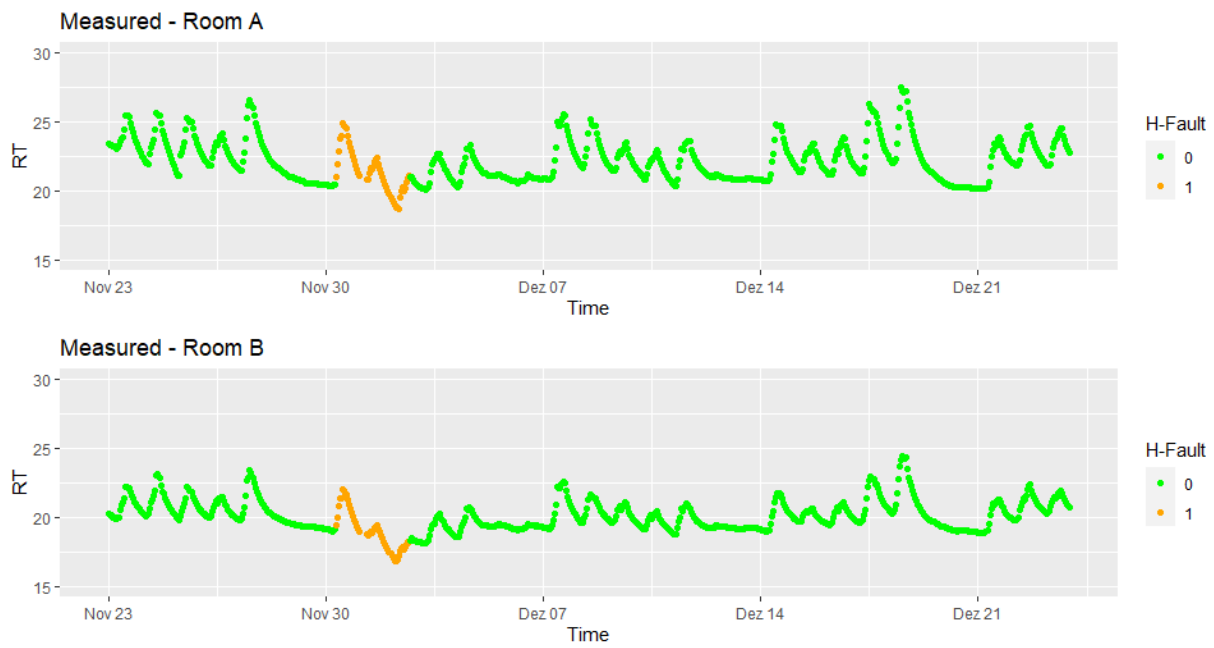


Figure 10.16: Heating fault: Room temperature Room A and room temperature in Room B
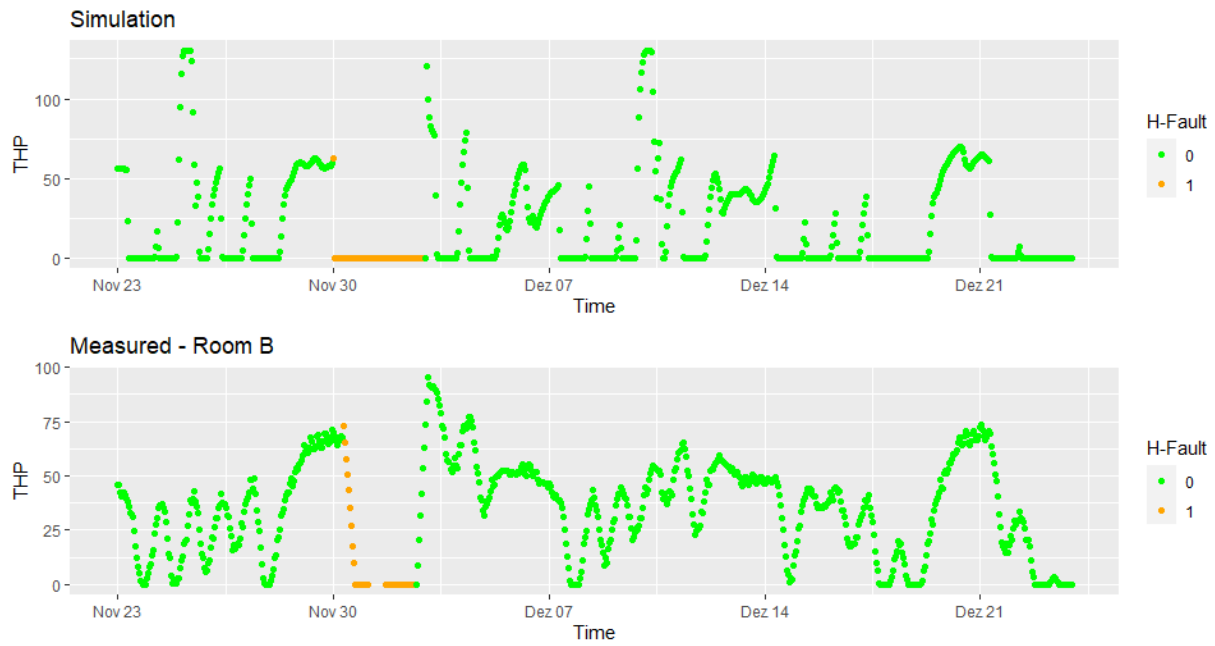
Figure 10.17: Heating fault: THP simulation and THP in Room B
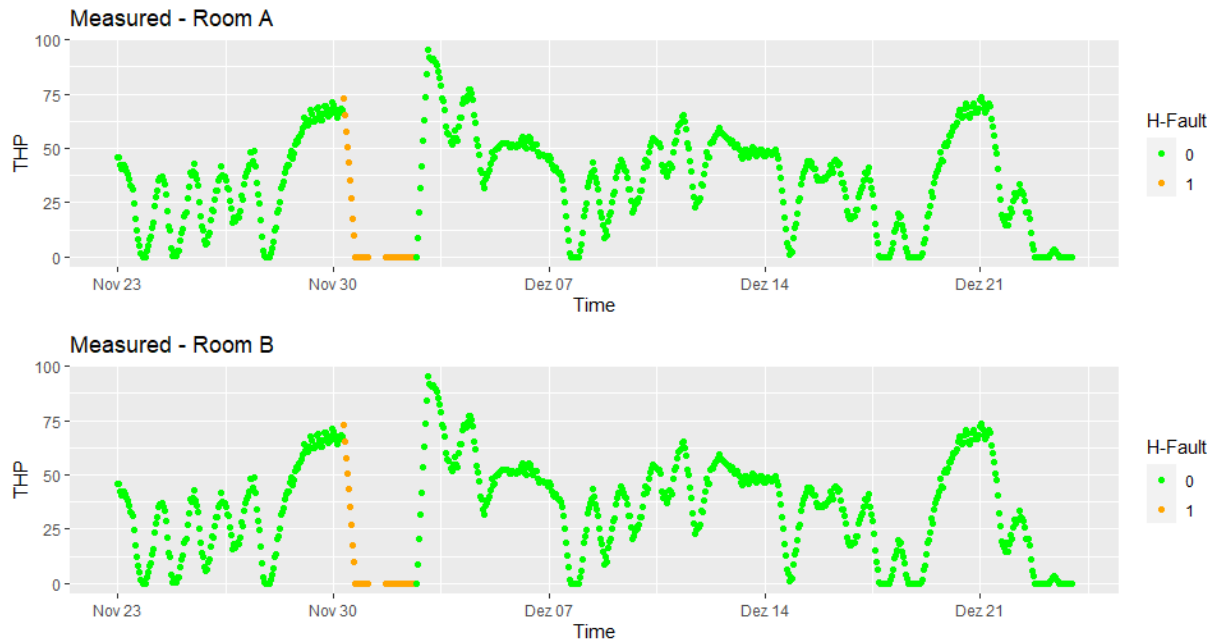


Figure 10.18: Heating fault: THP in Room A and THP in Room B

## H-RT
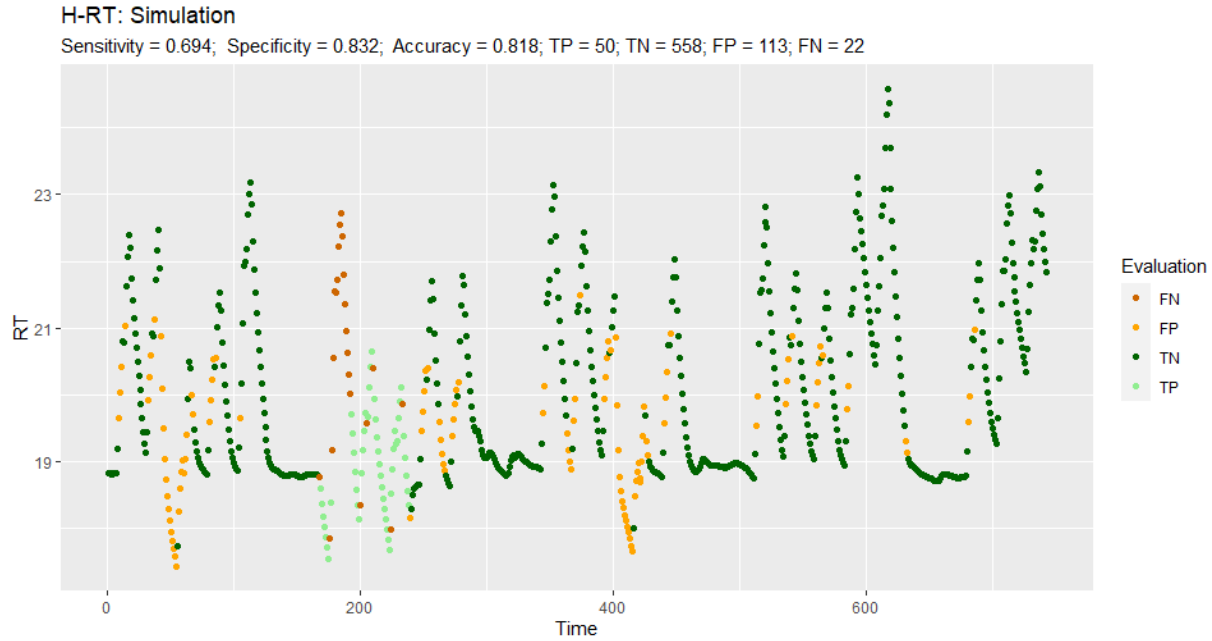


Figure 10.19: Results H-RT: Simulation

The first heating fault model focuses on the room temperature. Therefore it is called H-RT and is defined in definition 10.2.1. Figure 10.19 shows the result on the simulation data. The correct fault period is detected, but there are also many misclassifications.

---

**Definition 10.2.1** (H-RT). Let $\pi_t \in (0,1)$ be the probability for the heating fault to the time point $t$. Then the H-RT model is given by,

$$\text{logit}(\pi_t) = \beta_0 + \beta_1 \text{RT}_t + \beta_2 \text{RT}_{t-1} + \beta_3 \text{Pre}_{t-1}$$

where $\beta_i \in \mathbb{R}$ for all $i = 0, 1, 2, 3$.

---

## H-THP



**H-THP: Simulation**
Sensitivity = 0.986;  Specificity = 0.535;  Accuracy = 0.579; TP = 71; TN = 359; FP = 312; FN = 1
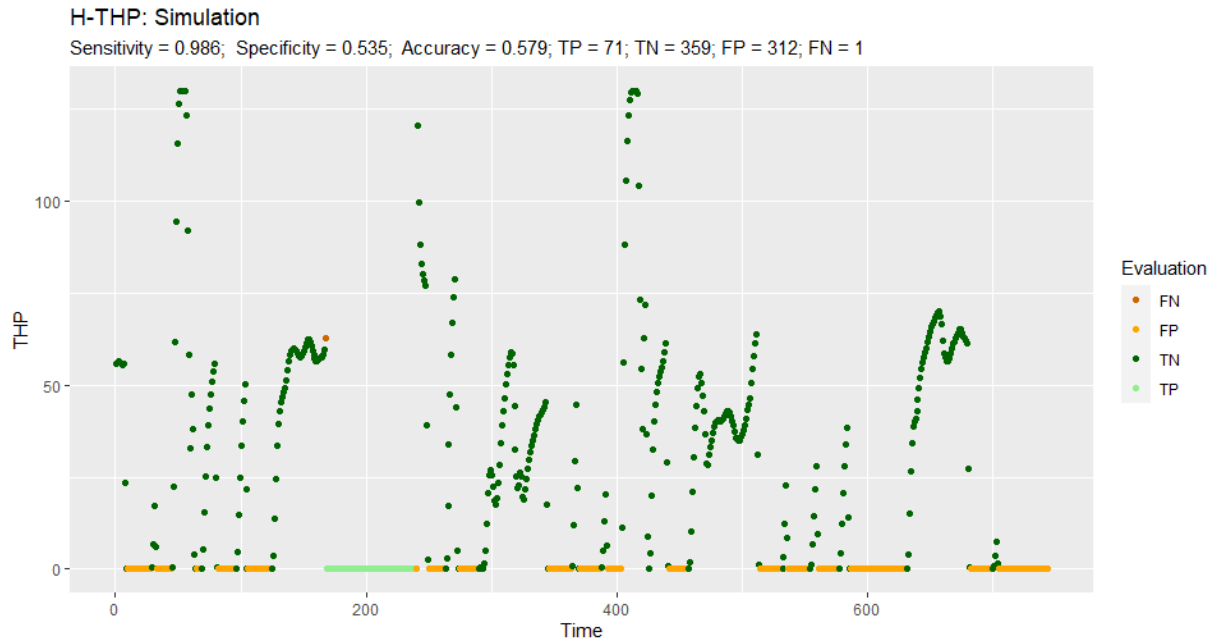
Figure 10.20: Results H-THP: Simulation

In the following approach, the total heating power is used instead of the room temperature see definition 10.2.2. Figure 10.20 shows that this model assumes a fault as soon as the heating power is at zero. However, there are also times when the heating power is zero, although there is no fault.

---

**Definition 10.2.2** (H-THP)**.** Let $\pi_t \in (0, 1)$ be the probability for the heating fault to the time point $t$. Then the H-THP model is given by,

$$\text{logit}(\pi_t) = \beta_0 + \beta_1 \text{THP}_t$$

where $\beta_i \in \mathbb{R}$ for all $i = 0, 1$.
(Adding time lags or the presence has barely changed the model.)

## H-Climate



**H-Climate: Simulation**
Sensitivity = 0.958; Specificity = 0.872; Accuracy = 0.88; TP = 69; TN = 585; FP = 86; FN = 3
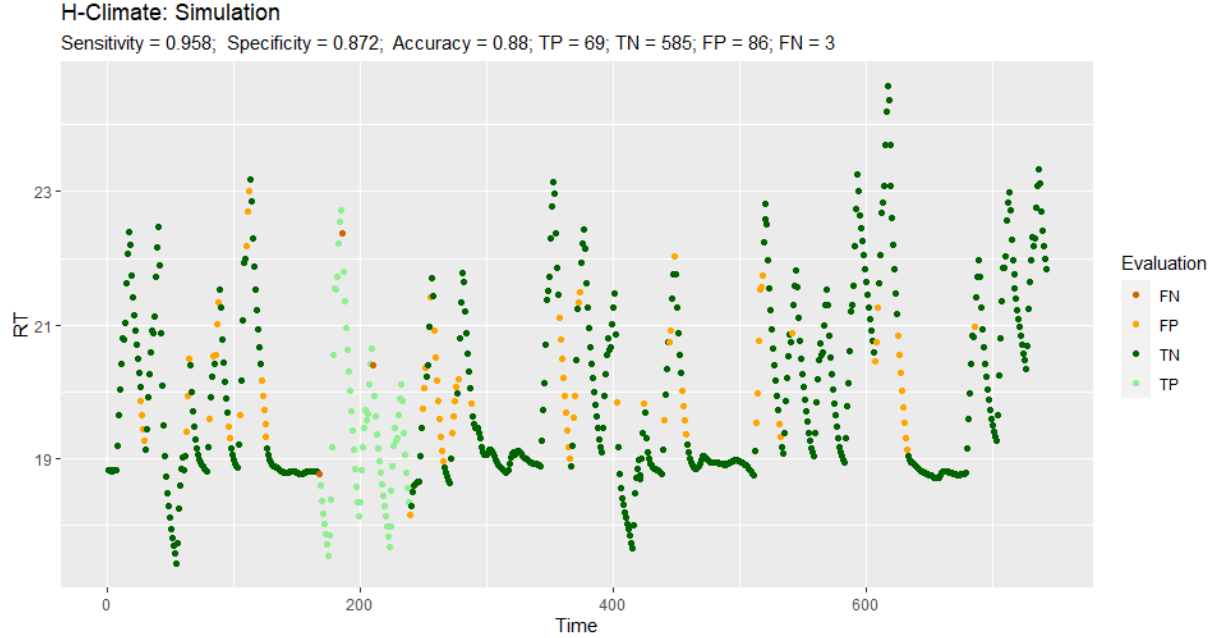
Figure 10.21: Results H-Climate: Simulation

The room temperature and heating power depending on the climate, which was not considered in the previous models. This changes with the H-Climate model from definition 10.2.3. This model has a sensitivity over 95% and specificity over 87% on the simulation data. The results are shown in figure 10.21. Therefore it seems like this model can detect the heating fault. The only question is how many hours of time delay should be considered and which coefficients should be included in the model. This question will be addressed in the following H-AIC model.

---

**Definition 10.2.3** (H-Climate). Let $\pi_t \in (0, 1)$ be the probability for the heating fault to the time point $t$. Then the H-THP model is given by,

$$\text{logit}(\pi_t) = \alpha + x_t \beta_1 + x_{t-1} \beta_2$$

where $x_t = (\text{RT}_t, \text{THP}_t, \text{OT}_t, \text{difR}_t, \text{dirR}_t, \text{Pre}_t)$, $\alpha \in \mathbb{R}$, $\beta_i \in \mathbb{R}^6$ for all $i = 1, 2$.

---

## Stepwise AIC

As with the ventilation fault, the forwarded stepwise AIC is used to find an appropriate model. The models defined in definition 10.2.4 would be the model if all covariates were included. The results are shown in table 10.11 and 10.12.

---

**Definition 10.2.4** (H-AIC)**.** Let $\pi_t \in (0, 1)$ be the probability for the heating fault to the time point $t$ and $l \in \{1, \ldots, 10\}$. Then the H-AIC models are given by,

$$\text{logit}(\pi_t) = \alpha + \sum_{i=0}^{l} x_{t-i}\beta_i$$

where $x_t = (\text{RT}_t, \text{THP}_t, \text{OT}_t, \text{difR}_t, \text{dirR}_t, \text{Pre}_t)$, $\alpha \in \mathbb{R}$, $\beta_i \in \mathbb{R}^6$ for all $i = 0, 1, \ldots, l$.

---

The question of the best model is very subjective. However, it is a good sign that the specificity is very high in each case. Interestingly, the specificity in the room with real users (Room A) is even better than in the room with artificial users (Room B). However, this does not apply to the sensitivity. As an example, figure 10.11 shows the results of the model with a time lag of five for Room B and figure 10.12 for Room A.

| | Sim_Sensitivity | Sim_Specificity | RoomB_Sensitivity | RoomB_Specificity |
|---|---|---|---|---|
| Lag_1 | 0.9583 | 0.8733 | 0.8372 | 0.9621 |
| Lag_2 | 0.9583 | 0.8761 | 0.8293 | 0.9591 |
| Lag_3 | 0.9306 | 0.9088 | 0.8000 | 0.9795 |
| Lag_4 | 0.9444 | 0.9311 | 0.7949 | 0.9824 |
| Lag_5 | 0.9722 | 0.9295 | 0.7895 | 0.9839 |
| Lag_6 | 0.9722 | 0.9369 | 0.6486 | 0.9794 |
| Lag_7 | 0.9861 | 0.9398 | 0.6389 | 0.9794 |
| Lag_8 | 0.9722 | 0.9518 | 0.6389 | 0.9794 |
| Lag_9 | 0.9861 | 0.9879 | 0.6000 | 0.9838 |
| Lag_10 | 0.9861 | 0.9849 | 0.6364 | 0.9897 |

Table 10.11: Results: Stepwise AIC - H-Fault - Room B

|  | Sim_Sensitivity | Sim_Specificity | RoomA_Sensitivity | RoomA_Specificity |
|---|---|---|---|---|
| Lag_1 | 0.9583 | 0.8733 | 0.6047 | 0.9942 |
| Lag_2 | 0.9583 | 0.8761 | 0.6098 | 0.9956 |
| Lag_3 | 0.9306 | 0.9088 | 0.4750 | 0.9971 |
| Lag_4 | 0.9444 | 0.9311 | 0.4615 | 0.9971 |
| Lag_5 | 0.9722 | 0.9295 | 0.4737 | 0.9956 |
| Lag_6 | 0.9722 | 0.9369 | 0.4595 | 0.9941 |
| Lag_7 | 0.9861 | 0.9398 | 0.4444 | 0.9912 |
| Lag_8 | 0.9722 | 0.9518 | 0.4167 | 0.9926 |
| Lag_9 | 0.9861 | 0.9879 | 0.3143 | 0.9912 |
| Lag_10 | 0.9861 | 0.9849 | 0.3030 | 0.9970 |

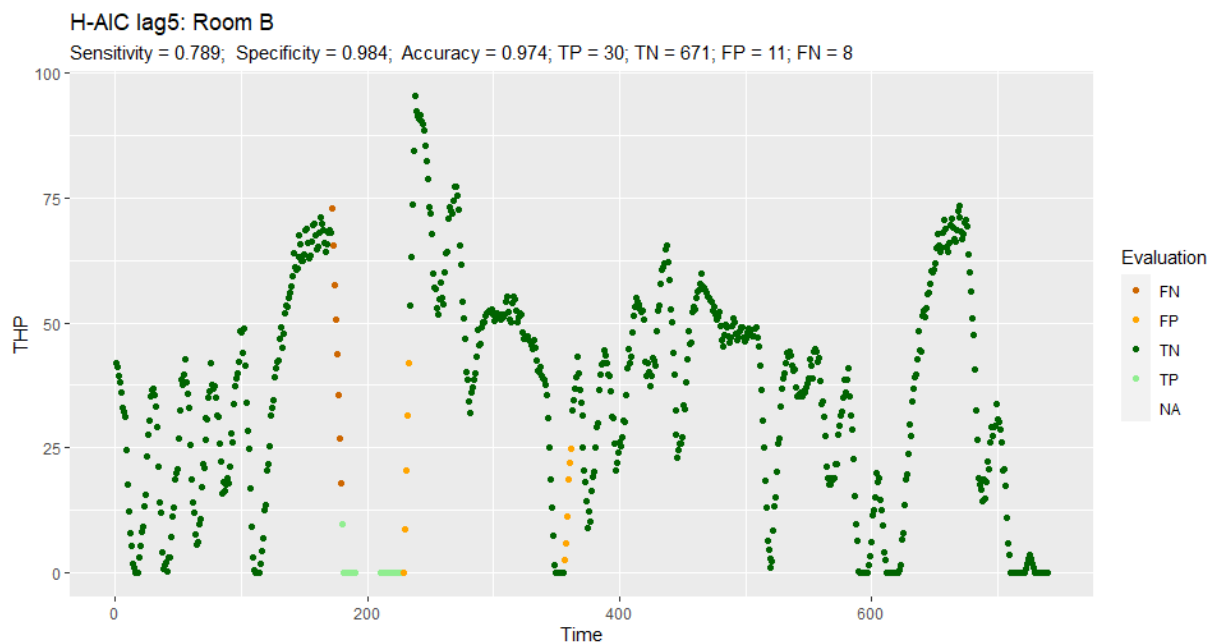Table 10.12: Results: Stepwise AIC - H-Fault - Room A



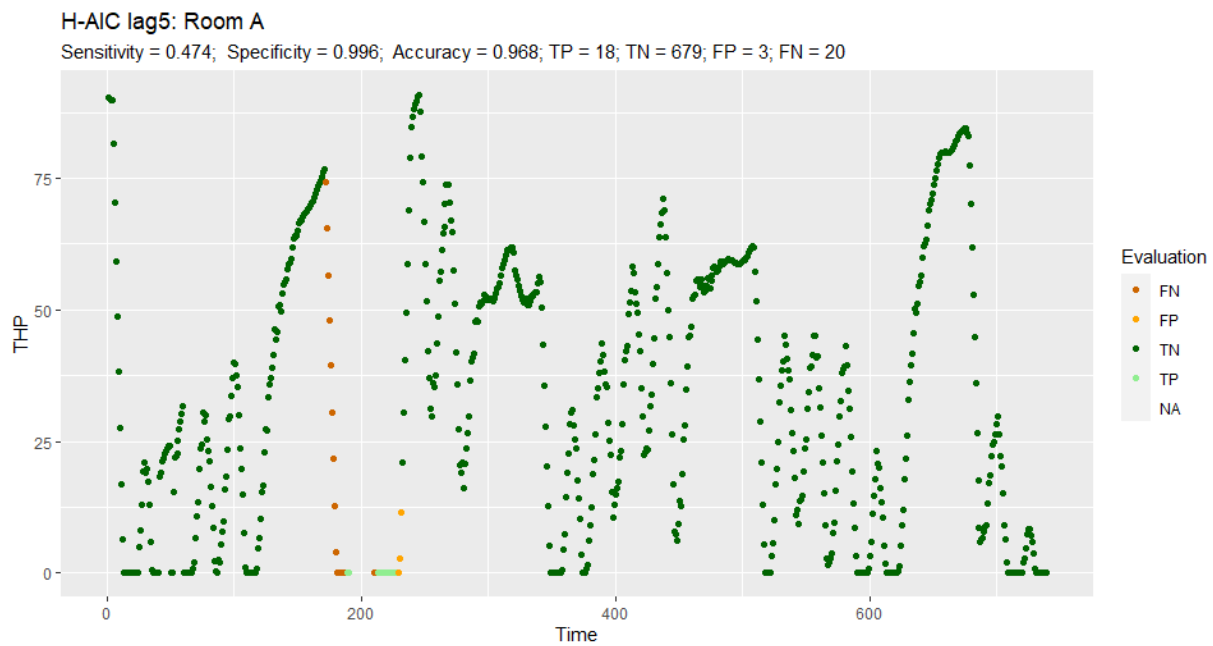Figure 10.22: Results H-AIC: Room B with a with a maximum time lag of five hours.

Figure 10.23: Results H-AIC: Room A with a with a maximum time lag of five hours.

## 10.3   Bypass fault

Finally, the bypass fault is examined. Similar to the heating fault, this one affects the room temperature and the heating power. For this reason, the 10.24 and 10.25 figures show the room temperatures and the figures 10.26 and 10.27 show the total heating powers. However, hardly any difference between fault and fault-free can be seen on the graphs at room temperature. The difference in heating power is only visible in the simulation (figure 10.26), but not in the measurements. This difference between simulation and measured data makes it hard to detect the fault on the measured data using the simulation data. Nevertheless, an attempt is made to detect the fault. For this, the forwarded stepwise AIC is used again to try out ten different models. The full models are defined in definition 10.3.1. The results given in table 10.13 and 10.14. For Room B, the sensitivity is usually just under 50% and the specificity about 80%, whereas for Room A, the specificity is sometimes over 90%. In these models, however, the sensitivity does not even reach 25%. Overall it seems like these models are not suitable to detect the fault on the measured data.

---

**Definition 10.3.1** (B-AIC)**.** Let $\pi_t \in (0, 1)$ be the probability for the bypass fault to the time point $t$ and $l \in \{1, \ldots, 10\}$. Then the B-AIC models are given by,

$$\text{logit}(\pi_t) = \alpha + \sum_{i=0}^{l} x_{t-i}\beta_i$$

where $x_t = (\text{RT}_t, \text{THP}_t, \text{OT}_t, \text{difR}_t, \text{dirR}_t, \text{Pre}_t)$, $\alpha \in \mathbb{R}$, $\beta_i \in \mathbb{R}^6$ for all $i = 0, 1, \ldots, l$.

---

|  | Sim_Sensitivity | Sim_Specificity | RoomB_Sensitivity | RoomB_Specificity |
|---|---|---|---|---|
| Lag_1 | 0.8333 | 0.9871 | 0.3929 | 0.8009 |
| Lag_2 | 0.8750 | 0.9957 | 0.3393 | 0.8405 |
| Lag_3 | 0.8750 | 0.9942 | 0.3036 | 0.8458 |
| Lag_4 | 0.8750 | 0.9986 | 0.3571 | 0.8486 |
| Lag_5 | 0.9375 | 0.9465 | 0.4643 | 0.7771 |
| Lag_6 | 0.9375 | 0.9551 | 0.4643 | 0.7961 |
| Lag_7 | 0.9375 | 0.9579 | 0.4643 | 0.7818 |
| Lag_8 | 0.9583 | 0.9549 | 0.4464 | 0.7872 |
| Lag_9 | 1.0000 | 1.0000 | 0.4821 | 0.8095 |
| Lag_10 | 1.0000 | 1.0000 | 0.4821 | 0.8049 |

Table 10.13: Results: Stepwise AIC - B-Fault - Room B

| | Sim_Sensitivity | Sim_Specificity | RoomA_Sensitivity | RoomA_Specificity |
|---|---|---|---|---|
| Lag_1 | 0.8333 | 0.9871 | 0.2143 | 0.9212 |
| Lag_2 | 0.8750 | 0.9957 | 0.1964 | 0.9165 |
| Lag_3 | 0.8750 | 0.9942 | 0.2143 | 0.9222 |
| Lag_4 | 0.8750 | 0.9986 | 0.1607 | 0.9280 |
| Lag_5 | 0.9375 | 0.9465 | 0.1964 | 0.9157 |
| Lag_6 | 0.9375 | 0.9551 | 0.2143 | 0.9063 |
| Lag_7 | 0.9375 | 0.9579 | 0.2857 | 0.8879 |
| Lag_8 | 0.9583 | 0.9549 | 0.2857 | 0.8967 |
| Lag_9 | 1.0000 | 1.0000 | 0.3750 | 0.8902 |
| Lag_10 | 1.0000 | 1.0000 | 0.4107 | 0.8476 |

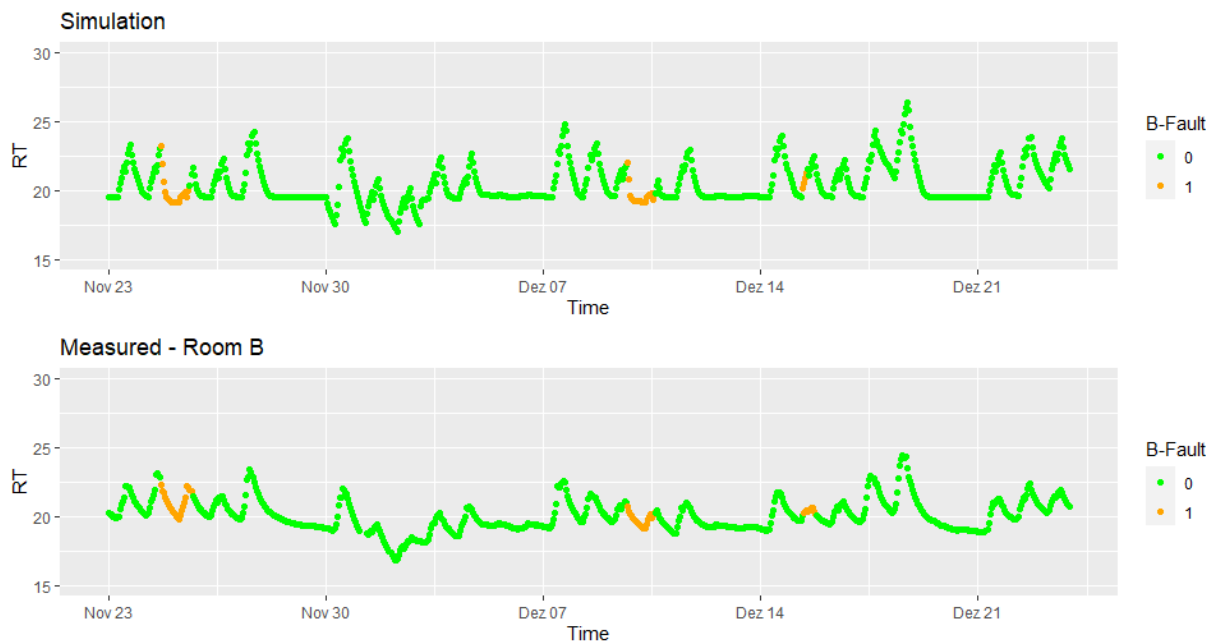Table 10.14: Results: Stepwise AIC - B-Fault - Room A



Figure 10.24: Bypass fault: Room temperature simulation and room temperature in Room B
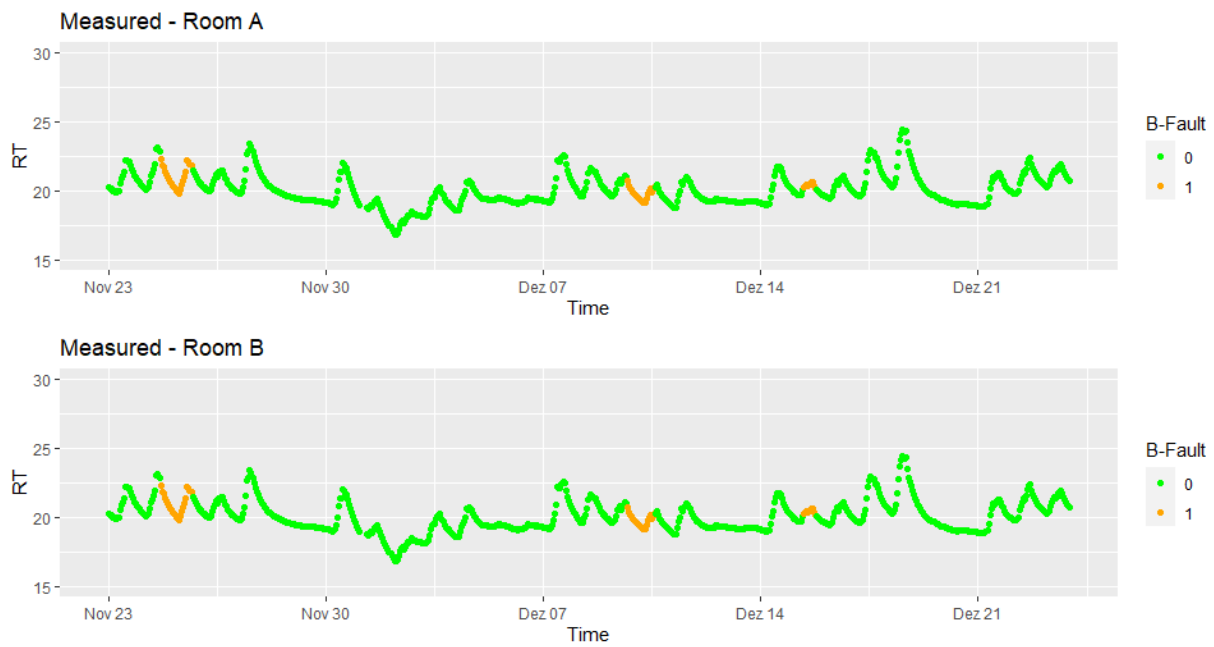
Figure 10.25: Bypass fault: Room temperature in Room A and room temperature in Room B
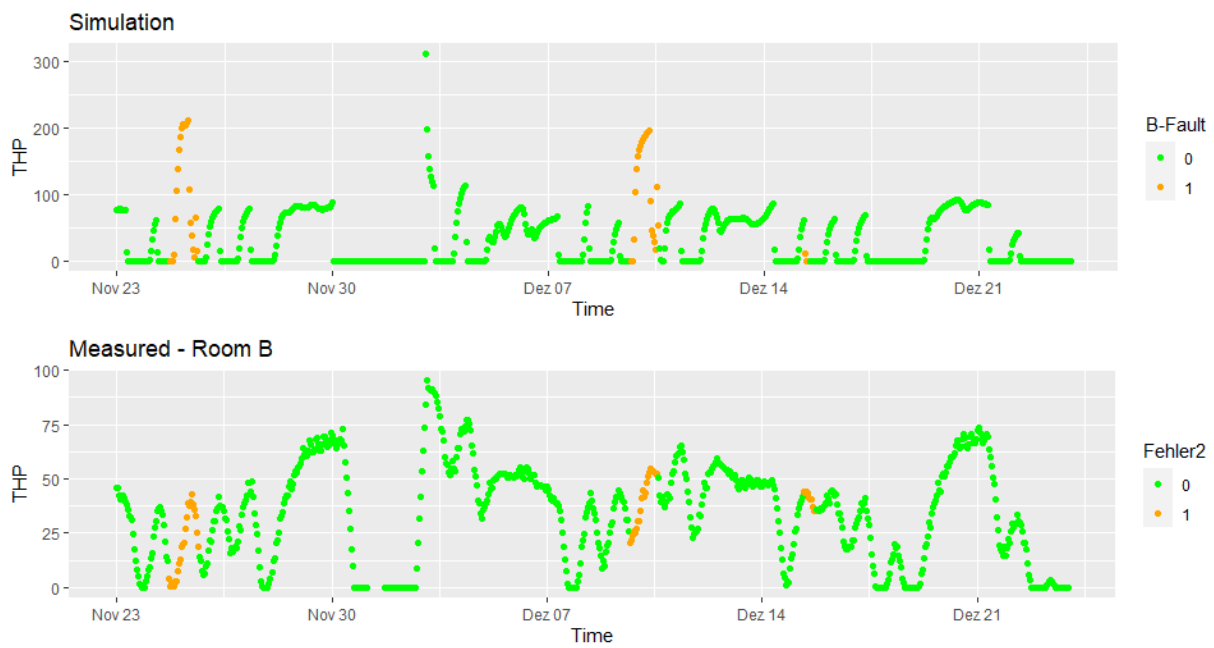


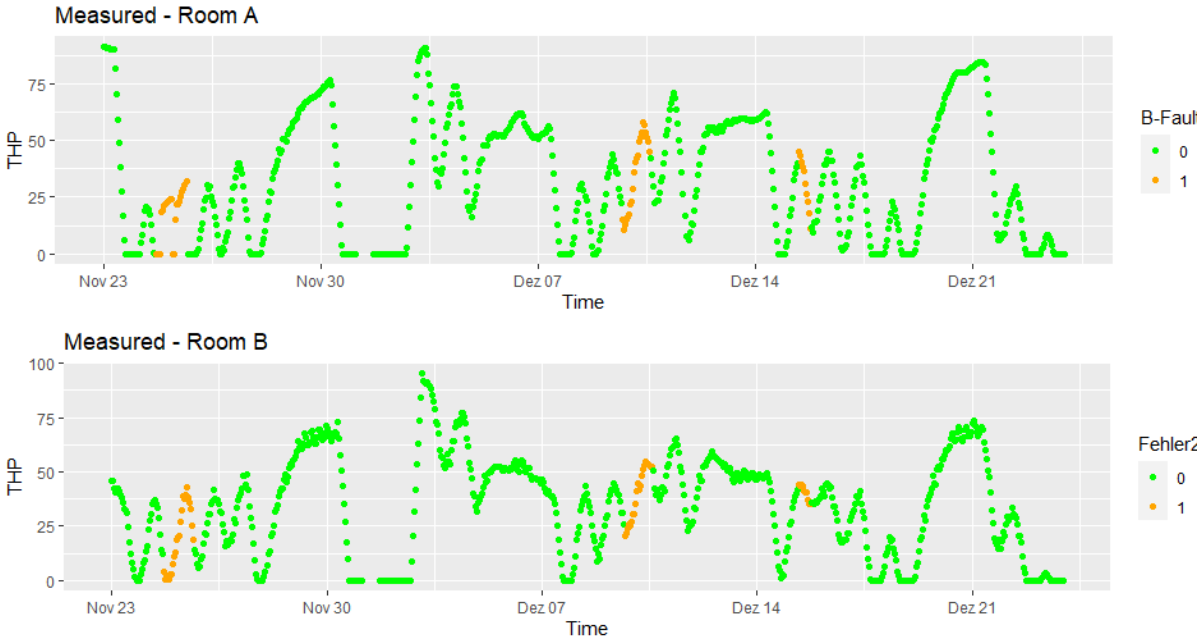Figure 10.26: Bypass fault: THP simulation and THP in Room B

Figure 10.27: Bypass fault: THP in Room A and THP in Room B

# 10.4   Comparison with different simulations

For the comparison with differences simulations, one model is used for each of the three faults. These models are the V-Categorial for the ventilation fault and the models via stepwise AIC for the other two faults.

## Connection with model-based comparison

Comparing the results of the fault detection with the results of the model-based comparison. See the tables 9.18 and 9.19. Then it is noticeable that the Room A results are worse than the Room B results in each case except for the Berlin data set. If the Berlin data set detects the heating fault, then a better result is obtained for Room A than for Room B. Also, for the similarity in the change of the total heating power, the similarity for Room A is higher with the Berlin data set. Therefore, however, it does not necessarily have to be a connection. The Vienna data set also has a higher similarity for Room A, but the results for fault detection are significantly worse.

Overall, it is difficult to draw a connection between model-based similarity and accuracy in fault detection. On the one hand, because the fault periods for the similarity were ignored. If these were not ignored, there would be problems for the realization in practice since the fault periods are unknown. On the other hand, there are still many ambiguities in model-based comparison. For example, whether the parameters should be weighted differently and, if so, how. Another question would be how to find the most suitable models for the comparison.

However, the model-based comparison still seems to be helpful. When using it, it should be clear what the limits are, i.e., the results should not be overinterpreted. Nevertheless, it is feasible to check how close the simulation is to reality.

## Room B results

The results for Room B are given in table 10.15, 10.16 and 10.17. The models for the ventilation fault all seem to be relatively identical, which is to be expected since time shifts and climate data should not affect the model. The models for the ventilation fault all appear to be relatively identical, which is to be expected since time shifts and climate data should not affect the model.

The models for the heating fault even seem to be more appropriate overall. However, there are different results here depending on the simulation. Since heating performance depends on climate data, this result is not surprising. It is even a good sign that the result is still usable with external climate data (Berlin/Vienna).

All the models for the bypass fault are unusable. This shows how important it is that the effects of the fault in the simulation are as close as possible to reality.

| Data | Sim_Sensitivity | Sim_Specificity | RoomB_Sensitivity | RoomB_Specificity |
|------|-----------------|-----------------|-------------------|-------------------|
| Original | 0.55 | 0.99 | 0.51 | 0.99 |
| Shift1 | 0.56 | 1.00 | 0.51 | 0.99 |
| Shift2 | 0.56 | 1.00 | 0.51 | 0.99 |
| Shift3 | 0.53 | 1.00 | 0.51 | 0.99 |
| Vienna | 0.53 | 1.00 | 0.51 | 0.99 |
| Berlin | 0.53 | 1.00 | 0.51 | 0.99 |

Table 10.15: V-Categorical model: Different simulations vs. Room B

| Data | Sim_Sensitivity | Sim_Specificity | RoomB_Sensitivity | RoomB_Specificity |
|------|-----------------|-----------------|-------------------|-------------------|
| Original | 0.93 | 0.91 | 0.80 | 0.98 |
| Shift1 | 0.93 | 0.92 | 0.65 | 1.00 |
| Shift2 | 0.85 | 0.97 | 0.68 | 1.00 |
| Shift3 | 0.90 | 0.98 | 0.60 | 1.00 |
| Vienna | 0.94 | 0.99 | 0.62 | 1.00 |
| Berlin | 0.94 | 0.96 | 0.57 | 0.98 |

Table 10.16: H-Fault model: Different simulations vs. Room B

| Data | Sim_Sensitivity | Sim_Specificity | RoomB_Sensitivity | RoomB_Specificity |
|------|-----------------|-----------------|-------------------|-------------------|
| Original | 0.46 | 0.03 | 0.93 | 0.19 |
| Shift1 | 0.18 | 0.47 | 0.00 | 0.89 |
| Shift2 | 0.09 | 0.47 | 0.00 | 0.91 |
| Shift3 | 0.00 | 0.41 | 0.00 | 0.87 |
| Vienna | 0.00 | 0.31 | 0.00 | 0.87 |
| Berlin | 0.50 | 0.50 | 0.50 | 0.50 |

Table 10.17: B-Fault model: Different simulations vs. Room B

## Room A results

There are no notable differences in the detection of the ventilation and the bypass fault for Room A.

For the detection of the heating fault, the sensitivity has again significantly deteriorated compared to Room B. Interestingly, the Berlin data set is the one with the highest sensitivity, even though it had the lowest sensitivity in Room B.

| Data | Sim_Sensitivity | Sim_Specificity | RoomA_Sensitivity | RoomA_Specificity |
|---|---|---|---|---|
| Original | 0.55 | 0.99 | 0.49 | 0.98 |
| Shift1 | 0.56 | 1.00 | 0.49 | 0.98 |
| Shift2 | 0.56 | 1.00 | 0.49 | 0.98 |
| Shift3 | 0.53 | 1.00 | 0.49 | 0.98 |
| Vienna | 0.53 | 1.00 | 0.49 | 0.98 |
| Berlin | 0.53 | 1.00 | 0.49 | 0.98 |

Table 10.18: V-Categorical model: Different simulations vs. Room A

| Data | Sim_Sensitivity | Sim_Specificity | RoomA_Sensitivity | RoomA_Specificity |
|---|---|---|---|---|
| Original | 0.93 | 0.91 | 0.47 | 1.00 |
| Shift1 | 0.93 | 0.92 | 0.30 | 1.00 |
| Shift2 | 0.85 | 0.97 | 0.32 | 1.00 |
| Shift3 | 0.90 | 0.98 | 0.20 | 1.00 |
| Vienna | 0.94 | 0.99 | 0.28 | 1.00 |
| Berlin | 0.94 | 0.96 | 0.65 | 0.95 |

Table 10.19: H-Fault model: Different simulations vs. Room A

| Data | Sim_Sensitivity | Sim_Specificity | RoomA_Sensitivity | RoomA_Specificity |
|---|---|---|---|---|
| Original | 0.46 | 0.03 | 0.59 | 0.51 |
| Shift1 | 0.18 | 0.47 | 0.05 | 0.88 |
| Shift2 | 0.09 | 0.47 | 0.04 | 0.88 |
| Shift3 | 0.00 | 0.41 | 0.11 | 0.83 |
| Vienna | 0.00 | 0.31 | 0.09 | 0.88 |
| Berlin | 0.50 | 0.50 | 0.50 | 0.50 |

Table 10.20: B-Fault model: Different simulations vs. Room A

# Chapter 11

# Outlook

The previous chapters have shown the potential to use simulated data to identify actual problems in the building system. This chapter shows possibilities for the optimization of fault detection with simulated and measured data. These improvements can concern the measurements, the simulation, the methods for comparisons, and the fault detection.

## Fault Monitoring Set-Up

Optimization of the fault monitoring set-up can be achieved by initially measuring the building for about two weeks without faults to validate the simulation models better. In practice, an expert who checks the building for faults within this period can implement the fault-free period. After that, the statistical models take over the role of the expert. Another suggestion for optimization would be to extend the experiment by examining a whole year or buildings instead of individual rooms.

## Simulation

The results show that the simulation of real people needs to be improved. One possibility is first to measure two weeks without a fault. Then gradually optimize real users' simulation during this period by measuring the usual $CO_2$ and their heat emission and use these for the simulation. In addition, further studies should simulate more variants to check how detailed the simulation must be to detect faults successfully. One research question is: What is the effect if the room is simulated too large or too small compared to reality? So far, the studies used one-hour values, and the measurement interval should be increased in future studies, as time delays are easier to detect. Smaller time steps will allow a better check of the similarity between simulation and measurement. In addition, the faults should not only be activated or not activated but simulated on different levels, since the question is not only whether a fault is present but how intensively a fault is pronounced.

# Comparison

The following shows some adaptation possibilities for the comparisons between simulation and measurement. One approach could first check data-based if the distance between simulation and measurement is considerable. In the next step, a model-based check can then be made for these periods to see if there are only different events or problems in the simulation. With model-based comparison, there is still the question of how to find the optimal models. This thesis creates simple models with low time lags for the exogenous variables and AR(1) error terms. Future studies can use stepwise AIC or other techniques to test better predictive models than this study. However, it is unclear whether better prediction automatically means better similarity testing. The model-based comparison only tested if the impact of individual variables was identical in the simulation and the measurement. It would be interesting, by contrast, to use one-dimensional normalized parameters that indicate how similar the two data sets are. Research could use a stepwise AIC, an optimal simulation model, and search the measurement data. Then, it can be checked how high the agreement between these two models is. Another simple approach,already used indirectly in this thesis, counts how many parameters match in simulation and measurement match. This approach can be extended by weighting important parameters higher than less relevant parameters. To obtain the importance of the influencing variables, the machine learning method Random Forest can be used [Bre01]. These *importance* measures can also be used directly for comparison. Overall, the goal is to derive future metrics that indicate the similarity between simulated and real data between zero and one. A high similarity means that fault detection can use the simulation data without problems, and a low one suggests that these data are unsuitable.

# Fault detection

In addition to Logistic Regression, future studies can use other methods. The question arises whether specific methods are particularly well suited to be trained with simulated data. A similar question arises for setting the threshold. For example, techniques that value the sensitivity higher than the specificity are more critical for the application. However, it was not necessary for this thesis since the models still provided a very high sensitivity, but it could be helpful for research with other faults. New techniques can take into account that the data observed for each fault has fewer periods than without them. The same applies to the time series structure and the fact that the method initially detects many faults with difficulty, and the state is not normalized directly after correction. Another approach would be to combine different simulation data or predictive models. For this purpose, simulation data with high similarity to the measured data could be weighted higher than data with low similarity. More generally, techniques should be developed that directly transfer the results from the model-based comparison for modeling the faults.

# Chapter 12

# Summary

This thesis dealt with the applicability of statistical models from simulated data to real data using automated fault detection in buildings. The Technical University of Applied Sciences in Rosenheim performed the data collection experiments at the test facility LiGT. The study presented both data-based and model-based methods to compare simulation and measured data. Data-based methods evaluated the quality of estimated radiations. However, the focus of this thesis was on the model-based comparison methods. The methods compared six different simulation data sets with two measurement data sets. The simulation generated one data set with the same fault periods and climate data as the measurements. The other five simulations used periods with shifted faults, and two of these five simulations used external climate data. The experiment took place in the test facility with a real user and with an artificial user. It turned out that hardly any differences were found between the simulation data sets when compared with the measured data sets on a model basis. However, when comparing the two measurement data sets with real and artificial users, it was found that the simulation showed a higher similarity with artificial users in most cases. These are precisely the expected results since the model-based comparison was defined to compensate for time shifts and different climate data as long as the rest of the simulation is identical.

This thesis tried different models for each fault type to perform fault detection. In most cases, it is necessary to decide between high sensitivity and high specificity for the detection. This study found at least one model with a specificity close to one and a sensitivity high enough to detect faults for the ventilation and heating unit faults. For the model used to detect the fault in the ventilation unit, there was hardly any difference between the detection with artificial users and with real users. The lack of differences could be since a simple model with categorical variables was used here. The detection results on the heating fault were less precise for the data with real users than for the data with artificial users. However, the specificity was still high for both cases. The models for this fault were generated using stepwise AIC and has used time lags. These time lags could explain the difference between the data sets. The methods used were not able to model the Bypass successfully. The reason was due to the differences in the effect between simulation and measurement. The study shows how vital the correct implementation of the fault is.

It could be shown that it is difficult to draw a connection between model-based similarity and accuracy in fault detection. The results with real users are worse than with artificial users for both cases. In addition, for both fault detection and model-based comparison, there were few differences within a measurement data set. An exception was the fault in the heating system. There was one simulation data set for each measured data set, which gave different results than the others.

The results presented in this study show that there is potential in the approach of using simulation data to find the faults in buildings. Future studies should develop the concepts further in all areas considered, including the simulation, the similarity, and the fault detection itself.

# Appendix A

# Digital appendix

The digital appendix consists of the following contents:

- Data: Folder with the used data.

- Fkt: Folder with the R functions used.

- Plot: Folder with Figures.

- R Analytics: Folder with the R programs used to compare simulation and measurement.

- R Climate customize: Folder with the R programs to clean the climate data. Therefore these can be used for the simulation.

- R Datenanpassungen: Folder with R programs to smooth the measured data to hourly observations.

- R Fault detection: Folder with R programs to train and apply the models for fault detection.

- R Plots: Folder with R programs to generate Figures.

- R Saves: Storage location for the results.

- Masterarbeit.Rproj: R Project file, which combines the other programs.

- Results.R: Overview of the final results.

# Appendix B

# R-Packages

The analyses were implemented with the programming language R. The following table shows the packages used.

| Package | Reference | Usage |
| --- | --- | --- |
| Fable | [OWHW20] | Estimation of coefficients for Dynamic Regression Models. |
| ggplot2 | [Wic16] | Creation of figures. |
| grid | [R C20] | Extension to ggplot2. |
| gridExtra | [Aug17] | Extension to grid. |
| MASS | [VR02] | Stepwise AIC |
| pROC | [RTH+11] | Find thresholds for logistic regression. |
| rjags | [Plu19] | MCMC for Bayesian analysis. |
| tsibble | [WCH20] | Create time series tables. |
| xtable | [DSR+19] | Export Tables to LaTeX. |

Table B.1: R packages which were used for the analysis

# Appendix C

# Fault periods for simulation

Here are the periods for the faults in the different simulations.

| Fault | Start | End |
|---|---|---|
| Ventilation | 07.12.20 10:00 | 07.12.20 17:00 |
| Ventilation | 17.12.20 18:00 | 18.12.20 17:00 |
| Ventilation | 22.12.20 09:00 | 15.12.20 14:00 |
| Heating | 30.11.20 00:00 | 02.12.20 23:00 |
| Bypass | 24.11.20 17:00 | 25.11.20 17:00 |
| Bypass | 09.12.20 17:00 | 10.12.20 14:00 |
| Bypass | 15.12.20 08:00 | 15.12.20 12:00 |

Table C.1: Temporal distribution of the faults in the Original simulation.

| Fault | Start | End |
|---|---|---|
| Ventilation | 02.12.20 19:00 | 03.12.20 18:00 |
| Ventilation | 08.12.20 11:00 | 08.12.20 18:00 |
| Heating | 11.12.20 01:00 | 14.12.20 00:00 |
| Bypass | 17.12.20 18:00 | 18.12.20 15:00 |

Table C.2: Temporal distribution of the faults in the Shift1 simulation.

| Fault | Start | End |
|---|---|---|
| Ventilation | 01.12.20 11:00 | 01.12.20 18:00 |
| Ventilation | 14.12.20 19:00 | 15.12.20 18:00 |
| Heating | 17.11.20 01:00 | 30.11.20 00:00 |
| Bypass | 03.12.20 18:00 | 04.12.20 15:00 |

Table C.3: Temporal distribution of the faults in the Shift2 simulation.

| Fault | Start | End |
|---|---|---|
| Ventilation | 04.12.20 11:00 | 04.12.20 18:00 |
| Ventilation | 20.12.20 19:00 | 21.12.20 18:00 |
| Heating | 04.12.20 01:00 | 07.12.20 00:00 |
| Bypass | 13.12.20 18:00 | 14.12.20 15:00 |

Table C.4: Temporal distribution of the faults in the Shift3, Vienna and Berlin simulation.

# Bibliography

[Aka74]    H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[Ann21]    Annex 71. Iea ecb annex 71. building energy performance assessment based on in-situ measurements. Website, 2016-2021. `https://www.ecbcs.org/projects/project?AnnexID=71`.

[ASH01a]   ASHRAE IWEC. ASHRAE IWEC2 Weather File for Berlin Schönefeld. IDA-ICE, 2001. `http://www.equaonline.com/iceuser/ASHRAE_IWEC.html`.

[ASH01b]   ASHRAE IWEC. ASHRAE IWEC2 Weather File for Vienna Schwechat. IDA-ICE, 2001. `http://www.equaonline.com/iceuser/ASHRAE_IWEC.html`.

[Aug17]    Baptiste Auguie. *gridExtra: Miscellaneous Functions for "Grid" Graphics*, 2017. R package version 2.3.

[BAH10]    Kenneth P. Burnham, David R. Anderson, and Kathryn P. Huyvaert. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, 65(1):23–35, August 2010.

[BC16]     William M. Bolstad and James M. Curran. *Introduction to Bayesian Statistics*. John Wiley & Sons, New York, 3 edition, 2016.

[BJRL15]   George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time Series Analysis - Forecasting and Control*. John Wiley & Sons, New York, 2015.

[Bre01]    Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[Bun15]    Bundesministerium für Wirtschaft und Engerie. Energieeffizienzstrategie Gebäude. Wege zu einem nahezu klimaneutralen Gebäudebestand, 2015.

[Cop21]    Copernicus Atmosphere Monitoring Service (CAMS). Copernicus Atmosphere Monitoring Service (CAMS). Website, 2021. `http://www.soda-pro.com/web-services/radiation/cams-radiation-service`.

[Deu]      Deutscher Wetterdienst. Climate Data Center. `https://opendata.dwd.de/climate_environment/CDC/`.

[DIN]      DIN EN. DIN EN 13779:2007-09, Lüftung von Nichtwohngebäuden - Allgemeine Grundlagen und Anforderungen für Lüftungs- und Klimaanlagen und Raumkühlsysteme.

[DSR⁺19]   David B. Dahl, David Scott, Charles Roosen, Arni Magnusson, and Jonathan Swinton. *xtable: Export Tables to LaTeX or HTML*, 2019. R package version 1.8-4.

[EQU95]    EQUA. IDA-ICE. Website, 1995. `https://www.equa.se/en/ida-ice`.

[Equ04]    EquaSimulation AB. Validation of ida indoor climate and energy 4.0 build 4 with respect toansi/ashrae standard 140-2004, 2004.

[ET94]     Bradley Efron and R.J. Tibshirani. *An Introduction to the Bootstrap*. CRC Press, Boca Raton, Fla, 1994.

[FKL07]    Ludwig Fahrmeir, Thomas Kneib, and Stefan Lang. *Regression - Modelle, Methoden und Anwendungen*. Springer-Verlag, Berlin Heidelberg New York, 2007.

[FKV⁺18]   António Figueiredo, Jérôme Kämpf, Romeu Vicente, Rui Oliveira, and Tiago Silva. Comparison between monitored and simulated data using evolutionary algorithms: Reducing the performance gap in dynamic building simulation. *Journal of Building Engineering*, 17:96–106, 2018.

[Gio09]    Toni Giorgino. Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software*, 31(7):1–24, 2009.

[GJ08]     Clive Granger and Roselyne Joyeux. An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1:15 – 29, 06 2008.

[HA18]     Rob J Hyndman and George Athanasopoulos. *Forecasting: principles and practice*, 2nd edition, 2018. OTexts: Melbourne, Australia. `OTexts.com/fpp2`. Accessed on 12.05.2021.

[HAB⁺21]   Rob Hyndman, George Athanasopoulos, Christoph Bergmeir, Gabriel Caceres, Leanne Chhay, Mitchell O'Hara-Wild, Fotios Petropoulos, Slava Razbash, Earo Wang, and Farah Yasmeen. *forecast: Forecasting functions for time series and linear models*, 2021. R package version 8.15.

[Hal13]    Peter Hall. *The Bootstrap and Edgeworth Expansion*. Springer Science & Business Media, Berlin Heidelberg, 2013.

[HB13]     Leonhard Held and Daniel Sabanés Bové. *Applied Statistical Inference - Likelihood and Bayes*. Springer Science & Business Media, Berlin Heidelberg, 2013.

[HBH+20] Sebastian Höfer, Kostas E. Bekris, Ankur Handa, Juan Camilo Gamboa Higuera, Florian Golemo, Melissa Mozifian, Christopher G. Atkeson, Dieter Fox, Ken Goldberg, John Leonard, C. Karen Liu, Jan Peters, Shuran Song, Peter Welinder, and Martha White. Perspectives on sim2real transfer for robotics: A summary of the R: SS 2020 workshop. *CoRR*, abs/2012.03806, 2020.

[HM82]     J A Hanley and B J McNeil. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology*, 143(1):29–36, 1982. PMID: 7063747.

[HMS06]    Wolfgang Karl Härdle, Marlene Müller, and Axel Sperlich, Stefan Werwatz. *Nonparametric and Semiparametric Models*. Springer, 01 2006.

[HTF09]    Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference and prediction*. Springer, 2 edition, 2009.

[JWHT14] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning - with Applications in R*. Springer New York, Berlin-Heidelberg, 2014.

[Kle95]    Jack P.C. Kleijnen. Statistical validation of simulation models. *European Journal of Operational Research*, 87(1):21–34, 1995.

[Kri07]    Hans-Peter Kriegel. ähnlichkeitssuche in zeitreihen. Website, 2007. `https://www.dbs.ifi.lmu.de/Lehre/STMD/SS2009/Skript/Kap04-Time_Series.pdf`.

[KZ03]     Sven Kropf and Gerhard Zweifel. Validation of the building simulation program ida-ice according to cen 13791 "thermal performance of buildings - calculation of internal temperatures of a room in summer without mechanical cooling - general criteria and validation procedures", 2003.

[Lee12]    Peter M. Lee. *Bayesian Statistics*. John Wiley & Sons, Incorporated, New Jersey, 4 edition, 2012.

[LWL+19] B. Luo, H. Wang, H. Liu, B. Li, and F. Peng. Early fault detection of machine tools based on deep learning and dynamic identification. *IEEE Transactions on Industrial Electronics*, 66(1):509–518, 2019.

[MÖ7]     Meinard Müller. Dynamic time warping. *Information Retrieval for Music and Motion*, 2:69–84, 01 2007.

[Mad07]    Henrik Madsen. *Time Series Analysis*. CRC Press, Boca Raton, Fla, 2007.

[MBF12] Tobias Maile, Vladimir Bazjanac, and Martin Fischer. A method to compare simulated and measured data to assess building energy performance. *Building and Environment*, 56:241–251, 2012.

[MOd+21] Mara Magni, Fabian Ochs, Samuel de Vries, Alessandro Maccarini, and Ferdinand Sigg. Detailed cross comparison of building energy simulation tools results using a reference office building as a case study. *Energy and Buildings*, 250:111260, 2021.

[OK74] W. Oberhofer and J. Kmenta. A general procedure for obtaining maximum likelihood estimates in generalized regression models. *Econometrica*, 42:579–590, 07 1974.

[OWHW20] Mitchell O'Hara-Wild, Rob Hyndman, and Earo Wang. *fable: Forecasting Models for Tidy Time Series*, 2020. R package version 0.2.1.

[Par62] Emanuel Parzen. On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, 33(3):1065 – 1076, 1962.

[PHS+20] M. Parzinger, L. Hanfstaengl, F. Sigg, U. Spindler, U. Wellisch, and M. Wirnsberger. Residual analysis of predictive modelling data for automated fault detection in building's heating, ventilation and air conditioning systems. *Sustainability*, 12:6758, August 2020.

[Plu19] Martyn Plummer. *rjags: Bayesian Graphical Models using MCMC*, 2019. R package version 4-10.

[Pow08] David Powers. Evaluation: From precision, recall and f-factor to roc, informedness, markedness & correlation. *Mach. Learn. Technol.*, 2, 01 2008.

[R C20] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020.

[RC12] A.C. Rencher and W.F. Christensen. *Multivariate Regression*, chapter 10, pages 339–383. John Wiley & Sons, Ltd, 2012.

[RH03] Robert Sollis Richard Harris. *Applied Time Series Modelling and Forecasting*. John Wiley & Sons., Chichester, UK, 2003.

[Rob] Rob J. Hyndman. The arimax model muddle. `https://robjhyndman.com/hyndsight/arimax/` (Last accessed 14.07.2021).

[RTH+11] Xavier Robin, Natacha Turck, Alexandre Hainard, Natalia Tiberti, Frédérique Lisacek, Jean-Charles Sanchez, and Markus Müller. proc: an open-source package for r and s+ to analyze and compare roc curves. *BMC Bioinformatics*, 12:77, 2011.

[Sar13]   R G Sargent. Verification and validation of simulation models. *Journal of Simulation*, 7(1):12–24, February 2013.

[Sch20]   Rainer Schlittgen. *Angewandte Zeitreihenanalyse mit R*. Walter de Gruyter GmbH & Co KG, Berlin, 4. auflage edition, 2020.

[SSB19]   Joachim Seifert, Lars Schinke, and Maximillian Beyer. *Grundlagen der Wärmephysiologie theoretische Grundlagen, stationäre und instatinäre Modelle*. ITM InnoTech Medien, 2019.

[Sti99]   David Stirzaker. *Probability and Random Variables*. Cambridge University Press, Cambridge, 1999.

[TGQS08]  Paolo Tormene, Toni Giorgino, Silvana Quaglini, and Mario Stefanelli. Matching incomplete time series with dynamic time warping: An algorithm and an application to post-stroke rehabilitation. *Artificial Intelligence in Medicine*, 45(1):11–34, 2008.

[TH-17]   TH-Rosenheim. InSitu Nachweisverfahren. Website, 2017. https://www.th-rosenheim.de/forschung-entwicklung/ kompetenzfelder-und-projekte/energietechnik-und-energieeffizienz/ insitu-nachweisverfahren/.

[VDI]   VDI-Fachbereich Technische Gebäudeausrüstung. VDI 2078:2015-06 - Berechnung der thermischen Lasten und Raumtemperaturen (Auslegung Kühllast und Jahressimulation).

[Vin68]   T. K. Vintsyuk. Speech discrimination by dynamic programming. *Cybernetics*, 4(1):52–57, 1968.

[VR02]   W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S*. Springer, New York, fourth edition, 2002. ISBN 0-387-95457-0.

[WCH20]   Earo Wang, Dianne Cook, and Rob J Hyndman. A new tidy data structure to support exploration and modeling of temporal data. *Journal of Computational and Graphical Statistics*, 29(3):466–478, 2020.

[Wic16]   Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.

[Woo06]   Simon Wood. *Generalized Additive Models - An Introduction with R*. CRC Press, Boca Raton, Fla, 2006.

[YSMA14]  Ke Yan, Wen Shen, Timothy Mulumba, and Afshin Afshari. Arx model based fault detection and diagnosis for chillers using support vector machines. *Energy and Buildings*, 81:287 – 295, 2014.

[YYK07]   Toshie Yamashita, Keizo Yamashita, and Ryotaro Kamimura. A stepwise AIC method for variable selection in linear regression. *Communications in Statistics - Theory and Methods*, 36(13):2395–2403, October 2007.