

# Grey-box models for prediction and control of solar district heat plants

by

**Linde Josien Frölke**

A thesis submitted to the Delft Institute of Applied Mathematics,  
in partial fulfillment of the requirements for the degree of

**Master of Science**

in Applied Mathematics  
specialization Computational Science & Engineering

at the Delft University of Technology, Faculty of EEMCS,  
to be defended publicly on Wednesday May 30, 2018 at 09:30 AM.

Supervisors:

Prof. dr. ir. A. W. Heemink, TU Delft  
Dr. ir. P. Bacher, TU of Denmark (DTU)

Other thesis committee member:

Prof. dr. ir. M. B. van Gijzen, TU Delft

*This thesis is confidential and cannot be made public until May 30, 2018.*

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.



# Abstract

It has become clear that the global energy system needs to shift away from fossil fuels towards clean, renewable energy sources. Experience in the past decades has shown that flat-plate solar heat panels can play a role in the energy system of the future, in particular in solar collector fields for district heat generation. This thesis concerns the dynamical modelling of such solar district heat plants (SDHPs). As the upswing of SDHPs occurred rather recently, research on their optimal control is ongoing. The main challenge for control is to adapt to fluctuations in solar radiation and other inputs in an optimal way, ensuring a high and stable outlet water temperature to the grid while minimizing flow fluctuations. Many modelling efforts of single collectors as well as full solar heat fields have been reported, although mostly in the form of detailed physical models.

In this thesis, a new approach for describing the dynamics of a large flat-plate solar field is proposed. We develop a continuous-discrete stochastic state space model suitable for prediction and control. This model form combines knowledge from physics and information from data, thereby allowing for relatively simple formulations while modelling complex dynamics. Retaining a physically meaningful model formulation has additional advantages for model development, as analysis of residuals and outputs can provide information on suitable model extensions. A basic model was formulated and systematically extended in a forward selection procedure, using a.o. likelihood ratio tests. The model development was based on May 2017 measurements from an Aalborg CSP solar heat plant in the municipality of Solrød, Denmark. The models are implemented using the R-package CTSM-R, which includes parameter estimation methods based on maximum likelihood estimation and the extended Kalman filter.

It is found that the developed model is suitable for very short term (minutes to hours ahead) to short term (day ahead) prediction, as needed for control and heat output forecasting of a SDHP. It includes several new aspects compared to existing models, the most notable being non-parametric modelling of shading effects and a split of total radiation into diffuse and direct components. Including these elements improves model predictions considerably, and allows for asymmetric panel efficiency over the day. Detailed analysis of the model's predictive performance is provided, including a comparison to current ISO standard model and the current Solrød control scheme, as well as a cross-validation on data from different seasons. In future work, the model's performance when using input predictions from weather forecasts should be tested. The model should further be used in a model predictive control scheme in order to improve current SDHP control strategies. This would lead to a smoother pump operation, thereby reducing electricity consumption, costs, and greenhouse gas emissions.

# Acknowledgements

This thesis was submitted as the final requirement for the Master degree Applied Mathematics at the Delft University of Technology. The main part of the research was conducted as a guest student at the Technical University of Denmark (DTU), at the department DTU Compute. The data set central to this research was obtained from CSP Aalborg, the commissioner of the studied solar heat plant.

I am taking this unique opportunity to thank some of the many people that helped me write this thesis. How often does one get such a wonderful chance to officially express some gratitude?!

First of all I would like to express my thanks to my two (!) supervisors Peder Bacher (DTU) and Arnold Heemink (TU Delft). Peder, I have enjoyed your supervision very much. In our frequent meetings you always managed to offer me both useful input, confidence in the project, and most likely a fun fact that made me leave with a smile on my face. You combine theoretical knowledge and practical insight in a way that I really hope to achieve one day too. In addition, you provided me with food for (non-mathematical) thought on among others roller skating to work, saving the world, and the proper use of smileys and ‘yeah!’ in emails. I am looking forward to working together more!

Arnold, thank you so much for all your support, not in the least for making it possible for me to write the thesis in Copenhagen. It has felt very reassuring to have such a knowledgeable and experienced researcher guarding the outcomes of this project. Thank you for reading my (usually slightly too elaborate) writings; it was impressive how you each time managed to point to parts of the text that I was unsure about myself. I imagine it would have been more fun for you to supervise in real life instead of over email, but I am very happy about the way the communication has worked out and hope you feel the same way too.

I will also use this opportunity to thank Martin van Gijzen for being so optimistic, cheerful, and personally interested in his students. It was your positive attitude and personal commitment that in the first place convinced me to come to Delft for my master, where I have felt welcome from the start. Although I have been coming up with weird ideas for study extensions and excursions from Delft, it was never too crazy for you, who simply replied that ‘it was your job to make things possible’. And you have!

Another great thanks to Henrik Madsen, who has been positive from the start about welcoming me at his department. I further greatly enjoyed your lectures in Advanced Time Series Analysis, and your written materials have been very helpful as well.

I want to thank CSP Aalborg for building solar heat plants, saving the world, and allowing me to work with their well-organised data set; it has been a pleasure. A special thanks to Andreas Zourellis for sharing his knowledge with me in many e-mails between us, and of course for showing me around at the Solrød solar heat plant on an especially cold and snowy day. It would be nice if we could

continue the project in the future.

Finally I would like to thank some people that have indirectly enabled me to write this thesis. My parents get infinite thanks simply for being my parents. You have been extremely supportive and patient with my long and windy study path and international adventures. You have taught me that it is never too late to change (studies), which has helped me a great deal when facing big choices. And maybe even more importantly, you have shown me the importance of making things ‘gezellig’ and having a nice meal with the people you love (including of course a ‘borrelhapje’). Steven, thank you for growing up to be such a sweet, reliable, not-so-little-anymore little brother. Coming home has not been the same without you there, and I am glad you will be back soon!

Lisa, thank you for sharing the sorrows of the many early train trips to Delft. I found it quite special how quickly we got to know each other well, and how we have so much in common (especially lunch preferences). To my sweet Danish roommate Ida, thank you for welcoming me in your hyggeligt apartment and bearing with me throughout my busy days. Many many thanks to every single member of the Familie Multirent for being such awesome friends. You gave me the best student life I could have wished for. Thanks Roshell, schaapie, for all the great days at Science Park and ‘koffie bij Sjeel’ that I will miss so much (and miss already), and all the times I could crash the amazing Stayokay bar. Huize Nasi, thanks for all the sleepovers, being so hospitable, and making me feel welcome and at home.

Lastly, thank you Tobias, for being the relaxed and nerdy and scouty and smart and sweet person that you are. I feel so lucky to have you in my life. Mm. Cheesy again.

*Linde Frölke  
Copenhagen, May 2018*

# Contents

<b>Nomenclature</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Solar Thermal Energy</b>	<b>5</b>
2.1 Solar Heat Panels . . . . .	5
2.2 Solar District Heat Plants . . . . .	5
2.2.1 Solar District Heating in Denmark . . . . .	6
2.3 Solar District Heat Plant Solrød . . . . .	6
<b>3 State of the art: modelling solar heat panels &amp; fields</b>	<b>8</b>
3.1 Single collector models . . . . .	8
3.1.1 ISO standard . . . . .	8
3.1.2 Grey box model . . . . .	10
3.2 Models for control of solar district heating plants . . . . .	11
3.2.1 Discretized PDE state space model . . . . .	11
3.3 Model currently used for control in Solrød . . . . .	11
<b>4 Stochastic State Space Models</b>	<b>13</b>
4.1 Introduction to state space formulation . . . . .	13
4.2 Shortcomings of deterministic state space models . . . . .	15
4.3 Stochastic state space formulation . . . . .	16
4.4 Motivation for using Stochastic State Space representation . . . . .	18

<b>5</b>	<b>State estimation in stochastic state space models</b>	<b>19</b>
5.1	Filtering in general state space model . . . . .	20
5.2	The Kalman Filter for linear time varying systems . . . . .	21
5.2.1	Prediction step . . . . .	21
5.2.2	Update step . . . . .	22
5.2.3	Kalman filter optimality . . . . .	22
5.3	Filtering and Prediction in nonlinear models: An Extended Kalman Filter . . . . .	23
5.4	Iterated Extended Kalman Filter . . . . .	24
<b>6</b>	<b>Maximum Likelihood Estimation in Stochastic State Space models</b>	<b>26</b>
6.1	Adapted likelihood function for robustness . . . . .	27
<b>7</b>	<b>The CTSM-R library</b>	<b>28</b>
7.1	Filtering methods . . . . .	28
7.2	Numerical stability of the Kalman filter . . . . .	29
7.3	Evaluation of matrix exponentials and their integrals . . . . .	29
7.4	Optimization details . . . . .	30
7.5	Parameter estimation . . . . .	31
7.6	Built in statistics for evaluation of fit . . . . .	31
7.7	Using and testing a fitted model . . . . .	33
<b>8</b>	<b>Exploration and preparation of the data set</b>	<b>34</b>
8.1	Data exploration . . . . .	34
8.2	Data preparation . . . . .	36
8.3	Data selection for parameter estimation . . . . .	37
<b>9</b>	<b>Model Selection and Extension</b>	<b>38</b>
9.1	First model formulation . . . . .	39
9.2	Extended models: formulation and fitting . . . . .	39
9.3	Selection of best extension using likelihood ratio tests . . . . .	39
9.4	Evaluation of selected model . . . . .	41

9.5	Indications for model extension . . . . .	42
<b>10</b>	<b>Formulation of grey-box models for the Solrød solar heat plant</b>	<b>43</b>
10.1	Base models $T_f$ and $T_o$ . . . . .	43
10.2	Including extra terms in heat balance . . . . .	44
10.3	Modelling shading: splines . . . . .	44
10.4	Splitting total radiation in its diffuse and direct components . . . . .	45
10.5	Models with incidence angle modifier . . . . .	47
10.5.1	IAM based on panel specifications . . . . .	48
10.6	Models with multiple compartments . . . . .	48
10.7	Models with a metal component . . . . .	49
<b>11</b>	<b>Parameter estimates according to EU performance test</b>	<b>51</b>
11.1	Results without split radiation: model $MLR1$ . . . . .	52
11.2	Results with split radiation, model $MLR2$ . . . . .	52
11.3	Discussion and comparison of model $MLR1$ and $MLR2$ . . . . .	53
<b>12</b>	<b>Model selection results</b>	<b>56</b>
12.1	Base model selection . . . . .	56
12.1.1	Model $T_f$ . . . . .	56
12.1.2	Model $T_o$ . . . . .	60
12.1.3	Motivation for model extensions . . . . .	61
12.2	First extension . . . . .	64
12.3	Second extension . . . . .	67
12.4	Third extension . . . . .	70
12.5	Fourth extension . . . . .	72
12.6	Fine-tuning the final model . . . . .	76
12.6.1	Adapted splines . . . . .	76
12.6.2	Delayed inlet temperature . . . . .	77
12.6.3	Combined effect . . . . .	80

12.7 Summary model sequence . . . . .	80
<b>13 Final model performance</b>	<b>95</b>
13.1 Final model predictive performance on fitting data . . . . .	95
13.2 Comparison to MLR2 and current Solrød scheme . . . . .	97
13.3 Cross-validation on other data sets . . . . .	99
13.3.1 Without re-estimation of the parameters . . . . .	99
13.3.2 Re-estimation of parameters . . . . .	100
<b>14 Conclusion, Discussion, &amp; Recommendations</b>	<b>104</b>
14.1 Conclusion . . . . .	104
14.2 Discussion . . . . .	106
14.3 Recommendations . . . . .	108
14.3.1 Next steps for model implementation . . . . .	109
14.3.2 Suggestions for future research . . . . .	109
<b>Appendices</b>	<b>114</b>
<b>A Derivation of the incidence angle</b>	<b>115</b>

# Nomenclature

$a_6$	wind speed dependence of the zero loss efficiency [s/m]
$a_7$	wind speed dependence of IR radiation exchange [W/(m <sup>2</sup> K <sup>4</sup> )]
$a_8$	radiation losses [W/(m <sup>2</sup> K <sup>4</sup> )]
$A_c$	total collector area [m <sup>2</sup> ]
$A_{f,m}$	area [m <sup>2</sup> ]
$b_0$	parameter used in incidence angle modifier [ - ]
$c_f$	collector fluid heat capacity [J/(kg·K)]
$C_{f,m}$	heat capacity of fluid or metal [J/(kg ·K) ]
$D_{i,o}$	inner or outer diameter [m]
$E_L$	long-wave irradiance ( $\lambda > 3\mu\text{m}$ ) [W/m <sup>2</sup> ]
$F'U_1$	temperature dependence of panel heat loss coefficient [W/(m <sup>2</sup> K)]
$F'U_0$	panel heat loss coefficient at $(T_a - T_f) = 0$ [W/(m <sup>2</sup> K)]
$F'(\tau\alpha)_{\text{en}}$	panel zero loss efficiency for radiation [ - ]
$F'U_u$	wind dependence of panel heat loss coefficient [J/((m <sup>2</sup> ·K)]
$F'\epsilon$	sky temperature dependence of the heat loss coefficient
$G$	both used for global radiation and for total radiation on collector plane [W/m <sup>2</sup> ]
$G_{b,\text{eff}}$	effective direct beam radiation [W/m <sup>2</sup> ]
$G_{b,i}$	$i^{\text{th}}$ splined component of direct beam radiation in collector plane [W/m <sup>2</sup> ]
$G_b$	direct beam radiation in collector plane [W/m <sup>2</sup> ]
$G_d$	diffuse radiation in collector plane [W/m <sup>2</sup> ]
$K$	Kalman gain matrix
$K_b(\theta)$	incidence angle modifier for direct radiation [ - ]
$K_d$	incidence angle modifier for diffuse radiation [ - ]
$(mC)_e$	effective thermal capacitance of the system [J/(m <sup>2</sup> K)]

$P$	state covariance matrix
$P$	useful power output per aperture area [W/m <sup>2</sup> ]
$P_{\max}$	maximal power output of collector field [W]
$Q_f$	volume flow per collector area [kg/(s·m <sup>2</sup> )]
$R$	output covariance matrix
$S$	measurement error covariance matrix
$s_i$	efficiency parameter of $i^{\text{th}}$ splined direct beam component $G_{b,i}$ [ - ]
$T_a$	ambient temperature [° C]
$T_f$	fluid temperature in center of collector row [° C]
$T_i$	inlet temperature [° C]
$T_m$	collector temperature in center of collector row [° C]
$T_o$	outlet temperature [° C]
$T_{\text{set}}$	temperature set point [° C]
$u'$	wind velocity [m/s]
$W_t$	Brownian motion
$yT_o$	measured outlet temperature [° C]
$\epsilon_{\text{obs}}$	Observation error [° C]
$\omega, \omega_{o,m}$	Brownian motion (related tot output or metal temperature)
$\rho_{f,m}$	density of fluid or metal [kg/m <sup>3</sup> ]
$\sigma$	Stefan-Boltzmann constant [ W/(m <sup>2</sup> K <sup>4</sup> )]
$\sigma_{o,m}$	parameter for diffusion part of solar SDE models (outlet or metal temperature )
$\theta$	incidence angle of total or direct radiation [degrees (°)]



# Chapter 1

## Introduction

### Background

The vast majority of the scientific community now acknowledges that human induced climate change is a fact. In 2013, the International Panel for Climate Change stated that “it is extremely likely that more than half of the observed increase in global average surface temperature from 1951 to 2010 was caused by the anthropogenic increase in greenhouse gas concentrations and other anthropogenic forcings together” [1]. Anthropogenic carbon dioxide emissions have primarily come from fossil fuel combustion and land use change [1].

As a global effort to respond to climate change, the Paris Agreement between members of the United Nations Framework Convention on Climate Change (UNFCCC) was established in 2015 and came into effect in November 2016. By ratifying the agreement, 175 parties have agreed to keep global temperature rise before 2100 well below 2 degrees Celsius above pre-industrial levels, with the intention to limit the increase to 1.5 degrees Celsius. In order to reach this goal, the emission of greenhouse gases should be reduced to zero in the upcoming decades [1]. This means that the global energy system needs to become more energy efficient and shift away from fossil fuels to clean, renewable energy sources. According to the United Nation Environment Program (UNEP), the Paris goals can be reached using a wide range of current technically feasible measures [2].

The shift is already visible in investment behaviour: in 2016, “renewables capture two-thirds of global investment in power plants as they become, for many countries, the least-cost source of new generation” [3]. Renewable sources for heat production also doubled in 2016. A number of recent studies come to the conclusion that district heating (DH) will play an important role in sustainable energy systems [4, 5]. A DH system consists of many buildings connected by a network of pipes, which allow for the distribution of heat generated at a centralized plant.

The usage of large solar collector fields for district heat generation has been explored in northern Europe since the 1980s, and has in the recent years expanded significantly in this region [6]. Such Solar District Heat Plants (SDHPs) usually consist of several rows of large flat-plate collectors connected in series. The panel’s absorber plate is heated by solar rays, and this heat is absorbed by a fluid flowing through the panel. A heat pump is used to transfer the generated heat to the district heating network. Solar heat fields reduce the amount of fossil fuel used for district heating, resulting in lower greenhouse gas emissions and operation costs.

## Problem statement

As the upswing of solar district heat plants (SDHPs) occurred rather recently, research on their optimal control is ongoing. The conditions of operation (sun, wind, ambient temperature) show a wide variety, which poses challenges to modelling the dynamics [7]. The main challenge for control, compared to fossil fueled plants, is that the source of energy cannot be manipulated and shows not fully predictable behaviour on short timescales. This is especially an issue in northern Europe, where weather conditions are often not stable over the day. The control has to adapt to fluctuations in solar radiation and other inputs in an optimal way, in order to ensure a high and stable outlet water temperature to the grid, and to prevent boiling.

Current control schemes often show large flow fluctuations, showing a need for more stable flow regulation. This need was also expressed by Aalborg CSP, commissioner of various solar heat plants including the Solrød SDHP studied in this research. Smoother pump operation can reduce the plant's electricity consumption, and thereby reduce costs and greenhouse gas emissions.

To prevent system overheating on sunny summer days, cool water from the storage tank can be used. If the storage tank is running out of cold water after several sunny days, the Solrød plant may need to perform night cooling in order to prevent overheating during the next day. Currently, the decision to activate night cooling is not based on model prediction, but on personal judgement based on weather forecasts. A model prediction of tomorrow's heat production could assist or automate this decisionmaking for night cooling. In addition, such forecasts are potentially valuable in smart grid applications, e.g. when planning the use of a heat pump so that it can use excess wind power.

Solar heat panels and fields are considered to show highly non-linear behaviour, with changing dynamics depending on the operating conditions [8, 9]. Changes in solar radiation can for example alter the dominant time constant of the process on a very short time scale [8, 10]. It is therefore widely recognized that traditional control methods are not suitable for the control of such systems, and that "the performance of PI and PID type controllers will be inferior to model based approaches" [8]. More advanced, non-linear controllers are therefore needed to control the system over the entire operation range [9]. Such controllers are based on non-linear models of the plant dynamics.

## Research Gap

In the past decades, much research has been done on the modelling of single collectors and solar fields. This has among others resulted in a widely used ISO standard for performance testing in the European Union. Most existing models for full plants are detailed physical models, often implemented in the simulation program TRNSYS. Their parameters are either derived from physics, obtained from physically detailed sub-models, or taken from performance tests. Such models either intend to estimate plant performance or to predict power output over a certain period. None of the plant models intended for power output predictions found in literature used statistical methods to estimate parameters for a particular plant. Furthermore, accounting for shading effects is a challenge in these white-box models.

Only few (nonlinear) models intended for the control of solar district heating plants have been described in the scientific literature. The existing models combine partial differential equations from physics with parameters fit from data. The models are rather elaborate, distinguishing two one-dimensional system states (continuous in space) which require a discretization in space. To our best knowledge, stochastic differential equations (SDE)-based models have never been used for this purpose.

SDE-based state space models combine physical knowledge with information from data using methods from statistics, and are therefore of grey-box form. In past research, such models have been successfully applied in heat dynamics applications, including the heat dynamics of buildings [11, 12], the heat dynamics of building integrated photovoltaic modules [13], and the heat dynamics of a single solar heat collector [14]. The promising results of the latter article suggest that the models can be used to describe the dynamics of entire SDHPs as well.

The stochastic nature of these models allows for the application of statistical methods for structural model identification, which is a challenging part of model building. The resulting models are relatively simple and data-driven, with parameter estimates that are easily interpretable due to the physics-inspired model equations [12]. The method further allows for parameter re-estimation from recent data, so that changes in the system and its performance can be monitored. Non-parametric methods can be applied in these models, for example to account for shading effects. Grey-box models can therefore be a promising contribution to solar heat plant modelling, both for performance monitoring, for the forecasting of heat power production, and for improving control.

## Research Aims

The aim of this thesis is to develop a continuous-discrete stochastic state space model of solar heat plant dynamics, which is suitable for the following applications:

1. Predicting the outlet temperature of a large solar collector field for improved control strategies (very short term prediction in the range of seconds to hours)
2. Predicting multiple day ahead heat power production for planning of operation (short term prediction)

The developed grey-box model should combine knowledge from physics and information hidden in data in an optimal manner, allowing for a formulation as simple as possible while modelling complex dynamics. A particular aspect of the models will be to account for shading effects in a data-driven manner, which is a simple way to improve model predictions considerably.

The result may be a single model suitable for both short and medium term prediction, or two separate models each suitable for one of the applications.

## Methodology

The models will be formulated in order to predict the outlet temperature of the solar field. All models will be implemented using the free software environment R [15] and the R-package CTSM-R [16]. The latter is documented in Kristensen et al. [17]. This software performs parameter estimation based on maximum likelihood estimation and the extended Kalman filter, and includes functions for simulation and prediction given a model and obtained parameter estimates.

A basic model will be formulated, which will then be systematically extended in a forward selection procedure. The likelihood ratio test and prediction residual plots will be used for model selection and validation. The final model performance will be evaluated by comparison to current standards and the control scheme as currently used in the Solrød plant. Furthermore, a cross-validation on data from different seasons is performed.

The model development will be based on May 2017 measurements from a solar heat plant located in Havdrup, a town in the municipality of Solrød, just South of Copenhagen, Denmark. The data was received from Aalborg CSP, the commissioner of this plant.

## Structure of this thesis

The first chapters of this thesis will introduce the reader to relevant (non-)mathematical theory. Chapter 2 will provide an introduction to solar thermal energy, including SDHPs and the solar heat plant in Solrød in particular. Next, Chapter 3 presents an overview of previous research on solar thermal modelling for parameter estimation and improved control. In Chapters 4 to 6, stochastic state space models and their filtering and parameter estimation methods will be discussed.

The next few chapters relate to methods used in this thesis. Chapter 7 briefly discusses the used software package CTSM-R. The data set is introduced and prepared for use in Chapter 8. Next, Chapter 9 describes a model selection procedure that has been used in past research and has been employed in this research. Chapter 10 introduces all model equations and components that have been used in the model sequence of this thesis.

Thereafter, the results of the modelling efforts of this thesis are presented. Chapter 11 first shows the results of MLR model fits as described in ISO standards. Next, Chapter 12 describes the model sequence that was followed to develop new grey-box models for the solar heat plant. The performance of the final developed model is elaborated on in Chapter 13, which will also include a cross-validation on data from other months. Final conclusions are drawn in Chapter 14, which will also include a critical discussion, directions for practical application, and recommendations for future research.

This thesis is written under the assumption that the reader has knowledge on basic probability theory and stochastic calculus. Theory on stochastic differentiation and integration will not be treated, but many good books on these topics exist. For example, Kloeden, Platen, and Schurz's *Numerical Solution of SDE Through Computer Experiments* extensively treats basic theory on probability, the Wiener process, stochastic integration, and stochastic differential equations (see reference [18]). Another good source is *Basic Stochastic Processes* by Zastawniak and Brzezniak (reference [19]).

## Chapter 2

# Solar Thermal Energy

This chapter will provide a brief introduction into solar thermal energy and solar district heating in general. In addition, the particular solar heat plant considered in this thesis will be described.

### 2.1 Solar Heat Panels

Solar panels capture energy from solar radiation. In solar photo-voltaic (PV) applications, the radiation energy is converted into electricity, while solar thermal panels or solar heat panels convert it into thermal energy of some working fluid. Hybrid (PVT) panels exist as well, which are able to produce both types of energy. This thesis will only concern solar heat panels. More specifically, so-called flat plate collectors will be discussed, which operate in the low to medium temperature range. The other common collector type is the parabolic trough collector, which can reach higher temperatures by concentrating the solar rays.

Typical flat plate collectors have a black background *absorber plate* covered with insulation layers and a glazing cover [20]. The glazing is usually made of one or more sheets of glass, which is a suitable material as it transmits short-wave radiation well while having a low transmissivity of long-wave radiation. The glazing should prevent convection losses, as well as reduce irradiation losses by reflecting radiation transmitted by the fluid (a green house effect).

Inside the absorber plate, a pipe containing the heat transfer fluid meanders horizontally, in order to capture as much heat as possible from the plate [21]. The heat needs to be absorbed rapidly, so that the system will not overheat [20]. The heat transfer fluid used in such systems is usually a mix of water and polypropylene glycol, which a.o. has the advantage of anti-freeze protection at low temperatures.

### 2.2 Solar District Heat Plants

When the heat generated by solar panels is used to supply district heating networks, the term Solar District Heat Plant (SDHP) is used to describe the system.

Solar collector fields for DH usually consist of several rows in which large flat-plate collectors are

connected in series. The rows are in turn connected in parallel. In this way, hundreds of collectors can be coupled. A heat pump then transfers the generated heat to the DH grid. District heating grids demand temperatures between 40-100° Celsius [22]. The supplied heat is used for space heating and domestic hot water.

The systems aim to reduce or replace fossil energy sources for district heating. Due to the intermittent nature of solar energy however, the existing fossil fuel based parts of the district heating network are usually still in use for period of low solar heat production. In order to keep this to a minimum, diurnal and/or seasonal heat storage may be installed. Seasonal storage may take the shape of water pit holes, which can store overproduced heat from summer season.

### 2.2.1 Solar District Heating in Denmark

Although Scandinavia is perhaps not the most obvious place to capture solar heat, the operation of SDHPs in this area is now commercially feasible. From the ranking list of European large scale SDHPs (last updated end of 2016), the 25 largest are located in Denmark [23]. A total of 106 Danish plants are listed on this ranking.

The fact that district heating networks are already present in many Danish municipalities has no doubt facilitated the rapid development of the technology in this country [24]. Approximately 60% of Danish buildings heat demand is supplied by district heating [22]. In addition, the Danish government has set itself the goal of being fossil fuel free by 2050. Due to high fossil fuel taxes, solar heat production has furthermore become economically beneficial [25]. Solar district heating is therefore a mature market in Denmark, which can exist without subsidies [22].

## 2.3 Solar District Heat Plant Solrød

This section is based on information provided by Aalborg CSP, the commissioner of the plant. The plant is located in Havdrup, part of municipality of Solrød, where it serves a community of 350 households. It was constructed in the period from October to December 2016, and by the end of January 2017 the plant started to produce energy. After a period of close monitoring and finetuning the system was delivered to the client, Solrød Fjernvarme, by the beginning of April.

The plant is a commercial project intended to reduce costs for the community, by reducing natural gas purchases. Another benefit of solar heat production is an ensured stable price of the heat for the entire lifetime of the project, independent of fluctuating gas prices. The operating costs of a SDHP are estimated around 0.5 percent of the total capital investment, and mainly come from the electricity cost for operating the pumps. Panel maintenance is minimal, as there is no need to clean the surface of the panels due to frequent rain.

The plant, depicted in Figure 2.1a, consists of 11 collector rows connected in parallel. Due to space constraints, the first two rows are 4 rows in space, but the water flows in series. The first three rows contain 16 collectors in series, the remaining rows contain 20. This adds up to a total of 208 collectors with a total area of 2569m<sup>2</sup>. The panels are GreenOneTec GK 3003 large size collectors, as depicted in Figure 2.1b. The first half of each row has a single glass glazing, whereas the second half holds a double glass layer. Furthermore, a 1250 m<sup>3</sup> accumulation tank is used for multiple-day storage.

A special feature of this plant is its operation as a stand alone system: the hot water is directly

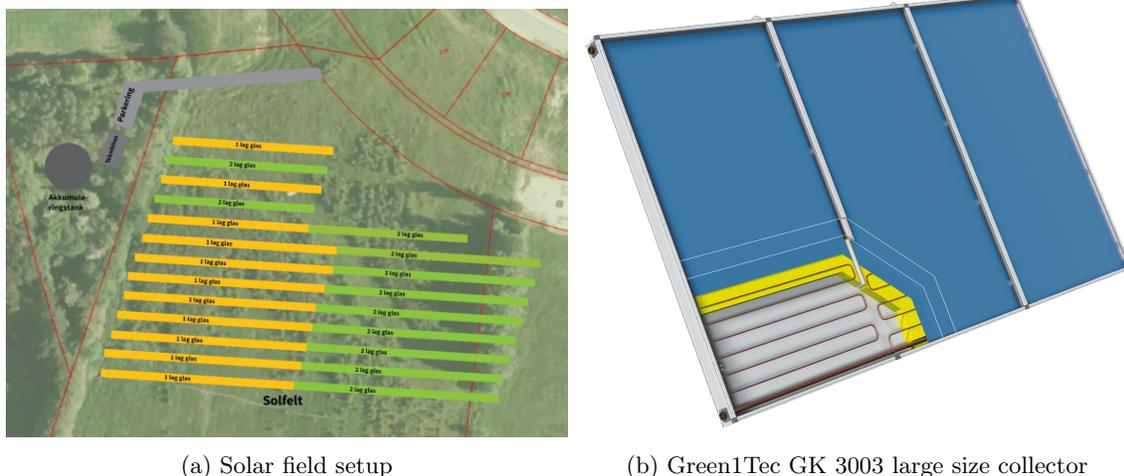


Figure 2.1: Solar field and solar collector from Solrød

supplied to the grid. Most plants in Denmark operate differently: they exchange heat with a central heating plant first, which in turn distributes the water to the connected houses [22].

During the summer months, the solar field can supply a large share of the heat demand. In 2017, almost the whole demand from May to the beginning of September was met by solar heat alone. The winter months can be productive too: this year's (2018) first production day was in mid-February, during which the plant produced more energy than in December and January together (almost 7 MWh). The field is designed to yield 1300 MWh annually, about 28 percent of the total heat consumption. The remainder of the heat demand is met using natural gas. Through reduced natural gas use, approximately 230 tonnes of CO<sub>2</sub> are saved annually. Additionally, the plant is expected to reduce energy costs significantly.

Water with 30% glycol is used as the heat transfer fluid. The piping has been designed to have an equal flow in all collector rows. An overview of important parameters (some are used by the plant for current control) is given in Table 2.1.

Parameter	Value	Unit
Total aperture area $A_c$	2569	m <sup>2</sup>
Latitude	-55.54	degree
Longitude	12.11	degree
Panel Slope	45	° wrt ground normal
Panel Azimuth	6	° wrt S (+ clockwise)
Fluid $c_f$	3980	J kg <sup>-1</sup> K <sup>-1</sup>
$F'(\tau\alpha)_{\text{en}}$ single glass	0.857	-
$F'(\tau\alpha)_{\text{en}}$ double glass	0.814	-
$F'U_0$ single glass	3.083	W/m <sup>2</sup> K
$F'U_0$ double glass	2.102	W/m <sup>2</sup> K
$F'U_1$ single glass	0.013	W/m <sup>2</sup> K <sup>2</sup>
$F'U_1$ double glass	0.016	W/m <sup>2</sup> K <sup>2</sup>
$P_{\text{max}}$	1.9	MW

Table 2.1: Overview of Solrød parameters

## Chapter 3

# State of the art: modelling solar heat panels & fields

This chapter will provide an overview of research on models for solar heat applications. Although this thesis aims to model a field of many connected collectors, two models for single collectors will be discussed as well. The widely used ISO-standard collector model will be discussed in Section 3.1.1, which represents the state-of-the-art and is the basis for many other models. Section 3.1.2 presents the grey-box SDE model for a single collector by Bacher et al., as this model formed the basis for the models developed in this thesis. In Section 3.2, an overview of modelling for control of solar district heat plants will be presented.

### 3.1 Single collector models

Models for single collectors have been in development over the last few decades. Such models are usually developed either for the estimation of panel performance parameters (i.e. performance testing), for the control of small, domestic installations, or for the prediction of heat output. Most models are based on physical principles, with parameters derived from physics or determined from data. Black-box approaches have also been reported, for example using Artificial Neural Networks in [26] or using fuzzy modelling in [27]. Grey-box methods are presented a.o. in [14, 28]. The black-box approaches will not be discussed in detail.

#### 3.1.1 ISO standard

In order to harmonize the European solar market, the region has set a standard for testing solar thermal collectors for determining performance parameters. The standard is revised every fifth year by a research group working for the International Organization for Standardization (ISO). The most recent version of the method is described in ISO standard 9806:2017 [29], which is considered the state-of-the-art model for solar collectors [30]. The tests result in the computation of certain parameters that are an indication of a collector's performance, among others the panel efficiency  $F'(\tau\alpha)_{\text{en}}$ , the heat loss coefficient  $F'U_0$ , and the temperature dependence of the heat loss  $F'U_1$ . In the equations discussed below, the interpretation of these parameters will become clear.

The ISO standard distinguishes between a steady state and a quasi-dynamic performance test method. The steady state model requires clear weather, small solar incidence angle, low diffuse radiation levels, and wind speed of around 3 m/s. The quasi-dynamic approach allows for a slightly higher wind speed. In the latter method it is important that the system is exposed to all weather conditions possible for normal operation of the plant: variability is needed to decouple parameter estimates.

At every time  $t$ , the useful power output per aperture area is defined as:

$$P = c_f Q_f (T_o - T_i) \quad (3.1)$$

where  $c_f$  is fluid heat capacity,  $Q_f$  is flow per unit aperture area of the collectors, and  $T_o$  and  $T_i$  are respectively the out- and in-flowing fluid. The quasi-dynamic model equation for  $Q$  is given by the following heat balance (see Nomenclature for overview of symbols):

$$\begin{aligned} P = & F'(\tau\alpha)_{en} [K_{\tau\alpha,b}(\theta)G_b + K_{\tau\alpha,d}G_d] - F'U_0(T_f - T_a) - F'U_1(T_f - T_a)^2 - F'U_u u'(T_f - T_a) \\ & + F'\epsilon(E_L - \sigma T_a^4) - (mC)_e \frac{dT_f}{dt} - a_6 u'G - a_7 u'(E_L - \sigma T_a^4) - a_8(T_f - T_a)^4 \end{aligned} \quad (3.2)$$

In the most recent version of the ISO standard, all parameters are labelled  $a_1$  up to  $a_8$ , but for clarity some of the original parameter names have been retained here.

The first term on the right hand side of equation (3.2) represents the energy input from radiation, which is split into beam and diffuse components. As the term implies, direct (beam) radiation comes in a direct line from the sun, whereas diffuse radiation is scattered from that beam by molecules, aerosols, and clouds. The portion of total solar radiation (diffuse + direct) that is diffuse is about 10% to 20% for clear skies and up to 100% for cloudy skies. The distinction is relevant for the developed models, because the panel's energy uptake from these terms depends on the solar incidence angle in different ways.

Both types of radiation are corrected for the incidence angle: the diffuse radiation is corrected by a (constant) parameter  $K_d$ , whereas the beam radiation is corrected using an Incidence Angle Modifier (IAM) function  $K_b(\theta)$ . Section 3.1.1 will provide a deeper discussion of this function. The parameter  $F'(\tau\alpha)_{en}$ , the zero-loss efficiency, expresses the efficiency of the collector when the temperature gradient to the ambient,  $T_f - T_a$ , equals 0.

The heat loss from the solar collector to the ambient is related to the temperature difference between the air and the collector with the heat loss coefficient  $F'U_0$ . As the collector is often warmer than the air and does not aim to capture heat from its surroundings, an ideal collector has a low heat loss coefficient. The second order term in equation 3.2 has parameter  $F'U_1$  which represents the temperature dependence of the heat loss coefficient. This parameter is often positive (f.e. in [31]), so that the fraction of heat lost to the environment increases with increasing temperature gradient ( $T_f - T_a$ ). The optional term  $a_8(T_f - T_a)^4$  allows to refine the temperature dependence of the heat loss coefficient further.

The time derivative of the collector fluid temperature  $T_f$  comes into the equation as  $(mC)_e \frac{dT_f}{dt}$  to reflect the energy used to heat up the fluid, where the parameter  $(mC)_e$  represents the thermal inertia of the system. The sky temperature dependence of the heat loss coefficient is accounted for with  $F'\epsilon(E_L - \sigma T_a^4)$ . This term is usually negative, "as the effective sky radiation temperature is lower than the ambient air temperature" [29].

Several terms correct for the wind dependence of parameters. The term  $F'U_u u'(T_f - T_a)$  corrects the wind dependence of the heat loss coefficient,  $a_6 u'G$  for the wind dependence of the zero-loss efficiency  $F'(\tau\alpha)_{en}$ , and finally  $a_7 u'(E_L - \sigma T_a^4)$  for the wind dependence of the long wave radiation exchange.

Chapter 11 will elaborate on the parameter estimation method as described in the ISO standard, and it will describe how this method was applied to the Solrød heat plant.

### Incidence angle modifier

The following formula has become accepted in scientific literature for the Incidence Angle Modifier (IAM) for the direct beam zero loss efficiency  $K_b(\theta)$  (a.o. in [31] )

$$K_b(\theta) = 1 - b_0 \left( \frac{1}{\cos(\theta)} - 1 \right) \quad (3.3)$$

The incidence angle  $\theta$  is defined as the angle between the direction of beam radiation (a straight line to the sun) and the normal to the (tilted) plane of the solar heat collector. Therefore,  $\theta$  depends on the time of the day, day of the year, and on the tilting angle of the solar collectors. The value of  $\theta$  is around  $90^\circ$  in the morning, around zero at noon, and increases to  $90^\circ$  again in the evening. The parameter  $b_0$  that determines the slope of the curve, needs to be estimated. For most flat plate collectors, this equation is considered sufficiently accurate for describing incidence angle effects [32].

Another way of modelling incidence angle effects on direct radiation uptake, is by using the so-called *angle-by-angle* method as proposed by Perers [32]. This method involves a splitting of the direct radiation into different components, each component corresponding to a certain range of incidence angle. A constant direct beam IAM value can then be estimated separately for each component.

### 3.1.2 Grey box model

Bacher et al. [14] report a grey-box model sequence for a single solar heat collector using a simplified heat balance discretized in space. The model is developed with the aim of obtaining accurate performance parameters. Stochastic state space models are fitted and compared to the 2003-ISO standard fitting procedure.

The base model in the sequence has the outlet temperature of the solar heat collector  $T_o$  as its single state variable. The system equation was derived from a simplified heat balance taken from the previous ISO standard equation [32], which was rewritten to the continuous-discrete grey box model equations

$$\begin{aligned} dT_o &= \frac{2}{(mC)_e} \left( F' U_0 (T_a - T_f) + c_f Q_f (T_i - T_o) + \right. \\ &\quad \left. F' (\tau \alpha)_{en} [K_b(\theta) G_b + K_d G_d] \right) dt + \sigma d\omega \\ yT_{o,k} &= T_{o,k} + \epsilon_k \end{aligned} \quad (3.4)$$

In this model,  $T_i$ ,  $T_a$ ,  $T_f$ ,  $Q_f$ ,  $G_b$ ,  $G_d$  are measured variables, and  $c_f$  is the known heat capacity. All other variables in Equation 3.4 are to be estimated.

The model is then extended to include multiple states in the flow direction of the panel, as well as a separate metal component in each of these compartments. All extensions were found to be significant, and it was concluded that further extensions are likely to improve the model. The parameter estimates were in agreement with ISO test results. The authors note that experimental conditions influence the parameter estimates.

## 3.2 Models for control of solar district heating plants

The vast majority of existing models for SDHPs are based on first principles, with parameters derived from physical equations rather than fit from data. Many of these models are developed using TRNSYS, a detailed ODE-based simulation program. Examples are found in [25, 33–35]. The mentioned papers include models for performance analysis as well as control. Recently, some nonlinear models based on MLR ([30, 36]) and transfer functions ([37]) have been reported.

As noted by Bava, most plant models are spatially lumped into one collector component. In addition, the wide majority of existing studies omit the distinction between beam and diffuse radiation, and do not take shadowing effects into account [25].

### 3.2.1 Discretized PDE state space model

Most non-linear statistical modelling for control has been done for high-temperature solar collector fields (parabolic troughs) in for example [10, 38] and summarized in the review [8]. To the best of our knowledge, the only nonlinear state space models for control of flat plate district heating applications have been developed by Pasamontes et al. [39] and further extended by de Andrade et al. [7].

The model includes two state variables: the temperature of the fluid ( $T_f$ ) and that of rest of the collector ( $T_m$ , for metal). The dynamics of these states are modelled by a set of partial differential equations

$$\rho_m C_m A_m \frac{\partial T_m}{\partial t}(t, x) = F'(\tau\alpha)_{\text{en}} G(t) - D_o \pi U_{\text{ma}} (T_m(t, x) - T_a(t)) - D_i \pi U_{\text{fm}} (T_m(t, x) - T_f(t, x)) \quad (3.5)$$

$$\rho_f C_f A_f \frac{\partial T_f}{\partial t}(t, x) = D_i \pi U_{\text{fm}} (T_m(t, x) - T_f(t, x)) - \rho_f C_f u(t) \frac{\partial T_f}{\partial x}(t, x) \quad (3.6)$$

This model is continuous in space and time. The outlet temperature  $T_o(t)$  equals the fluid temperature at the end of the collector field,  $T_f(t, L)$  where  $L$  is the length of the tube. The model parameters are efficiency parameter  $F'(\tau\alpha)_{\text{en}}$ , the heat loss coefficient  $U_{\text{ma}}$  from metal to ambient, and the coefficient  $U_{\text{fm}}$  for heat exchange between metal and fluid. All remaining factors are constants.

Andrade et al. solved this system of PDEs using a discretization in space, where derivatives are replaced by their finite difference approximations. The resulting ODE system was solved numerically using an ODE solver. Parameter values were obtained by a genetic algorithm fitting the integral of squared error criterion. It was assumed that the parameters “do not depend on state variables, which has demonstrated to be a valid approximation for the operation range in these kinds of low-temperature applications” [39]. The model was successfully applied for simulation and control [7].

This model does not take into account shading and incidence angle effects, nor does it account for the different effects of diffuse and direct radiation.

## 3.3 Model currently used for control in Solrød

The control aim of the plant is to maintain a stable outlet temperature  $T_o$  at given set point  $T_{\text{set}}$  that is usually around 95 °C. Besides that, the fluid temperature should not exceed the boiling

temperature, which would lead to glycol evaporation. Such an event is very undesirable, as the system would require a full re-start to redistribute the glycol into the water.

In order to reduce outlet temperature fluctuations to a minimum, the water flow is adapted every second. The desired value of the flow  $Q_f$  is computed based on the measured variables  $G$ ,  $T_i$ ,  $T_a$ ,  $T_{\text{set}}$  as well as some parameters and constants. For the solar radiation, the average over the past 30 seconds is used as an input, in order to reduce flow fluctuations. However, strong variations in the flow are still seen on intermittently cloudy days, and these variations are often redundant.

The current power output  $P$  is first estimated using the heat balance

$$P = A_c \cdot (F'(\tau\alpha)_{\text{en}}G - F'U_0(T_f - T_a) - F'U_1(T_f - T_a)^2) \quad (3.7)$$

where the fluid temperature is computed as the average of the inlet and outlet temperature. Based on this power, the desired flow needed to reach output temperature equal to  $T_{\text{set}}$  is computed as

$$Q_f = \frac{P}{c_f(T_{\text{set}} - T_i)} \quad (3.8)$$

The parameters in Equations 3.7 and 3.8 are taken from the panel specifications, as given in Table 2.1. For  $F'(\tau\alpha)_{\text{en}}$ ,  $F'U_0$ , and  $F'U_1$ , the average of the single and double glazed panels are taken.

It can be seen in Figure 8.1 that the control succeeds reasonably well in maintaining a constant outlet temperature. However, even on clear sunny days there it takes some time after the pre-heating phase before the outlet temperature reaches the set point. This is unnecessarily long, given that the flow in this period is not at its minimal value. The same effect is seen in the afternoon, where the flow not at its minimum, but the outlet temperature decreases away from its set point.

From these observations, it becomes clear that the model used for current control can be improved. The current model structurally overestimates the flow needed to reach the set point temperature.

## Chapter 4

# Stochastic State Space Models

This chapter will outline theory on continuous-discrete State Space Models. The basic formulation of state space models is mainly inspired on Madsen (2007) [40], Maybeck (1979) [41], and Jazwinsky (1970) [42].

The motivation for developing this type of model has been to find a model formulation that accounts for uncertainties in a proper manner, and remains suitable for use of known statistical methods such as maximum likelihood estimation. The state space form is in general appropriate when some relations from physics are known, as the model is often obtained directly from physical equations [40].

### 4.1 Introduction to state space formulation

A state space model describes the dynamics of a physical system and how it is observed. The system is represented by a *state vector*: a set of quantities that represents the entire system at any moment in time [43]. Moreover, the state at any time  $t$  should include enough information to determine the system behaviour for all later times  $\tau > t$ . In order to have a complete model, the state vector must include quantities that together describe the system uniquely at a certain time [40]. This means that two equal realizations of the state vector containing these representative quantities should correspond to only one state of the complete system. For example, the state vector of a system describing a particle may include the position and velocity of the particle.

The states of the system may also be *hidden* in the sense that they are not directly measured. A Solar Heat Plant model could for example include the state ‘temperature of the fluid in collector 1’, even when the temperature is only measured in collector 5 .

In the state space models discussed in this thesis there will be an explicit distinction between measured input variables, measured output variables, and ‘true’ or *internal* states. The model describes the evolution of the measured output variable and the internal state variable in two different equations. The first, called the *system equation*, describes the evolution of the state vector  $x_t$  in time. The second, called the *observation* or *measurement equation*, relates the state vector  $x_t$  to the vector of observable outputs  $y_t$ . In the following, only state space formulations of the continuous-discrete type will be discussed, meaning that the system equation will always be continuous while the measurements take place at discrete points in time. This is the most natural way to represent

physical systems, where we assume that the physical process occurs continuously in time, while measurements can in practice only be discrete.

The system is called *forced* when the evolution of the system state depends on measured *inputs*. State vectors will always be denoted by the letter  $x$ , inputs by the letter  $u$ , and outputs by  $y$ . The input  $u$  is always a known, deterministic variable obtained from measurements. To keep notation as compact as possible, the model's dependence on parameter vector  $\theta$  will not be explicitly written. The reader should keep in mind that functions and matrices may, in all cases presented, depend on the parameters.

**Definition 4.1** (Deterministic Linear State Space Model). The system equation is of the form

$$\dot{\mathbf{x}} = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) \quad (4.1)$$

where  $\mathbf{x}$  is the state vector,  $\mathbf{u}$  the input, and  $\mathbf{F}$  and  $\mathbf{B}$  are matrices that possibly depend on time. The observation equation is of the form

$$\tilde{\mathbf{y}}_k = \mathbf{H}(t)\mathbf{x}_k \quad (4.2)$$

where we denote  $y_{t_k} = y_k$  and similarly  $x_{t_k} = x_k$  for simplicity. In order to ensure the existence of a solution for this system, it is supplemented with an initial condition for the state vector

$$\mathbf{x}(t_0) = \mathbf{x}_0. \quad (4.3)$$

It should be noted that the linear observation equation can be generalized by adding a term  $\mathbf{D}(t)\mathbf{u}(t)$  to the right hand side of equation (4.2). However, such a linear system can always be rewritten to the form without this term by redefining  $\tilde{y} = y - \mathbf{D}(t)\mathbf{u}(t)$ . In the rest of this thesis, the input-dependent term in linear measurement equations will therefore be omitted. The model structure can be generalized by allowing any function of  $x, u$  in the system and observation equation, as follows.

**Definition 4.2** (Deterministic General State Space Model). The system equation is of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \quad (4.4)$$

where  $\mathbf{x}$  is the state vector,  $\mathbf{u}$  the input, and  $f$  a function with vector-form output. The observation equation is of the form

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k) \quad (4.5)$$

and the initial condition is

$$\mathbf{x}(t_0) = \mathbf{x}_0. \quad (4.6)$$

A state space model explicitly uses information about the system state for the description of its dynamics and is therefore not of the input-output type, but rather the input-state-output type. This feature is not present in many other statistical models (such as regression models), which relate input and output directly. However, when the matrices in Definition 4.1 do not depend on time, this state space formulation can be shown to be equivalent to linear, time-invariant transfer function models [41].

This common model formulation uses Ordinary Differential Equations (ODEs) to describe a dynamical system. State space formulations are frequently used for control purposes. However, the deterministic modelling of physical systems has several shortcomings. The next section presents this discussion as a motivation for developing a stochastic state space model form.

## 4.2 Shortcomings of deterministic state space models

An interesting discussion of the intuitive arguments for adding stochasticity to a model is found in Maybeck [41]. Here it is argued that “to assume perfect knowledge of all quantities necessary to describe a system completely and/or to assume perfect control over the system is a naive, and often inadequate, approach”. One reason for doubt is that a deterministic ODE formulation allows for an infinite deterministic solution of the state given an initial condition. In practice it would never be considered feasible for a model to give predictions for infinite time.

There are several reasons why the system state predicted by a model can differ from reality:

1. Models leave out the effect of some variables. This may be because it is practically impossible to include them, they are not measured, or the effect of such variables is negligible
2. Models are inherently simplifications of reality. Even when the effect of a certain input is included in the model, the form of the equation is an approximation of reality. Equations from physics are not universal truths, but rather adequate approximations of what is observed in nature
3. Measured inputs  $u$  are subject to noise and uncertainty
4. The system state itself is influenced by uncontrolled, random disturbances

This calls for the addition of a stochastic term to the state equation. Moreover, the measurement of output variables is subject to noise, which could be modelled by adding a noise term to the measurement equation as well. As we will see in the stochastic model formulation the next section, the errors have distinct meanings. The measurement error accounts for deviations of a measured value from the ‘true’ value of the output, while the system error represents actual deviation of the system state  $X$  from its deterministic value.

A particularly desirable result of including system or process noise in the model is that it accounts for the auto-correlation of prediction errors often seen in deterministic models. As these models do not take into account that the error in the previous step may have been a deviation in the actual state of the system rather than just a measurement error, it cannot include the error information in its prediction for the next step. This can for example lead to a situation in which an overestimation of the state at time  $t_k$  will likely be followed by another overestimation of the state (or negative error) at time  $t_{k+1}$ . This is depicted in Figure 4.1.

When the system and observation equation include error terms, the state and output vector are stochastic and the model is a *stochastic state space model*. The next section will explore the form such a system should have in order to satisfy our needs.

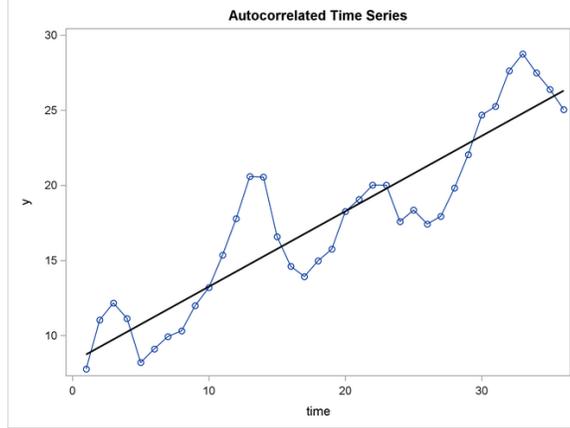


Figure 4.1: Auto-correlated errors, source = [http://support.sas.com/documentation/cdl/en/etsug/66840/HTML/default/viewer.htm#etsug\\_autoreg\\_gettingstarted01.htm](http://support.sas.com/documentation/cdl/en/etsug/66840/HTML/default/viewer.htm#etsug_autoreg_gettingstarted01.htm)

### 4.3 Stochastic state space formulation

This section will introduce the stochastic state space model form. As argued in the previous section, two types of errors will be added to the model equations. A straightforward way to include the system noise is by adding a term to the system equation:

$$\dot{\mathbf{X}}_t = \mathbf{f}(\mathbf{X}_t, \mathbf{u}) + \mathbf{g}(\mathbf{X}_t, \mathbf{u}_t, t)\mathbf{Z}_t \quad (4.7)$$

where the probability law of the noise process  $Z_t$  should be known. The addition of a stochastic error term  $Z_t$  makes  $X_t$  itself a stochastic process. The equation models the evolution of the state's probability density over time, rather than the evolution of the state itself.

Unfortunately, a model of this general form is not guaranteed to be useful. The reason for this is that the probability distribution of  $X_t$  and  $Y_t$  cannot in general be obtained from measurements. If, for example,  $Z_t$  is uniformly distributed, very little can be known about the distributions of  $X_t$  and  $Y_t$  [41]. To obtain a useful model formulation, there are two requirements for the process  $Z_t$ . First of all, it should lead to a well-defined differential equation for  $X_t$  which can only be the case if the integral  $\int g(X_t, t)Z_t dt$  is defined in some way. Second, we would like the resulting  $X_t$  to be a Markov process in order to have a model that is useful for predictions. This property allows for predictions of  $X_t$  given a previous state value  $X_s$  (so  $s < t$ ) with knowledge of  $X_s$  only. If  $X_t$  were not Markov, additional information from states before  $X_s$  (so states  $\{X_j | j < s\}$ ) may be needed for this prediction, which would be impractical; or even future information on  $X_t$  may be necessary, which would be impossible in applications. If  $Z_t$  is correlated (i.e. does not have independent increments), the process  $\{X_t\}$  is not a Markov process [42]. However, when  $Z_t$  is assumed to be Gaussian white noise, it can be shown that the resulting  $X_t$  is a Markov process ([42], page 92).

By assuming that  $Z_t$  is white Gaussian noise ( $N_t$ ), which we recall has the property  $N_t \sim \frac{dW_t}{dt}$ , the equation can also be rewritten in a form that is mathematically well-defined. Then equation (4.7) can then be rewritten to a stochastic differential equation of the form

$$d\mathbf{X}_t = \mathbf{f}(\mathbf{X}_t, \mathbf{u}_t)dt + \mathbf{g}(\mathbf{X}_t, \mathbf{u}_t, t)dW_t \quad (4.8)$$

Note that this special case of equation (4.7) has meaning due to the development of stochastic integrals. In this formulation, the stochastic term is free to depend on the state  $X_t$ . In the rest of this thesis, the model form will be restricted to have function  $g$  independent of the system states.

Therefore, the Itô and Stratonovich interpretation of the integrals are equivalent. The diffusion term is also taken independent of the states in CTSM, because “this renders parameter estimation more feasible” [17]. As found by Nielsen and Madsen [44], the use of a Lamberti transformation makes it possible to incorporate some types of state dependency in the diffusion term, but this is outside the scope of this thesis as such models have not been used.

This type of system equation will now be used to define the type of state space model that will be used for the modelling part of this thesis, the continuous-discrete stochastic state space model. The bold notation for vectors will from now be dropped, but the processes  $X_t$ ,  $W_t$  and the functions  $f$ ,  $h$ , and  $g$  may still be vector-valued.

**Definition 4.3** (Continuous-discrete stochastic state space model). The system equation for state vector  $X_t \in \mathbb{R}^d$  is of the Itô SDE form

$$dX_t = f(X_t, u_t, t)dt + g(u_t, t)dW_t, \quad (4.9)$$

where  $f : \mathbb{R}^d \times [t_0, t] \rightarrow \mathbb{R}^d$  is the drift term and  $g$  is the diffusion term, a matrix valued function  $\mathbb{R}^d \times [t_0, t] \rightarrow \mathbb{R}^{d \times m}$ . The discrete measurement equation is of the form

$$Y_k = h(X_k, u_{t_k}, t_k) + \epsilon_k \quad (4.10)$$

where  $k$  is used as an index to indicate discrete numbered time points  $t_k$  for  $k = \{0, 1, \dots\}$ , which may be irregularly spaced in time. The measurement error  $\epsilon_k \sim N(0, S_k(u_{t_k}, t))$  is a white noise sequence of vector form. The variance of the measurement error,  $S_k$ , is thus free to depend on  $t$  and  $u_{t_k}$ . It is assumed that  $\{W_t\}$  and  $\{\epsilon_k\}$  are independent. It is further assumed that the initial state  $X_0$  is a random variable with a given probability density  $p_0(x)$ .

Note that the Itô interpretation of the SDE is chosen, as this is more suitable for parameter estimation than the Stratonovich interpretation [17].

We will introduce two special case of this general model, as solutions methods implemented in CTSM differ for the different model types.

**Definition 4.4** (Linear time varying (LTV) stochastic state space model). A continuous-discrete stochastic state space model is *linear time-varying (LTV)* if it can be written as a sum of matrix products with  $X$  and  $u$ . This corresponds to the following system and observation equation:

$$\begin{aligned} dX_t &= (A(u_t, t)X_t + B(X_t, u_t, t)u_t)dt + \sigma dW_t \\ Y_k &= C(u_{t_k}, t_k)X_{t_k} + \epsilon_k \end{aligned}$$

Note that the diffusion term  $\sigma dW_t$  is now assumed independent of  $u$  as well. Further assumptions are equal to those in the Definition of the general model (Definition 4.3).

**Definition 4.5** (Linear time invariant (LTI) stochastic state space model). A continuous-discrete stochastic state space model is *linear time invariant (LTI)* if the matrices are only functions of system parameters

$$\begin{aligned} dX_t &= (AX_t + Bu_t)dt + \sigma dW_t \\ Y_k &= CX_{t_k} + \epsilon_k \end{aligned}$$

Further assumptions are equal to those in the Definition of the general model (Definition 4.3).

Input measurements  $\{u_t\}$  are usually known at discrete times  $\{t_1, t_2, \dots\}$  only, while the continuous system equation requires a value for  $u_t$  at each time  $t$ . Two possible assumptions to solve this problem are implemented in CTSM: a *zero* or *first order hold* on  $u$ . Zero order hold keeps  $u$  constant until the next measurement is taken. When first order hold is used,  $u_s = u_k \frac{u_{k+1} - u_k}{t_{k+1} - t_k} (s - t_k)$ , corresponding to linear interpolation between measurements. Only zero order hold has been used in this thesis.

## 4.4 Motivation for using Stochastic State Space representation

The *stochastic* state space formulation has several advantages, which will be outlined here.

Stochastic state space models are sometimes also referred to as *grey box models*. As the name suggests, a grey box model is of a form in between a white box and a black box model, using both physical knowledge and available data. A white box model is entirely based on equations from physics, and does not include statistical variations and uncertainties. Once its parameters are determined (estimated or obtained from physics), it is therefore completely deterministic. A black box model is identified using data only, without taking physical meaning into account. Each type of model has its advantages and disadvantages, and the most suitable of the models therefore depends on the circumstances. When both physical knowledge of the system and data of its behaviour are available, a grey box model may be appropriate as it can take advantage of both sources of information. Kristensen, Madsen, & Jørgensen (2004) describe this mix as “a way of combining the advantages of both model types by allowing prior physical knowledge to be incorporated and statistical methods for parameter estimation to be applied”. It can be shown that the translation of the ODE model into a stochastic state space model does not affect the parameters of the drift term, which means that their physical interpretation is preserved [17]. Therefore, stochastic state space models provide a way to combine physical knowledge and statistics. The advantage of including the uncertainty explicitly is that statistical methods can be employed to provide estimates of the uncertainty of the model and its parameters. Based on the estimated uncertainties, possibly valuable model extensions can also be identified. Such procedures will be discussed in Section 9.5.

The stochastic state space formulation allows the uncertainty in the model to explicitly depend on system variables. Furthermore, it distinguishes between system (or process) error and measurement error explicitly. As said, the process noise can be shown to account for auto-correlation of prediction errors that are frequently observed in practice. Also due to this explicit handling of noise, the parameter estimates for grey box models tend to be less dependent on data, “because random effects due to process and measurement noise are not absorbed into the parameter estimates but specifically accounted for by the diffusion and measurement noise terms” [17].

Compared to black and white box models, the semi-physical grey box model is often economical in the number of parameters needed to attain a certain accuracy and of a lower computational complexity [28].

## Chapter 5

# State estimation in stochastic state space models

Given a set of measurements and a (stochastic state space) model formulation, the next step is to extract the value of quantities of interest: the system states. This chapter will deal with the question of estimating these system states in an optimal way.

Depending on the way the measurements are used for state estimation, we distinguish between filtering (or reconstruction), prediction, and smoothing (or interpolation). Assuming that a time series of measurements  $Y$  is available until time  $t_k$ , filtering concerns the problem of obtaining  $X_{t_k}$ , the system state at time  $\tau$ , from measurements  $\{Y_i | t_i \leq t_k\}$  obtained up until time  $\tau$ . This set of measurements up to  $t_k$  is denoted  $\mathcal{Y}_k$ . If the information in  $\mathcal{Y}_k$  is used to obtain an estimate of a future state  $X_\tau$  for  $\tau > t_k$ , the procedure is called prediction. Finally, if a past value  $X_\tau$  should be reconstructed (so  $\tau < t_k$ ), the process is called smoothing. The latter will not be discussed here.

The stochastic differential system equation as formulated in Definition 4.3 describes the propagation of a stochastic process  $X_t$  with probability distribution  $p(x_t, t)$ . The SDE is driven by a Wiener process which has Gaussian increments. Section 5.1 will present general properties of the propagation, which will form the basis for explicit filtering methods in the sections that follow.

In the special case of linear stochastic differential equations as in Definition 4.4, it can be shown that the propagated conditional density remains Gaussian for all  $t$ . This property was the basis for the development of the *Kalman Filter* for linear stochastic state space models. It was proposed by Rudolf Kalman in 1960 [45]. The method estimates the conditional probability density of the state  $X$ , given the measurements of  $Y$ . This density contains all information about the state of the system [42]. For linear systems, this method can be shown to be optimal in several way. The basic Kalman Filter and its optimality properties are treated in Section 5.2. For nonlinear models, the filtering procedure becomes more complicated, as in this case only the full probability density contains all possible information about the system state. An approximate filter to be used in such cases, the Extended Kalman filter, will be discussed in Section 5.3.

## 5.1 Filtering in general state space model

This Section will consider filtering in the most general state space model formulation. For convenience this nonlinear continuous-discrete model setup is repeated here

$$\begin{aligned} dX_t &= f(X_t, u_t, t)dt + g(u_t, t)dW_t \\ Y_k &= h(X_{t_k}, u_{t_k}, t_k) + \epsilon_k. \end{aligned} \quad (5.1)$$

The filtering approach discussed here will consider the state estimation in two steps. First of all, what would be the optimal *prediction* of the state  $X$  at time  $t_k$  using previously obtained measurements  $Y_k$  at all times up until  $t_{k-1}$ ? Secondly, how can the new measurement obtained at time  $t_k$  be used to *update* this prediction? In the following, the estimate of  $X_t$  given  $\mathcal{Y}_k$  for  $t > t_k$  will be denoted  $\hat{X}_{t|t_k}$ . The estimate of  $X_{t_k}$  after incorporating observation  $Y_k$  is written  $\hat{X}_{t_k|t_k}$ .

The most informative outcome of a filter would be to obtain the full probability density of  $X_{t_k|t_k}$  at each time  $t_k$ . From that, it would be possible to define estimators that are optimal with respect to a chosen criterion. The filter should start from an initial condition for the probability density of  $X_0$ , namely  $p(x, t_0)$ . We would in the prediction stage like to derive an expression for  $p(x, t|Y_{k-1})$  with  $t > t_{k-1}$ . The probability density of a stochastic process generated by an SDE of the form in Equation 5.1 satisfies the forward Kolmogorov equation. Thus, this distribution propagates from initial condition  $p(x, t_{k-1}|Y_{k-1})$  while satisfying

$$\frac{\partial p(x, t|\mathcal{Y}_{k-1})}{\partial t} = -\frac{\partial [p(x, t|\mathcal{Y}_{k-1}) \cdot f]}{\partial x} + \frac{1}{2} \frac{\partial^2 [p(x, t|\mathcal{Y}_{k-1}) \cdot g^2]}{\partial x^2} \quad (5.2)$$

This equation could in theory be used to derive an expression for the time propagation of the probability density in between measurements.

At an observation, Bayes' rule gives the expression

$$p(x, t_k|\mathcal{Y}_k) = \frac{p(y_k|x_{t_k}, \mathcal{Y}_{k-1})p(x, t_k|\mathcal{Y}_{k-1})}{p(y_k|\mathcal{Y}_{k-1})} \quad (5.3)$$

This expression should be rewritten to contain only known information from the prediction step. Because  $X_t$  is Markov,  $p(x, t|\mathcal{Y}_k) = p(x, t|Y_k)$ . Using the relation between joint and marginal densities, the denominator of equation (5.3) is

$$p(y_k|\mathcal{Y}_{k-1}) = \int p(y_k, x|\mathcal{Y}_{k-1})dx = \int p(y_k|x)p(x, t_k|Y_{k-1})dx \quad (5.4)$$

where the second equality follows from Bayes' rule.

However, as argued in Maybeck [41] (page 215), it is not generally possible to obtain the full propagated probability density. This holds both for the prediction and the update step. In the prediction step the problem is that the Kolmogorov forward equation cannot generally be solved analytically.

One might think a solution could be to find a propagation equation for some of the moments of the state  $X$ . However, as shown in Maybeck ([41] page 199), it is not possible "to generate proper differential equations for these moments themselves without knowledge of the entire density." In fact, the density  $p(x(t), t)$  needs to be known for all  $t$  just to obtain the first or second moment. In the linear case however, the problem is simplified and it is possible to obtain exact expressions for the propagation of the conditional probability density and/or estimates of the state's first and second moment. This approach is taken in the Kalman filter, which will be discussed in the next Section 5.2.

## 5.2 The Kalman Filter for linear time varying systems

This Section contains a derivation of the Kalman Filter, a method to do both prediction and filtering on a linear continuous-discrete state space model.

The Kalman filter considers the propagation of estimates between two measurement times,  $t_k$  and  $t_{k+1}$ . For a better intuitive understanding, the algorithm is split into a *prediction step* and an *update step*, following Maybeck [41] and Madsen [40]. In the prediction step, which takes place directly after  $y_k$  is measured, the algorithm estimates the system state  $X_{t_{k+1}}$  based on measured output values  $\mathcal{Y}_k = \{y_k, \dots, y_0\}$ . The corresponding estimator, denoted  $\hat{X}_{t_{k+1}|t_k}$ , is obtained by maximizing the state's conditional likelihood function. In the following update step, which takes place just after  $y_{k+1}$  has been observed, the estimate  $\hat{X}_{t_{k+1}|t_{k+1}}$  is obtained by correcting the predicted  $\hat{X}_{t_{k+1}|t_k}$  using the new information in  $y_{k+1}$ . As in the prediction step, the estimate  $\hat{X}_{t_{k+1}|t_{k+1}}$  is computed using the conditional likelihood function.

For convenience, the linear time-variant (LTV) model, of which the LTI model is a subcase, is repeated here.

$$\begin{aligned} dX_t &= (A(u_t, t)X_t + B(X_t, u_t, t)u_t)dt + \sigma dW_t \\ Y_k &= C(u_{t_k}, t_k)X_{t_k} + \epsilon_k \end{aligned} \quad (5.5)$$

$$\epsilon_k \sim N(0, S_k(u_{t_k}, t_k)) \quad (5.6)$$

In the following,  $C_{t_k}$  will be used as a shorthand for  $C(u_{t_k}, t_k)$ .

### 5.2.1 Prediction step

**Theorem 5.1.** *Between measurements, the conditional mean and conditional variance of the LTV model satisfy the ordinary differential equations*

$$\frac{d\hat{x}_{t|t_k}}{dt} = A\hat{x}_{t|t_k} + Bu_t \text{ for } t \in [t_k, t_{k+1}) \quad (5.7)$$

$$\frac{dP_{t|t_k}}{dt} = AP_{t|t_k} + P_{t|t_k}A^T + \sigma\sigma^T \text{ for } t \in [t_k, t_{k+1}) \quad (5.8)$$

where short hand notation is used for matrices in the LTV system and observation equations

$$A = A(u_t, t)$$

For a proof of this theorem, see for example Jazwinsky [42], Maybeck [41] or the CTSM reference guide [46].

These equations can be solved to have solution

$$\hat{x}_{t|t_k} = e^{\int_{t_k}^t A(s)ds} \hat{x}_{t_k|t_k} + e^{\int_{t_k}^t A(s)ds} \int_{t_k}^t e^{\int_{t_k}^s A(u)du} B(s)ds \quad (5.9)$$

and using short hand notation  $\Phi(t) = e^{\int_{t_k}^t A(s)ds}$  the solution for  $P$  is

$$P_{t|t_k} = \Phi(t)P_{t_k|t_k}\Phi_t^T + \Phi(t) \left[ \int_{t_k}^t e^{\int_{t_k}^s -A(u)du} \sigma(s)\sigma(s)^T e^{\int_{t_k}^s -A(u)du} ds \right] \Phi(t)^T \quad (5.10)$$

To obtain these solutions, integrals over matrix exponentials have to be evaluated. This is a subject of its own, and will not be discussed here. The derivation is presented in the CTSM mathematics guide. Section 7.3 will outline how CTSM deals with the computation of the integrals.

For LTV systems, the matrices are constant, and therefore equations (5.9) and (5.10) can be simplified by replacing the integrals in the exponents. For example,  $e^{\int_{t_k}^t A(s)ds}$  can be replaced by  $e^{At}$ .

Predictions for the next measurement  $\hat{y}_{k+1|k}$  and its predicted variance  $R_{k+1|k}$  are also computed from conditional expectation, which gives

$$\hat{y}_{k+1|k} = C_{t_{k+1}} \hat{x}_{t_{k+1}|t_k} \quad (5.11)$$

$$R_{k+1|k} = C_{t_{k+1}} P_{t_{k+1}|t_k} C_{t_{k+1}}^T + S_k(u_{t_{k+1}}, t_{k+1}) \quad (5.12)$$

## 5.2.2 Update step

Following Jazwinsky (reference [42] section 7.2), the change in the state  $x_{t_k}$  estimated after adding an observation  $y_k$  is computed using Bayes rule. The form of the update step was already given in Equation (5.3). The densities on the right hand side can now be evaluated explicitly, using that the densities of  $X$  and  $Y$  remain Gaussian in the LTV case.

The density  $p(x, t_k | \mathcal{Y}_{k-1})$  is given from the prediction step, a normal distribution with mean and variance given by  $\hat{x}_{t|k}$  and  $P_{t|k}$  from equations (5.9) and (5.10).

The density  $p(y, t_k | x_{t_k}, \mathcal{Y}_{k-1})$  is the density of the predicted measurement, conditioned all previous measurements and assuming that the realization of  $x_{t_k}$  is known to have a certain value, say  $x_{t_k} = \tilde{x}$ . By the measurement equation (5.5), which has a normally distributed error term, this conditional density is normally distributed with mean  $C(u_{t_k}, t_k) \tilde{x}$  and variance  $S_k$ .

The final density in Equation (5.3),  $p(x, t_k | \mathcal{Y}_{k-1})$ , does not assume that the realization of  $X_{t_k}$  is known. The density can be shown to be Gaussian as well, see for example Maybeck ([41], page 211). It is computed using the system equation and the independence of  $\mathcal{Y}_k$ ,  $\sigma$ , and  $\epsilon_k$ . This gives that  $p(y_k | \mathcal{Y}_{k-1})$  has mean and covariance equal to the predicted  $\hat{y}_{k|k-1}$  and  $R_{k|k-1}$  from equation (5.11). This derivation can be found in more detail in Maybeck ([41], page 211).

Combining these densities, it turns out that  $p(x, t_k | \mathcal{Y}_k)$  is Gaussian. This is not a trivial result, but follows from completing squares as done for example in Jazwinsky ([42], page 197). The mean and variance of the updated conditional density  $p(x, t_k | \mathcal{Y}_k)$  are given by

$$\hat{x}_{t_k|t_k} = \hat{x}_{t_k|t_{k-1}} + K_{t_k} (y_k - \hat{y}_{k|k-1}) \quad (5.13)$$

$$P_{t_k|t_k} = P_{t_k|t_{k-1}} - K_{t_k} C_{t_k} P_{t_k|t_{k-1}}^T \quad (5.14)$$

which together form the update equations. Here,  $K$  is the Kalman gain matrix defined as

$$K_{t_k} \equiv P_{t_k|t_k} C_{t_k}^T R_{k|k-1}^{-1} \quad (5.15)$$

## 5.2.3 Kalman filter optimality

It can be shown that the linear continuous-discrete system can be rewritten as a discrete-discrete system, by introducing a so-called *state transition matrix*. For the discrete-discrete version of the

Kalman filter, several optimality properties have been derived. For a proof of these properties, the reader is referred to Jazwinsky ([42] section 7.3).

When the Kalman filter is derived from a Bayesian point of view, the conditional mean arises as an optimal estimator. because it equals the mean, mode, and median of the propagated density function of  $X_t$  conditional on all previous measurements. [41] In addition, it has been shown that the estimator  $\hat{X}_t = \mathbb{E}[X_t|\mathcal{Y}_{t-1}]$  is the estimate that minimizes the Mean Square Error. In other words, the error of the estimator has the lowest possible variance among all linear estimators. The Kalman filter in the discrete-discrete LTI system minimizes the Mean Square Error (MSE) among all linear estimators.

The Kalman filter can be derived from a least square and a maximum likelihood perspective as well. Further, the estimate of the Kalman filter is the orthogonal projection of the true state onto the approximation space. This means that the state estimate and the error of this estimate are independent.

### 5.3 Filtering and Prediction in nonlinear models: An Extended Kalman Filter

The theory in this section is mainly based on standard works by Maybeck [41] and Jazwinsky [42], as well as the mathematics guide for CTSM [46] and a paper presenting CTSM [17].

A stochastic process driven by a nonlinear SDE is not Gaussian, and therefore it cannot be fully represented by its mean and covariance. However, “by sampling fast enough relative to the time constants of the system, it is then reasonable to assume that the conditional density is approximately Gaussian” [46]. In that case, the continuous-discrete Kalman filter may still be used for prediction and filtering.

An Extended Kalman Filter (EKF) is a version of the Kalman filter that is adapted to nonlinear state space models. The EKF version presented here uses the nonlinear equation to produce the most accurate *prediction* of the state, while an local linear approximation is used to produce an approximate version of the *update* equation. The EKF thus attempts to include nonlinearities into the model while still exploiting linear estimation techniques. Section 5.4 presents the Iterated EKF, which attempts to better include nonlinear effects in the filter.

For convenience, the nonlinear continuous-discrete model setup is repeated here.

$$\begin{aligned} dX_t &= f(X_t, u_t, t) + \sigma(u_t, t)dW_t \\ Y_k &= h(X_k, u_k, t_k) + \epsilon_k \end{aligned} \tag{5.16}$$

#### The prediction step

For the prediction step, the model is linearized around the most recent prediction of the state  $X_{t|k}$ . The approximation is written

$$f(X_t, u_t, t) \approx f(X_{t|k}, u_t, t) + \tilde{A}(t) \cdot (X_t - \hat{x}_{t|k}) \tag{5.17}$$

where

$$\tilde{A}(t) = \frac{\partial}{\partial X} f(X, u_t, t)|_{X=\hat{x}_{t|k}} \tag{5.18}$$

The linearized version of the system equation can be written in the LTV format with  $A = \tilde{A}(t)$  and  $B = f(X_{t|k}, u_t, t) - \tilde{A}(t)\hat{x}_{t|k}$ . Using the equations from Theorem 5.1, predictions for the conditional mean and variance are then given by

$$\frac{d\hat{x}_{t|k}}{dt} \approx f(X_{t|k}, u_t, t) \text{ for } t \in [t_k, t_{k+1}) \quad (5.19)$$

$$\frac{dP_{t|k}}{dt} \approx \tilde{A}(t)P_{t|k} + P_{t|k}\tilde{A}(t)^T + \sigma\sigma^T \text{ for } t \in [t_k, t_{k+1}) \quad (5.20)$$

These ODEs for the propagated conditional mean  $\hat{x}_{t|k}$  and covariance  $P_{t|k}$  are now approximations, whereas they were exact for the LTI and LTV systems.

## The update step

The update equations of the extended Kalman filter are of the same form as for the LTV system, as given in equation (5.13). The Kalman gain matrix includes the matrix  $C$ , which for the nonlinear system is approximated as

$$\tilde{C}_k(\hat{X}_{t_k|t_{k-1}}, u_{t_k}, t_k) = \frac{\partial}{\partial X} h(X, u_{t_k}, t_k)|_{X=\hat{X}_{t_k|t_{k-1}}} \quad (5.21)$$

Similar to the matrix  $\tilde{A}(t)$ , this corresponds to a linearization of the observation equation about the most recent state estimate, which in the case of the update step is the predicted state after the previous measurement.

The Gaussianity of the updated conditional density is now an approximation. If the observation equation contains strong nonlinearities in the state variables, the approximation may be too crude and the filter may not perform well.

## 5.4 Iterated Extended Kalman Filter

In the prediction step of the extended Kalman filter, the matrix  $\tilde{A}(t)$  and  $\tilde{C}_k$  are linearizations around the most recent state estimate. To reduce the inaccuracy of this approximation due to the influence of nonlinearities in the function  $h$ , an iterative scheme may improve the approximation of matrix  $\tilde{C}$  used in the update step. This option is implemented in CTSM as well.

Suppose we have measured  $y_k$ , and predicted  $\hat{x}_{t_{k+1}|t_k}$ . The latter is used in the computation of  $\tilde{C}_{k+1}$ , which combined with measurement  $y_{k+1}$  allows for the updated estimate  $x_{t_{k+1}|t_{k+1}}$ . As more information is included for the updated estimate  $\hat{x}_{t_{k+1}|t_{k+1}}$ , it is probably closer to the ‘true’ value of the state at time  $t_{k+1}$ . Therefore, if  $\tilde{C}_{k+1}$  is computed by evaluating the derivative in Equation (5.21) in  $\hat{x}_{t_{k+1}|t_{k+1}}$ , the approximation of the updated state may improve. This new  $C$  may then be used to perform the update step again, perhaps leading to a state estimate  $\eta$  different from the original estimate  $\hat{x}_{t_{k+1}|t_{k+1}}$ . This new estimate may again be used to recompute  $\tilde{C}$ , and the procedure is repeated until the new estimates of the updated state no longer differ from one another (i.e. not more than some small threshold number  $\approx 10^{\dots}$ ). The procedure is called the *Iterated Extended Kalman Filter*.

This repetition of the update step, starting from  $\eta_1 = \hat{x}_{t_{k+1}|t_k}$  can be written as:

$$\eta_{i+1} = \hat{x}_{t_{k+1}|t_k} + K_{k+1}^i (y_{k+1} - \hat{y}_{k+1|k}^i) \quad (5.22)$$

where

$$\hat{y}_{k+1|k}^i = h(\eta_i, u_{t_{k+1}}, t_{k+1}) \quad (5.23)$$

$$C_i = \frac{\partial}{\partial X} h(X, u_{t_{k+1}}, t_{k+1}) \Big|_{X=\eta_i} \quad (5.24)$$

$$K_{t_{k+1}}^i = P_{t_{k+1}|t_k} C_i^T R_{k+1|k}^i{}^{-1} \quad (5.25)$$

$$R_{k+1}^i = C_i P_{t_{k+1}|t_k} C_i^T + S_{k+1} \quad (5.26)$$

The updated covariance matrix  $P_{t_{k+1}|t_{k+1}}$  can be computed when the iterations have finished at iteration  $N$  by

$$P_{t_{k+1}|t_{k+1}} = P_{t_{k+1}|t_k} - K_{t_{k+1}}^N R_{k+1|k}^N (K_{t_{k+1}}^N)^T \quad (5.27)$$

## Chapter 6

# Maximum Likelihood Estimation in Stochastic State Space models

This section is largely based on the CTSM-R reference guide [47] and an article by Madsen et al. [17].

Maximum likelihood theory can be used to estimate parameters in a state space model. Maximum Likelihood Estimation (MLE) makes use of the likelihood function as a measure of how likely a set of parameters is to have produced the data, given a model. The likelihood function equals the joint probability density of this set of measurements, viewed as a function of the parameters  $\theta$  while the data is viewed as fixed. The MLE estimator  $\hat{\theta}_{\text{MLE}}$  is defined as the set of parameter values that maximizes the likelihood function. In the following, this will be formalized.

Suppose a measured time series of the output state vector is available, written as

$$\mathcal{Y}_N = \{y_0, y_1, \dots, y_N\} \quad (6.1)$$

For such a set, the likelihood function  $L(\mathcal{Y}_N, \theta)$  equals the joint probability of these measurements:

$$L(\theta; \mathcal{Y}_N) = p(\mathcal{Y}_N | \theta) \quad (6.2)$$

As we view the likelihood as a function of the parameters, rather than a function of the data, the parameter vector  $\theta$  is part of the equation explicitly. The equality  $P(A, B) = P(A|B)P(B)$  can be used to rewrite the joint probability density in Equation (6.2) as a product of the probability densities of each measurement individually.

$$L(\theta; \mathcal{Y}_N) = p(y_N | \mathcal{Y}_{N-1}, \theta) p(\mathcal{Y}_{N-1} | \theta) = p(y_0 | \theta) \cdot \prod_{j=1}^N p(y_j | \mathcal{Y}_{j-1}, \theta) \quad (6.3)$$

Then the maximum likelihood estimator is defined as

$$\hat{\theta}_{\text{MLE}} = \arg \max_{\theta \in \Theta} \{L(\theta; \mathcal{Y}_N)\} \quad (6.4)$$

To evaluate this likelihood function for a given  $\theta$ , the conditional densities in Equation (6.3) have to be found. As discussed, this density can generally not be determined exactly.

The Kalman filter and extended Kalman filter are therefore used to iterate through the data sequence and estimate the conditional mean and covariance at each time point. As these filters assume normality of the conditional densities, which are completely characterized by their mean and covariance,

the likelihood function can be rewritten to

$$L(\theta; \mathcal{Y}_N) = \left( \prod_{k=1}^N \frac{\exp(-\frac{1}{2} \epsilon_k^T R_{k|k-1}^{-1} \epsilon_k)}{\sqrt{\det(R_{k|k-1})} (2\pi)^l} \right) p(y_0|\theta) \quad (6.5)$$

where  $l$  is the dimension of  $y$ . Note that this expression is exact for LTI and LTV systems, but an approximation for nonlinear systems. The prediction errors  $\epsilon_k = y_k - \hat{y}_{k|k-1}$  and their covariance matrices  $R_{k|k-1}$  are computed by the Kalman filter. If we now condition the likelihood function in equation (6.5) on  $y_0$ , the probability  $p(y_0|y_0, \theta) = 1$  and drops out of the expression. Consequently taking the negative logarithm results in the loglikelihood function

$$-\ln(L(\theta; \mathcal{Y}_N)|y_0) = \frac{1}{2} \sum_{k=1}^N \left( \epsilon_k^T R_{k|k-1}^{-1} \epsilon_k + \det(R_{k|k-1}) \right) + \frac{1}{2} N \cdot l \cdot \ln(2\pi) \quad (6.6)$$

As the logarithm is strictly increasing, the logarithm of any function assumes its optimal values at the same values as the original function. The loglikelihood can therefore be used instead to find the MLE estimator as defined in equation (6.4). In fact, CTSM uses the following form of the optimization problem

$$\hat{\theta}_{\text{MLE}} = \arg \min_{\theta \in \Theta} \{-\ln(L(\theta; \mathcal{Y}_N))\} \quad (6.7)$$

This form is often favoured over the likelihood in equation (6.5) for numerical considerations: as the likelihood is a product of many probability densities, its value is often very close to zero. The larger the data set, the smaller it will be, leading to floating point precision related problems. The logarithm will convert these small values into moderately large (negative) numbers, which will enhance the precision of the method.

The negative log-likelihood function will in the following also be referred to as the *objective function*, as it will be the objective function for the optimization procedure.

## 6.1 Adapted likelihood function for robustness

It is desirable for a model and its parameter estimation procedure that the outcome is not too sensitive to small deviations in assumptions or in the data set. A few outliers should influence the outcome of the estimation as little as possible. However, it is unknown which measurements are outliers and which show actual meaningful deviation, perhaps due to unmodelled system behaviour. Still, the estimation procedure should be protected in some way against the influence of outliers.

Outliers have a large value of  $\epsilon_k$ , the prediction error. As the likelihood function is quadratic in  $\epsilon_k$ , outliers have a large influence on the value of this likelihood function. This is not desirable, as this favours the parameter values that fit these outliers well. One solution is to modify the objective function for values of the quadratic term  $\epsilon_k^T R_{k|k-1}^{-1} \epsilon_k$  above some threshold value  $c^2$ . As suggested by Huber, the likelihood of this outlier can be made linear in  $\epsilon_k$  above this threshold [48]. The resulting robust likelihood for observation  $k$  is of the form

$$\phi(\epsilon_k^T R_{k|k-1}^{-1} \epsilon_k) = \begin{cases} \epsilon_k^T R_{k|k-1}^{-1} \epsilon_k & \text{if } \epsilon_k^T R_{k|k-1}^{-1} \epsilon_k < c^2 \\ c \left( 2\sqrt{\epsilon_k^T R_{k|k-1}^{-1} \epsilon_k} - c \right) & \text{if } \epsilon_k^T R_{k|k-1}^{-1} \epsilon_k \geq c^2 \end{cases} \quad (6.8)$$

This function is continuous and continuously differentiable, and the optimization of the objective function remains feasible [46].

# Chapter 7

## The CTSM-R library

This chapter will discuss some aspects of the methods implemented CTSM, the R package that was used for this thesis.

The package *Continuous Time Stochastic Modelling for R* (CTSM-R or short CTSM) has been developed by DTU Compute as a tool to estimate parameters in state space models from data. It further contains functions to simulate, predict, and filter the state and output variables from new data sets, using the fitted model. The package is freely available at <http://ctsm.info/>.

The performance of the implemented algorithms has been tested in several applications [11–14]. Furthermore, an article by Kristensen, Madsen, and Jørgensen compares the performance in parameter and state estimation of CTSM to another tool, with good performance results for nonlinear systems with significant diffusion terms [17].

Documentation in the form of a user guide and reference manual are available at the CTSM website. This chapter is further based on an older version of the Mathematics Guide [46], and the article by Kristensen, Madsen, and Jørgensen (2004, [17]) which also presents the mathematics behind most of CTSM's methods.

### 7.1 Filtering methods

CTSM distinguishes between LTI, LTV, and nonlinear models in the same way as they were defined in chapter 4. The filtering method depends on the type of model. For LTI models, the basic Kalman filter can be used, whereas the other two models require usage of the Extended Kalman Filter. More specifically, the iterative EKF is used. The EKF is exact for LTV models, whereas it is an approximation for nonlinear models.

The iterations between measurements continue until the maximum number of iterations is reached (default is 10), or when the desired tolerance of the measurement equation is met (default is  $10^{-12}$ ).

## Filtering initialization

Initial states should either be supplied, or estimated along with the other parameters. In this thesis only the latter method has been used.

The initial state covariance is computed as:

$$P_0 = P_s \int_{t_0}^{t_1} e^{As} \sigma \sigma^T (e^{As})^T ds, \quad (7.1)$$

where  $A = A(\hat{x}_{t_0}, u_0, t_0)$  or in the nonlinear case it is defined as the Jacobian of  $f(x, u, t)$  evaluated at  $t_0$ . The form of this equation for the initial covariance is derived from the prediction step as presented in Equation (5.10), where the previous covariance  $P_{t_j|t_j}$  is set to zero. The prediction of the initial covariance is scaled by a pre-specified scaling factor  $P_s \geq 1$  (default = 1).

## 7.2 Numerical stability of the Kalman filter

The Kalman filter can be unstable due to rounding errors. Covariance matrices are by definition symmetric and positive definite, but due to rounding errors they can become non-positive definite because of rounding errors. To prevent this from happening, CTSM replaces all covariance matrices by their square root free Cholesky decompositions:  $P = LDL^T$ . Here  $L$  is a lower triangular matrix and  $D$  is a diagonal matrix with strictly positive entries. This decomposition is known to exist for all symmetric positive definite matrices. It can be computed without taking square roots, as opposed to the regular Cholesky decomposition.

To ensure that the updated covariance matrix remains positive, a method proposed by Thornton and Bierman (1980, [49]) is implemented in CTSM. This method is based on the knowledge that factorized covariance matrices can be updated as

$$P' = P + GD_g G^T \quad (7.2)$$

where  $G$  is a full matrix, and  $D_g$  is diagonal. To ensure that  $P'$  will be positive definite, it is found by solving the equation:

$$(P' =) L' D' (L')^T = P + GD_g G^T \quad (7.3)$$

for a lower triangular  $L'$  and a diagonal  $D'$  with positive diagonal values. The Cholesky-factorized form of the covariance matrix allows for an easy evaluation of  $\det P$  as well, from the insight that  $\det(LDL^T) = \sum_i d_{ii}$ . This determinant is needed in the objective function, i.e. the loglikelihood function.

## 7.3 Evaluation of matrix exponentials and their integrals

To obtain estimates of the states in the Kalman filter, integrals over matrix exponentials have to be evaluated. These integrals were presented in Chapter 5. As will not be shown here, the integrals

can be rewritten so that the final integrals that need to be computed are:

$$\begin{aligned} & \int_0^{\tau_k} e^{A \cdot s} s ds \\ & \int_0^{\tau_k} e^{A \cdot s} ds \\ & \int_0^{\tau_k} e^{A \cdot s} \sigma \sigma^T (e^{A \cdot s})^T ds \end{aligned} \tag{7.4}$$

The evaluation of such integrals is a difficult task. First of all, it is difficult to find the exponent of a matrix in an efficient way, and much research has been done to find methods for doing so. For LTV systems, it is assumed that the matrices are constant between samples, so that the integrals of exponents can be replaced by  $e^{At}$ . To ensure that this assumption does not influence the solution too much, the time interval between measurements is sub-sampled and at each sub-interval the matrices are recomputed.

A Pade approximation with repeated scaling and squaring is used to approximate the matrix exponential  $e^{As}$ , as presented by van Loan (1978, [50]). The implementation in CTSM is based on the implementation of Sidje (1998, [51]), which will not be further discussed here. If  $A$  is singular, a singular value decomposition of  $A$  is performed to transform the matrix and thereby the integrals. To determine efficiently whether or not  $A$  is singular, it is tested whether the reciprocal condition number:

$$\kappa^{-1} = \frac{1}{|A||A^{-1}|} \tag{7.5}$$

is below a threshold (of around  $10^{-12}$ ). The condition number is computed using the 1-norm of the matrix  $A$  and an estimate of the 1-norm of  $A^{-1}$ . The expressions obtained for the integrals are omitted here as they are lengthy and do not illustrate the method further; for further information see [46]. Separate expressions are obtained for the basic and extended Kalman filter.

## 7.4 Optimization details

For parameter estimation, the objective function as given in Equation (6.7) needs to be minimized. CTSM uses a quasi-Newton method to solve this nonlinear optimization task. To solve

$$\min_{\theta} \mathcal{F}(\theta) \tag{7.6}$$

the algorithm aims to find  $\theta$  such that the gradient of  $\mathcal{L}$  satisfies

$$g(\theta) \equiv \frac{\partial}{\partial \theta} \mathcal{F}(\theta) = 0. \tag{7.7}$$

Such minima are sought in an iterative manner. Assume that  $\theta_{i+1} = \theta_i + \delta$ . By computing the first order Taylor expansion of  $g$  around the current estimated solution  $\theta_i$ , and assuming that  $g(\theta_{i+1}) = 0$ , the following relations hold:

$$g(\theta_i + \delta) = g(\theta_i) + \frac{\partial g(\theta)}{\partial \theta} \Big|_{\theta=\theta_i} \delta = 0 \tag{7.8}$$

which can be solved for  $\delta$ , while denoting the partial derivative in the previous equation by  $H_i$ :

$$\delta = -H_i^{-1} g(\theta_i) \tag{7.9}$$

When the approximations have been made and  $\delta$  has been found, the next parameter estimate is computed as a scaled step from  $\theta_i$  in the direction of  $\delta$

$$\theta_{i+1} = \theta_i + \lambda\delta \tag{7.10}$$

for some factor  $\lambda > 0$  such that the objective function is decreased at the new estimate: i.e.  $\mathcal{F}(\theta_{i+1}) < \mathcal{F}(\theta_i)$ . Finding a suitable value of  $\lambda$  is a *soft line search* in the direction of  $\delta$ . For more information on this procedure, the reader is again referred to the CTSM mathematics guide. Alternatively, the book by Dennis and Schnabel (1983, [52]) provides further information and proofs.

## 7.5 Parameter estimation

The parameters are estimated by solving the optimization problem as given in Equation (6.7), maximizing the loglikelihood. The objective function is modified so that known upper and lower bounds for parameters can be set. In applications there are often certain limits for the parameter vector  $\theta$  to be of physical meaning. Let these boundary vectors be given by  $\theta^{min}$  and  $\theta^{max}$ . For each parameter  $\theta_i$ , rather than minimizing the objective function of the original parameter, the function is minimized with respect to a transformed parameter  $\theta_i^t$  defined as

$$\theta_i^t \equiv \log \left( \frac{\theta_i - \theta_i^{min}}{\theta_i^{max} - \theta_i} \right) \tag{7.11}$$

The optimization can be unconstrained for  $\theta_i^t$ , because the transformation ensures that the corresponding  $\theta_i$  will be within the bounds. This is seen from the inverse transformation

$$\theta_i = \frac{e^{\theta_i^t} \theta_i^{max} + \theta_i^{min}}{e^{\theta_i^t} + 1} \tag{7.12}$$

which for  $\theta_i^t \in (-\infty, \infty)$  has range  $(\theta_i^{min}, \theta_i^{max})$ . A problem is that for  $\theta_i$  close to either boundary, the finite difference derivative  $\theta_i^t$  may be close to zero due to the logarithm's symmetry in  $\theta_i^{min}$  and  $\theta_i^{max}$  (see Equation (7.11)). It then seems that the optimum is found.

In case the parameter estimate is close to the boundary, the boundary should be moved if it is not a hard boundary.

## 7.6 Built in statistics for evaluation of fit

After the parameters have been estimated, CTSM provides several statistics to evaluate the result.

### Estimate of parameter standard error

The uncertainty of parameter estimates can be approximated using the fact that the maximum likelihood estimator is asymptotically Gaussian with mean  $\theta$  and covariance  $H^{-1}$ . This  $H$  is the *Fisher information matrix*, a matrix with elements

$$H_{ij} = -\mathbb{E} \left[ \frac{\partial^2}{\partial \theta_i \partial \theta_j} L \right] \tag{7.13}$$

where  $L$  is the loglikelihood [53]. Since the true  $\theta$  is unknown,  $H$  needs to be estimated, which is done by evaluating the derivatives at the parameter estimates  $\hat{\theta}$

$$H_{ij} \approx - \left( \frac{\partial^2}{\partial \theta_i \partial \theta_j} L \right) \Big|_{\theta=\hat{\theta}}. \quad (7.14)$$

To obtain estimates for the uncertainties of individual parameters, a decomposition of the covariance matrix  $H^{-1}$  into a diagonal matrix  $\sigma_{\hat{\theta}}$  and a matrix  $R$  is used

$$H^{-1} = \sigma_{\hat{\theta}} R \sigma_{\hat{\theta}} \quad (7.15)$$

where  $R$  is interpreted as the correlation matrix (with entries  $|R_{ij}| \leq 1$ ), and  $(\sigma_{\hat{\theta}})_{ii}$  the standard deviation of parameter estimate  $\hat{\theta}_i$ .

The correlation matrix  $R$  can be studied on its own. A high correlation indicates that certain parameter *estimates* are highly interrelated, which is an indication that the different effects of these parameters on model outputs cannot be distinguished given the data.

## Significance of parameters

For each parameter, CTSM additionally outputs the t-value of a statistic testing its significance, and the probability that this parameter equals zero. The hypotheses for this test are

$$H_0 : \theta_j = 0 \quad \text{versus} \quad H_1 : \theta_j \neq 0 \quad (7.16)$$

The method is further based on the assumption that  $\theta_{MLE}$  is approximately Gaussian. Under this assumption and  $H_0$ , the following statistic is t-distributed

$$z(\hat{\theta}_j) = \frac{\hat{\theta}_j}{\hat{\sigma}_{\hat{\theta}_j}} \quad (7.17)$$

The t-distribution has degrees of freedom equal to the number of observations (excluding missing observations) diminished by the number of parameters. When the t-distribution's degrees of freedom go to infinity, it approaches the standard normal distribution.

The statistic  $z$  expresses how many standard deviations the estimated  $\hat{\theta}$  is away from zero. When  $z$  is large (in absolute sense), the probability that the null hypothesis holds is low. The probability to obtain a value of  $|z| \geq \tilde{z}$  can be calculated using the t-distribution's cumulative probability function. This function is however not known explicitly. Therefore, the probability of obtaining  $|z| \geq \tilde{z}$  is estimated using the transformed statistic  $z^N$ , which has a standard normal distribution under  $H_0$ :

$$z^N(\hat{\theta}_j) = z(\hat{\theta}_j) \cdot \frac{1 - \frac{1}{4DF}}{\sqrt{1 + \frac{z(\hat{\theta}_j)^2}{2DF}}} \quad (7.18)$$

From  $z^N$ 's cumulative probability function, we can compute the probability of obtaining the estimated value of  $\hat{\theta}_j$  under the assumption that it is zero. If this probability is lower than a certain significance level  $\alpha$ , often taken at 5%, the null hypothesis should be rejected, and the parameter is judged significantly different from zero.

## Evaluation of optimization results

Finally, the summary of a fit gives the user some insight in the shape of the objective function and the penalty function at the estimated value of  $\theta$ .

The derivative of the objective function with respect to  $\theta_i$ , named  $dF/dPar$ , shows the steepness of the objective function in the direction of each parameter. If this value is not close to zero, the optimum was steep, and may not be a true optimum.

The derivative of the penalty function, named  $dPen/dPar$ , should be low as well. If this value is high, one of the estimated parameter values may be close to its limits. The penalty function may in this case have influenced the estimated parameters. If possible, the constraints on the corresponding parameter should be loosened; otherwise the one parameter should be estimated in another way (using for example profile likelihood).

## 7.7 Using and testing a fitted model

The package offers functions to simulate, filter, and predict using a parameterized model on the data used for fitting, or on new data. In the case of new data, it is possible to provide a new initial value of the state estimates.

The function *simulate* allows for a simulation of the *mean* of state and output variables, along with their standard deviations. The estimates are generated by the (extended) Kalman filter without updating. These estimates show how the *mean* state and standard deviation evolve over time, rather than representing a single realization of the modelled states and outputs. The latter would be a *stochastic* simulation, which could be implemented by means of for example a stochastic Euler scheme. This is not implemented in CTSM now.

The CTSM-R package further contains functions for *filtering*, *smoothing*, and *prediction*. The predict-function allows for *n*-step-ahead prediction.

## Chapter 8

# Exploration and preparation of the data set

The main data set used for model building consists of measurements from the Aalborg CSP solar heat plant in Solrød. It covers the period from May 23 up to and including May 31 2017. The measurements from this period will be discussed in detail in Section 8.1. After model building, cross-validation is done on data from the same plant, dated 1-19 June, 1-7 August, 5-11 September, and 5-11 October 2017. These sets will not be discussed in detail. Section 8.2 outlines how the data has been manipulated for use in this thesis. This procedure was the same for all different data sets.

### 8.1 Data exploration

The data set contains many different measured quantities. The inputs used in this thesis are

- Total solar radiation in the plane of the collector
- Temperature set point for the control scheme
- Fluid flow through entire field
- Fluid inlet and outlet temperature in 4 different rows of the collector field
- Ambient temperature
- The control mode of the system (idle, warm-up, etcetera)

The time between values in the data set is 1 minute (although measurements at the plant are taken each second). All numerical values are thus 1 minute averages. Every hour, there is one exception to this, as the value for the 59<sup>th</sup> minute of each hour is missing due to a programming mistake in the system.

The most important quantities for this project are plotted as a full time series in Figure 8.1, and for an individual sunny day (Figure 8.2a) and cloudy day (Figure 8.2b). The diffuse and direct radiation included in the plot are computed from total radiation using a method described later in Section 10.4. All other quantities are measured.

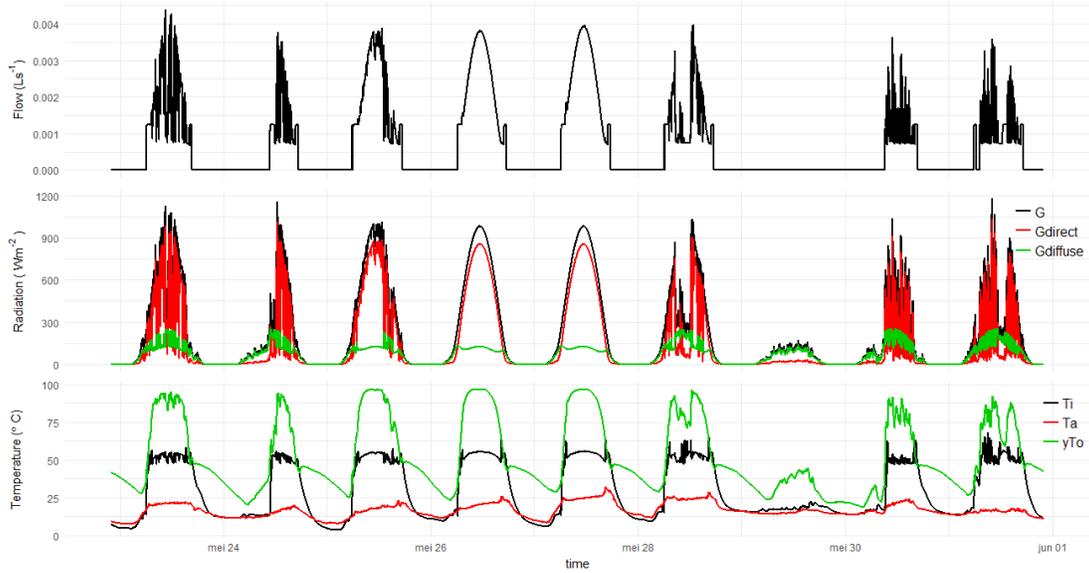


Figure 8.1: Overview of important variables in data set: fluid flow, radiation (total, diffuse, direct), temperatures ( $T_a$ ,  $T_i$ ,  $T_o$ ). Time series from May 23 up to May 31 2017

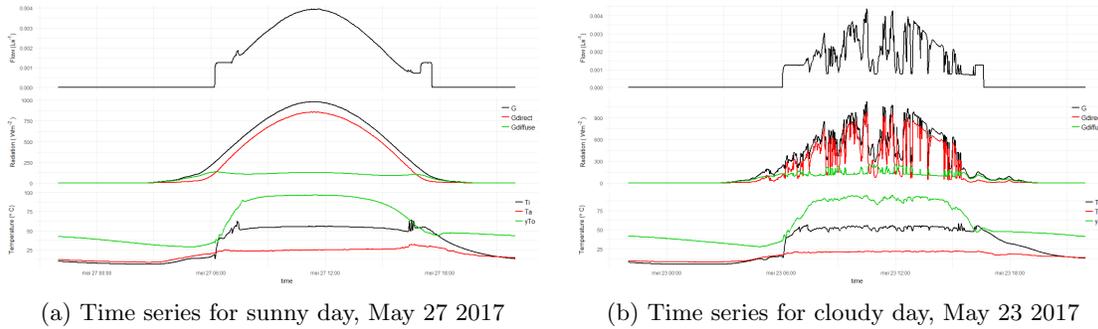


Figure 8.2: Overview of important variables in data set: fluid flow, radiation (total, diffuse, direct), temperatures ( $T_a$ ,  $T_i$ ,  $T_o$ ).

## Weather

Weather conditions varied over the measurement period. As seen in Figure 8.1, the 23, 24, 28, 30 and 31 of May have quite intermittent but high solar radiation intensity, implying partly cloudy days. On 25, 26, and 27 May the radiation is quite smooth, characteristic for clear sunny days with little to no clouds. May 29 was cloudy, and the plant was not operating.

## Mode of operation

The plant's control scheme switches between several modes in a 24-hour period. When global radiation is zero or below a fixed value, the system goes into idle mode. When the system restarts after a period of idleness, there is a preheating mode in which the flow is kept at a constant (low) value, until a certain fluid outlet temperature is reached. The preheating phase on the studied days typically lasts about an hour, but this time varies depending on weather conditions and the time of

day. In an hour of preheating, the fluid can go through the field around 3 times.

## Inlet, outlet, and ambient temperature

After the preheating phase, the inlet temperature remains fairly constant at  $55^{\circ}C$ . It however shows some more fluctuations on cloudy than on sunny days (compare Figure 8.2a and 8.2b). At the start and end of the day, when the preheating phase is activated, some fluctuations in inlet temperature occur.

The inlet temperature decreases to values around  $15^{\circ}C$  in several hours after the plant shuts down. In this it differs from the outlet temperature, which remains above  $25^{\circ}C$  during the night.

The outlet temperature is the controlled variable. Between 10 AM and 4 PM, it is close to its set point on most days. It fluctuates more on intermittent days than on clear sunny days.

The ambient temperature is between 8 and 15 degrees Celsius at night, and on most days increases rapidly in the morning to values from 15-25 degrees. It slowly increases further during the day, followed by a rapid decrease after sunset again.

## Time of passage

The flow in the main circuit is adapted between a minimum of  $7m^3/h \sim 1.9L/s$  and a maximum of  $40m^3/h \sim 11L/s$  at periods of high radiation. There is a minimum flow to ensure turbulent flow. The piping is designed so that the main circuit's flow divides equally over the eleven parallel rows.

The volume of one collector is 11.35 liter, so that a row of 20 collectors (neglecting connecting pipes between the panels) can contain around 230 liters of fluid. This means that it can take between 3 minutes (for high flow) and 20 minutes (in the preheating stage) for the fluid to pass through the rows of the solar field.

## 8.2 Data preparation

In order to prepare the data for model development and parameter estimation, a few steps have been taken:

- The time stamps were converted from local time to GMT (local time = GMT +2)
- For the 59<sup>th</sup> minute of each hour, a data point was added. For numerical values, this was done by linear interpolation between the nearest two data points, whereas factors were set equal to the previous value.
- The unit of the flow was converted from  $m^3/h$  to  $kg/(s \cdot m^2)$ , equal to the flow per aperture area of the solar field.
- The incidence angle between direct solar rays and the panel is added. The computation of this angle is described in Appendix A

In order to model the field as one compartment, one input and one output temperature measurement were chosen: those of the field's third collector row. The model may later be extended to describe the four measured rows separately, predicting four different output temperatures.

### 8.3 Data selection for parameter estimation

A subset of the data suitable for parameter fitting was selected. As the model is intended for usage during operation, the idle mode dynamics are of interest and these periods are therefore excluded. Moreover, this part of the data set has different dynamics than during operation: it reflects the cooling of a still pool of water. As a result of this selection step, one day is removed completely from the fitting data set, as the plant was not operating.

Furthermore, the preheating phase was removed from the data set. In this period the system dynamics are quite different from the rest of the day, mainly due to a more variable inlet temperature. Even when radiation is strictly increasing, the inlet temperature can oscillate slightly in this period, as some parts of the fluid have heated up more than others.

For the introduction of an Incidence Angle Modifier (IAM) as discussed in Section 3.1.1, the data was further filtered based on incidence angle. Data points with large incidence angle  $\theta$  such that  $\cos(\theta) < 0.2$  (or  $\theta > 78.5^\circ$ ) were excluded as well. This choice allows the parameter  $b_0$  to be of values in  $[0, 0.25]$  for the IAM to remain positive.

In the end of some of the days, the plant switches back and forth from the preheating to the operating mode. Such periods were also removed from the data set.

The resulting data set is in the remainder of this thesis referred to as  $D$ . The set contains 3900 values of each measured variable. The number of measurements included differs for each of the 7 days in the set. The fitting procedure will in some cases be repeated for single days from data set  $D$  for comparison. A typical sunny day will be taken as May 27th, denoted  $D_s$ , whereas a typical intermittent day will be May 23rd, denoted  $D_i$ . These dates will also be used for plotting model performance on different weather types.

## Chapter 9

# Model Selection and Extension

Model selection is a difficult step in the process of model development. The difficulty lies, among others, in a lack of objective criteria for model choices. Further, the quality judgement depends on the model's purpose. "An purely algorithmic and exhaustive selection procedure is seldomly feasible, hence iterative methods, in which the modeller is partly involved in the selection, are commonly applied" [11].

This Chapter will describe a method to perform model selection and extension on a sequence of stochastic state space models, which was proposed by Bacher and Madsen [11]. Figure 9.1 shows the model selection steps as presented in the paper by Bacher and Madsen.

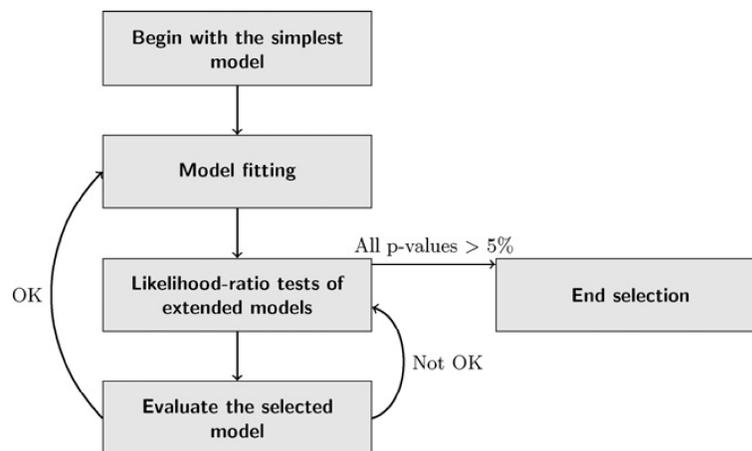


Figure 9.1: Model selection steps

## 9.1 First model formulation

An idea central to grey box models is to combine physical knowledge and statistical methods [53]. Therefore, an initial stochastic state space model formulation is usually based on differential equations known from physics. In a forward selection procedure, it is useful to start with the simplest possible model that produces reasonable output and follows the main dynamics of the process. A suitable starting model should provide insight in possible extensions and improvements. It is good practice to start by using a simple diffusion term, independent of inputs and states. This can be modified later on, if evidence for input or state dependent diffusion is present. However, if for example a mass balance is to be maintained by the model, the drift term may be made state dependent as done in for example [54].

The model should be formulated such that parameters are structurally identifiable. For example, if one parameter multiplies all of the others, the parameter set that maximizes the likelihood is not unique, and one parameters should be fixed.

## 9.2 Extended models: formulation and fitting

In a forward selection procedure, the model is iteratively extended. The extended models should be generalizations of the previous iteration's model outcome, and the extensions should be as small as possible.

### Parameter estimation & evaluation

After the extensions have been formulated, their parameters need to be estimated. The outcome of the estimation procedure can be evaluated separately. There are cases in which the estimation procedure does not result in a solution. This may be a result of several problems, such as an explosion of the state variable, a covariance matrix that is not positive definite, or slow convergence, so that the maximum number of iterations is reached before finding the solution [46]. This may be a first indication that a model is not suitable.

If the parameters are successfully estimated, it should be checked that their values are not too close to the user-set boundaries as described in Section 7. If the boundary is not a strict one, the parameter estimation should be repeated with relaxed boundaries for the parameter in question [46].

Once all parameters have been estimated well within the set boundaries, some other statistics may be of use. If the model is accurate enough for the assumption of Gaussian conditional probability densities to more or less hold (which will probably not be the case for the first modelling attempts), a t-test can be performed to determine the significance of parameters [53], as explained in Section 7.6.

## 9.3 Selection of best extension using likelihood ratio tests

The Likelihood Ratio Test is used as a first comparison of model performance. The most significant extension (with lowest p-value) is chosen. If several extensions have equal or similar p-values, the extensions with the highest loglikelihood *or* with the least parameters is chosen. The chosen model

will be evaluated based on other criteria in the next step of the model selection procedure. If none of the extended models is significant, the model selection procedure is terminated, and the current model is accepted.

The Likelihood Ratio Test (LRT) is a method to determine whether an extended model performs significantly better than a simpler version. When extending a model to include more parameters, the space over which the likelihood function can be minimized is extended. Therefore, the maximum of the likelihood function of the extended model will always be larger or equal to that of the reduced model. The likelihood ratio test uses hypothesis testing to determine whether the increase in likelihood is significant.

The LRT can only be used for nested models, where one is an extended version of the other. Assume the simple model has parameter values  $\theta_0 \in \Omega_0 \subset \mathbb{R}^{d_0}$ , and the extended model has  $\theta \in \Omega \subset \mathbb{R}^d$ . Then the simple model is a sub-model of the second if

$$\Omega_0 \subset \Omega \tag{9.1}$$

and  $d_0 < d$ . In other words, the simple model is a *special case* of the extended model, and can be obtained from the extended model by setting some parameters to zero.

As the null hypothesis it is assumed that the simple model is the ‘true’ model, and the true parameter set describing the data lies in the parameter space of the simple model, that is:

$$H_0 : \theta \in \Omega_0. \tag{9.2}$$

The alternative hypothesis equals

$$H_a : \theta \in \Omega \setminus \Omega_0 \tag{9.3}$$

A quantity used to test the hypothesis is

$$\Lambda(\mathcal{Y}_N) = \frac{\sup_{\theta \in \Omega_0} L(\theta, \mathcal{Y}_N)}{\sup_{\theta \in \Omega} L(\theta, \mathcal{Y}_N)} \tag{9.4}$$

As the notation shows, the statistic depends on the data. It expresses how many times less likely the measurements are observed under the best simple model (with optimal parameters) than under the extended model (with optimal parameters).

Under the null hypothesis, it has been shown that the test statistic

$$-2 \log(\Lambda(\mathcal{Y}_N)) = -2 \log \left( \frac{\sup_{\theta \in \Omega_0} L(\theta, \mathcal{Y}_N)}{\sup_{\theta \in \Omega} L(\theta, \mathcal{Y}_N)} \right) \tag{9.5}$$

converges in distribution to a  $\chi^2$  distribution with  $(d - d_0)$  degrees of freedom as  $N \rightarrow \infty$ .

This means that if  $N$  is large enough, the probability of the null hypothesis being true while obtaining the value  $x$  of the test statistic (i.e.  $-2 \log(\Lambda) = x$ ) can be computed from the cumulated probability function of this distribution. If the probability value is smaller than some significance level (for example 0.05 for a 5% significance level) then it is significantly more likely to observe the data with the extended model than with the simple, and the alternative hypothesis should be accepted.

The test cannot prove that  $H_1$  is actually true, it only indicates that  $H_1$  is more likely to hold than  $H_0$ , under the assumption that  $H_0$  holds and given the data.

## 9.4 Evaluation of selected model

As a first step, it should be checked that the model satisfies some basic requirements, e.g. positive states that remain finite [54]. Further, the selected model is validated/evaluated based on:

1. Residual white noise properties
2. Residual plots

If the model is judged as unsuitable, the next most significant model extension from the previous step should be selected and evaluated instead.

Depending on the application of the model, it may be important to check simulation,  $n$ -step-ahead prediction, or both. If the model will be used in a closed loop system, with intermediate measurements becoming available while predicting the state,  $n$ -step prediction will be important. This is for example the case when modelling for control purposes [53]. If the model intends to predict longer-term system behaviour (given predicted inputs), correct simulation is of greater importance for model evaluation.

The residuals are defined as:

$$\epsilon_k = \mathbf{y}_k - \hat{\mathbf{y}}_{m|k} \quad (9.6)$$

where  $m = k + n$  for  $n$ -step ahead prediction residuals. For residuals of simulation the value of  $k = 0$ , that is, the estimates are only based on known initial measurement of  $y_0$ .

### Checking white noise assumptions

Several of the methods used rely on the assumption that the residuals are white noise. If this is not the case (at least to some extent), parameter estimates will be biased, and the LRT is strictly speaking not valid. During the model sequence, we will not attach too much value to these properties, but will check as a validation step that they residuals are approximately white.

The white noise properties of the residuals can be tested both in the time and in the frequency domain. The (partial) auto-correlation function is used for the former, while the cumulated periodogram is used for the latter.

The sample auto-correlation function (ACF) and sample partial auto-correlation function (PACF) of the residuals are typically used to assess the predictive quality of the model. The value of these functions indicates the amount of variation in residual  $\epsilon_t$  that can be explained by variation in the lag- $k$  residual  $\epsilon_{t-k}$ . For formal definition of the ACF and PACF, as well as the cumulated periodogram, see for example [40]. This text also explains how confidence intervals can be computed for these diagrams.

For white noise, the ACF and PACF only show lag-dependencies outside the 5% confidence band for 1 in every 20 lags. If correlations between current and  $k$ -lagged values of the residuals exist, there will be a significant lag  $k$  dependency. In that case the predictive capabilities of the model could be improved, because the residual at time  $t - k$  apparently contains unused information that can help to improve the prediction at step  $t$  [53].

If there is no evidence against the hypothesis of white noise residuals, then the model may sufficiently

describe the data. There is no evidence that there are systematic errors in the model that can still be captured [46].

It must however be noted that in practice, residuals are almost never (completely) white. The residual white noise check is therefore not used as a model selection criterion in the model sequence of this thesis. Instead, we will check whether the final model's residuals are at least reasonably close to being white. Attaching too much value to having white(r) residuals does not generally benefit the development of a good model. It may for example happen that a beneficial model extension reduces the residual sizes, which makes smaller residual patterns more visible in the ACF and cumulative periodogram, thereby appearing to have worse white noise properties.

## Residual time series analysis

Time series of the residuals should be plotted and compared to time series of the in- and outputs. A good model's residuals depend on the in- and outputs as little as possible, which means that the the input variables' effect on the output variable are included in the model.

It is further useful to compare the residuals of the previous model and the extended model, in order to evaluate if the extension lead to the desired improvement.

## 9.5 Indications for model extension

First of all, identified relations between inputs and residuals through plotting them against one another can help the modeller to find useful extensions to the model. If the residuals show a clear relation to a variable, this may indicate that this variable's effects are not described (well) by the model.

Comparing residual and input/output time series plots can also be used to reveal remaining dependencies, and thereby indicate what effects are inaccurately captured by the model [11].

Large diffusion coefficients indicate model deficiencies in the corresponding state. The presence of significant parameters in the diffusion term is an indication that the corresponding drift term may be incorrect, which in turn provides an uncertainty measure that allows model deficiencies to be detected. [53]

If the model is accurate enough for the assumption of Gaussian conditional probability densities to hold (which will probably not be the case for the first modelling attempts, and perhaps not hold for the whole sequence), a t-test can be performed to determine the significance of parameters [53], as explained in Section 7.6. The presence of an insignificant parameter is an indication of model over-parameterization, and reducing the model may improve it. However, it is important to keep in mind that removal of insignificant parameters may be physically unreasonable or undesirable due to model requirements, in which case it may be retained despite the statistical testing [54]. Again, this is a subjective decision to be made by the modeller.

## Chapter 10

# Formulation of grey-box models for the Solrød solar heat plant

As described in the previous chapter, we will iteratively extend a basic model to find the best possible model for a SDHP. Extensions come in various forms: additional inputs variables can be included, additional parameters may be added, simple parameters may be turned into functions (of time, or of state and input variables). Finally, the spatial setup of the model may be altered by adding more state variables.

This chapter will introduce the different model extensions that have been considered in the model sequence for the solar heat plant. The resulting fitted models will be presented and evaluated in the next chapter. In particular, Section 12.1.3 will provide a motivation for the chosen extensions.

### 10.1 Base models $T_f$ and $T_o$

The initial models for the solar heat plant lump the entire field into one single compartment with one corresponding state variable. It is further assumed that the fluid temperature in the collector rows increases linearly between the inlet fluid temperature  $T_i$  and the outlet fluid temperature  $T_o$ , so that the average temperature in the compartment  $T_f = (T_o + T_i)/2$ . The first model equation is derived as done by Bacher, Madsen, and Perers [55] as described in Section 3.1.2. We however modified their equation to take into account that the current data set has variable inlet temperature  $T_i$ , as opposed to the data used by Bacher et al. This variable inlet temperature implies that the relation  $\frac{dT_f}{dt} = \frac{1}{2} \frac{dT_o}{dt}$  no longer holds.

To derive a state equation, the following heat balance is taken from Perers [32]

$$P = c_f Q_f (T_o - T_i) = F'(\tau\alpha)_{\text{en}} K_{\theta}(\theta) G - F' U_0 (T_f - T_a) - F' U_1 (T_f - T_a)^2 - (mC)_e \frac{dT_f}{dt} \quad (10.1)$$

This equation was discussed in Section 3.1.1, where the physical interpretation of each of the terms can be found.

## Model $T_f$

This model takes the average temperature in the collector row as the only system state ( $T_f$ ). We substitute  $T_o = 2T_f - T_i$  in the heat balance, and neglect the second order heat loss term. Rewriting gives state equation

$$dT_f = \frac{1}{mC} \left( F'(\tau\alpha)_{\text{en}}G + F'U_0(T_a - T_f) + 2c_fQ_f(T_i - T_f) \right) dt + \sigma_f d\omega_f \quad (10.2)$$

From the assumption  $T_f = (T_o + T_i)/2$  we derive observation equation

$$T_{o,t_k} = 2T_{f,t_k} - T_{i,t_k} + \epsilon_{\text{obs},t_k} \quad (10.3)$$

where the variance  $\mathbf{V}[\epsilon_{\text{obs}}] = \sigma_{\text{obs}}$ .

In the system equation (10.2), the measured quantities are  $T_a$ ,  $Q_f$ ,  $T_i$ , and  $G$ . The parameter  $c_f$  is further assumed given. The incidence angle  $\theta$  is a model input directly computed from the data. The quantities  $(mC)_e$ ,  $F'U_0$ ,  $F'(\tau\alpha)_{\text{en}}$ ,  $\sigma_f$  and  $\sigma_{\text{obs}}$  are parameters to be estimated. For a list of symbols, see the Nomenclature.

## Model $T_o$

Another way to derive a model equation from equation (10.1) is to approximate the  $T_f$  for the heat loss term by  $T_f = (T_o + T_i)/2$ , while assuming  $\frac{dT_f}{dt} \approx \frac{dT_o}{dt}$ . The outlet temperature  $T_o$  is the only system state.

The resulting system equation then becomes

$$dT_o = \frac{1}{(mC)_e} \left( F'U_0(T_a - T_f) + c_fQ_f(T_i - T_o) + F'(\tau\alpha)_{\text{en}}G \right) dt + \sigma_o d\omega_o \quad (10.4)$$

and the observation equation is

$$yT_{o,t_k} = T_{o,t_k} + \epsilon_{\text{obs},t_k} \quad (10.5)$$

where  $\mathbf{V}[\epsilon_{\text{obs}}] = \sigma_{\text{obs}}$ .

## 10.2 Including extra terms in heat balance

In the simplest model, the temperature dependence of the heat loss has been neglected. An extension to the models can thus be to include such effects by introducing  $F'U_1(T_f - T_a)^2$  into the equation.

Such models will have  $FU1$  as a part of their name, for example a model could be  $To\_FU1$ .

## 10.3 Modelling shading: splines

Spline functions will be used to split the global or direct beam radiation in  $K$  components, which peak during different times of day. For each component individually, we will estimate the efficiency

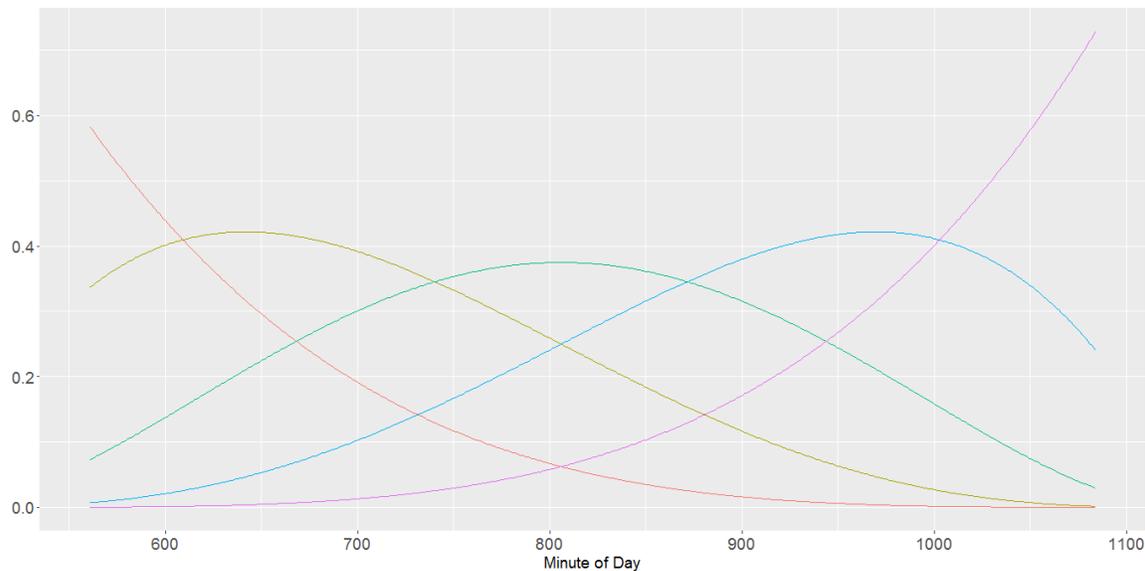


Figure 10.1: Constructed spline functions,  $K = 5$

parameter  $\eta_i$ . To illustrate the method, the splines on direct radiation will be used as an example here.

The splines are defined as B-splines based on Time of Day, with boundary knots at sunrise and sunset (for the May data set, this corresponds to the minute of day 429 and 1182). In a later model, the choice of boundary knots is changed to the moment of day where  $\cos(\theta) = 0.2$ . The splines had no pre-specified knots within the boundary knots. Depending on the desired number of splines, the degree of the used splines was altered. To illustrate the result, the spline functions in the 5-spline case are shown in Figure 10.1.

The total effective beam radiation at time  $t$  is then computed as

$$G_{b,\text{eff}}(t) = \sum_{i=1}^K s_i \cdot G_{b,i}(t) \quad (10.6)$$

where  $G_{dir,i}$  is the component of the radiation computed using spline functions  $S_i(t)$

$$G_{b,i}(t) = G_b(t) \cdot S_i(t) \quad (10.7)$$

The term in equation (10.6) either replaces the previous effective radiation term  $F'(\tau\alpha)_{\text{en}}G_b$  or it may replace this term *and* incidence angle dependencies together by replacing  $K_b(\theta)F'(\tau\alpha)_{\text{en}}G_b$ . Both options will be explored. The models with splines will have *Kspline* in their name, where  $K$  is the number of splines used for one solar day.

## 10.4 Splitting total radiation in its diffuse and direct components

As the name implies, direct beam radiation comes in a direct line from the sun. For sunny days with clear skies, most of the solar radiation is direct beam radiation. On overcast days, the sun is

obscured by the clouds and the direct beam radiation is zero. Diffuse radiation is scattered out of the direct beam by molecules, aerosols, and clouds. Because it comes from all regions of the sky, it is also referred to as sky radiation. The portion of total solar radiation that is diffuse is about 10% to 20% for clear skies and up to 100% for cloudy skies.

At many solar heat plants, including the one in Solrød, only the total radiation in the plane of the collectors is measured. It would be beneficial to measure the diffuse component of the solar radiation as well, but as such sensors are very costly it is often decided against this to reduce the installment costs. Moreover, the complex diffuse radiation measurement devices have a high maintenance cost [56].

In order to account for the different heating effects of the diffuse and direct components of solar radiation, we will use a method proposed by Ruiz-Arias et al. ([56], 2010) to split global radiation into an estimated diffuse and direct component. Note that in Solrød, not global but total *tilted* radiation is measured.

The method relies on a regression equation relating the diffuse fraction  $k_d$  to the so-called clearness index  $k_t$ . This clearness index and diffuse fraction are defined as

$$k_t = \frac{I_G}{I_0 \cos \phi_Z} \quad (10.8)$$

$$k_d = \frac{I_D}{I_G} \quad (10.9)$$

where  $I_G$  is the horizontal global radiation per  $m^2$ ,  $I_0$  is the extraterrestrial radiation per  $m^2$ ,  $I_D$  is the diffuse radiation per  $m^2$  in the horizontal plane, and  $\phi_Z$  is the zenith angle of the sun. Note that both  $k_t$  and  $k_d$  usually have values in  $(0, 1)$ , with some exceptions that will not be discussed here.

The extraterrestrial “clear sky” radiation can be computed as

$$I_G = 1367 \cdot (1 + 0.033 \cdot \cos(360 \cdot n/365)) \cos(\phi_Z) \quad (10.10)$$

where  $n$  is the day of the year. This formula is based on the rotation of the Earth around the Sun, which is ellipsoidal.

Ruiz-Arias et al. derive that the form of the relation between  $k_t$  and the diffuse fraction  $k_d$  is of the form

$$k_d = 0.952 - 1.041 \cdot e^{-e^{2.300 - 4.701 k_t}} \quad (10.11)$$

which is a sigmoid curve. The form of this equation was selected based on a derivation assuming uniform Rayleigh scattering, which involved Taylor expansions for the approximation of a logarithm and a factor  $1/k_t$ . The numbers in the equation were estimated specifically for Europe, based on data sets from 3 stations in Spain and 3 stations in Germany.

The sigmoid form was deemed reasonable based on scatter plots of  $k_t$  against  $k_d$  from data. In addition, sigmoid curves have the nice properties of being real-valued and differentiable with either a non-negative or non-positive first derivative and one inflection point. Under highly cloudy conditions, it would be physically reasonable for  $k_d \rightarrow 1$  as  $k_t \rightarrow 0$ . This condition is nearly fulfilled by the sigmoid curve, which takes value  $k_d \approx 0.95$  in this case. “On the other hand, for clear days, the diffuse fraction is expected to tend to small values, albeit strictly never equals to zero (note that even a completely clear atmosphere will scatter some amount of solar radiation).” The sigmoid curve takes values  $k_d = 0.001$  in this case.

The resulting split radiation is shown in Figure 10.2

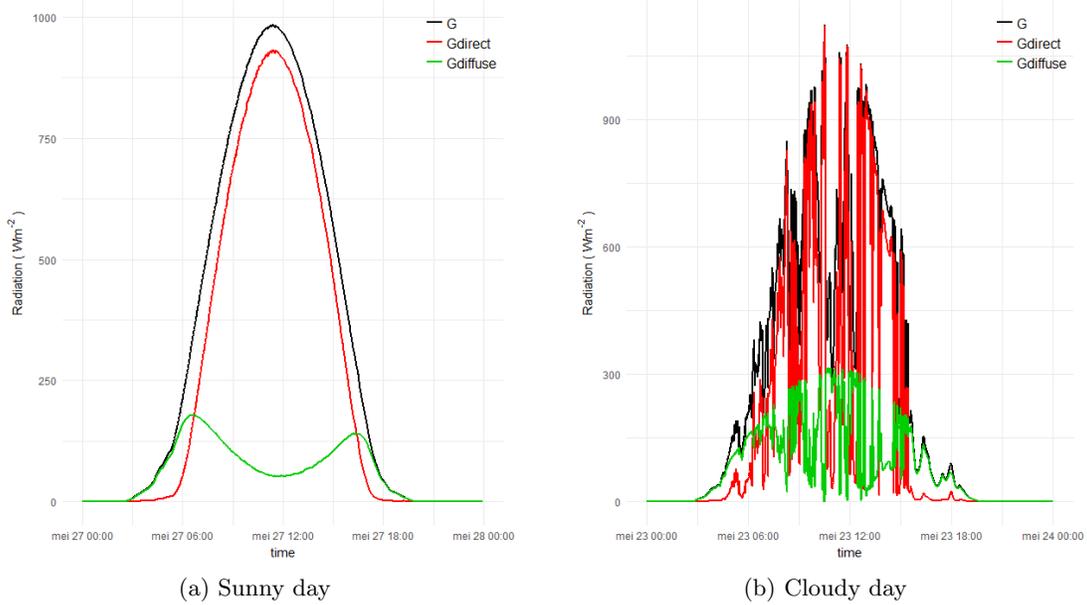


Figure 10.2: Split radiation into diffuse and direct components

The authors note that especially for intermediate values of the clearness index  $k_t$ , the possible values of the diffuse fraction  $k_b$  are in a wide range. “For instance, for a clearness index of 0.5 [...] the interval of diffuse fraction ranges approximately from 0.3 to 0.8” [56]. Other predictors such as temperature or humidity could improve the predictions, however, such measurements are often not available and the clearness index is the best we can obtain from available measurements.

The models extended with diffuse and direct radiation use the following substitute for the total radiation term in the heat balance

$$F'(\tau\alpha)_{\text{en}}G = F'(\tau\alpha)_{\text{en}}(K_b(\theta)G_b + K_dG_d) \quad (10.12)$$

which has additional parameters  $K_d$  and  $K_b(\theta)$ . Parameter  $F'(\tau\alpha)_{\text{en}}$  is set to 1 in some of the extended models, in order to have a unique solution to the parameter estimation procedure. However, when  $K_b(\theta)$  is replaced by an IAM function, parameter  $F'(\tau\alpha)_{\text{en}}$  can in some cases be varied again. Models that rely on this split will have the component  $G_{\text{split}}$  in their name.

## 10.5 Models with incidence angle modifier

The influence of the solar incidence angle on the absorbed radiation can be modelled using IAM functions, as discussed in Section 3.1.1. Such models replace either the parameter  $F'(\tau\alpha)_{\text{en}}$  (in case of models with unsplit total radiation  $G$ ) or the parameter  $K_b$  (in models with split  $G$ ) with a function of the solar incidence angle  $\theta$ . This angle is defined as the angle between the panel normal and the solar rays.

When an IAM function is included, the parameter  $F'(\tau\alpha)_{\text{en}}$  does not need to be fixed, because the function  $K_b(\theta)$  is restricted to specific values in the interval  $[0, 1]$ .

The next Section 10.5.1 will explore the form of the function  $K_b$  that would be suitable for the solar heat plant modelled in this thesis, by plotting the function for several values of its parameter  $b_0$

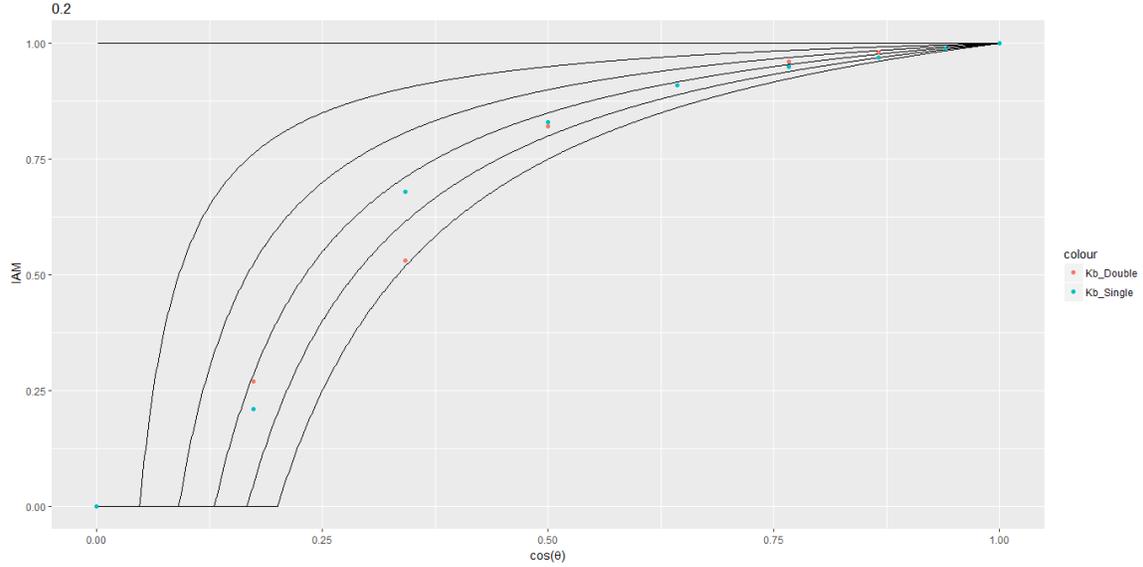


Figure 10.3: IAM functions for  $b_0 = \{0, 0.5, 0.1, 0.15, 0.2, 0.25\}$ . The dots represent values from the panel specifications, calculated using ISO standard test results.

against the values expected from panel specifications.

### 10.5.1 IAM based on panel specifications

The two panel types used in Solrød have specifications for the incidence angle correction for the incidence angles  $\{0, 20, 30, 40, 50, 60, 70, 80, 90\}$ . Plotting the given values against several forms of the incidence angle modifier, a reasonable range for  $b_0$  can be identified for the initialization and estimate evaluation of the grey box model.

Figure 10.3 shows the IAM for several possible values for  $b_0$ , based on  $c_{\min} = 0.20$ . The highest line represents  $b_0 = 0$ , the next  $b_0 = 0.5$ , and so forth for all  $b_0 \in \{0, 0.05, 0.1, 0.15, 0.2, 0.25\}$ . The panel specifications correspond to  $b_0$  between 0.15 and 0.2, with an outlier close to 0.25. Bacher et al. [14] found values around 0.2.

## 10.6 Models with multiple compartments

This section will introduce models with multiple compartments in the flow direction. It is assumed that these compartments are *equal*, that is, they have the same properties and therefore the same performance parameters. The resulting system will have  $N$  system states (or  $2N$  if combined with a metal compartment model, see next section): the outlet temperature from each compartment  $T_{o,i}$  for  $i \in \{1, \dots, N\}$ .

To illustrate how a model can be extended this way, the system equations for the  $N$ -compartment

model *ToIAM\_NComp* are written out here (for  $j = 1 \dots N$ ).

$$dT_{o,j} = \frac{1}{(mC)_e} \left( F'U_0(T_a - T_{f,j}) + c_f Q_f(T_{o,j-1} - T_{o,j}) + F'(\tau\alpha)_{en} (K_{\tau\alpha,b}(\theta)G_b + K_{\tau\alpha,d}G_d) \right) dt + \sigma_{o,j} d\omega_{o,1} \quad (10.13)$$

where  $T_{o,0} \equiv T_i$ . The observation equation involves only the final outlet temperature  $T_{o,N}$ :

$$yT_{o,t_k} = T_{o,N,t_k} + \epsilon_{\text{obs},t_k} \quad (10.14)$$

where  $\mathbb{V}[\epsilon_{\text{obs}}] = \sigma_{\text{obs}}$ .

The introduction of more compartments in the flow direction has two expected effects. First, it increases the freedom in possible temperature profiles in the collector row: from linear to a continuous combination of  $N$  linear functions (still assuming linear temperature profiles in each compartment, for the heat loss to the ambient). Therefore, the heat loss to the ambient can be modelled more closely.

Second, and most importantly, the effect of the inlet temperature  $T_i$  on the outlet temperature at the end of the row is expected to be less instantaneous, as the in-flowing water is mixed into the first compartment before it can affect the temperature in further compartments. This is the reason that the multi-compartment models have been introduced: previous models showed overly strong reactions to changes in inlet temperature.

## 10.7 Models with a metal component

The collector field can in addition be split in the direction perpendicular to the flow direction. The water in the collector and the panel itself can be modelled separately, adding the panel temperature  $T_m$  as a hidden state. For a model with  $N$  compartments in the flow direction, this extension adds a metal component (and thus a hidden state  $T_{m,j}$ ) for each of them. In this case it is assumed that no heat exchange takes place between the metal compartments.

Adding a metal compartment aims to improve the modelled response of the outlet temperature to variations solar radiation. The metal would act as a buffer, so that changes in solar radiation affect the water outlet temperature less instantly. The extension is thus mainly relevant for partly cloudy conditions.

For simplicity, the one-compartment metal model equations will be stated here, but they can be generalized to the  $N$  compartment form. The system equations for  $T_m$  is

$$dT_m = \frac{1}{C_m} \left( U_{fm}(T_f - T_m) + U_{ma}(T_a - T_m) + F'(\tau\alpha)_{en} (K_b(\theta)G_b + K_d G_d) \right) dt + \sigma_m d\omega_m \quad (10.15)$$

and one for  $T_o$ :

$$dT_o = \frac{1}{(mC)_e} \left( U_{fm}(T_m - T_f) + c_f Q_f(T_i - T_o) \right) dt + \sigma_o d\omega_o \quad (10.16)$$

where we still approximate  $T_f = (T_o + T_i)/2$ .

The metal compartment has heat capacity  $C_m$  that differs from the fluid's heat capacity  $(mC)_e$ . Further new parameters in the model are  $U_{fm}$  and  $U_{ma}$ , the respective heat loss coefficients for fluid to metal and metal to ambient. These have replaced the parameter  $F'U_0$ .

## Chapter 11

# Parameter estimates according to EU performance test

This chapter presents the process of applying EU ISO standard [29] performance tests to the studied solar field. Applying this method resulted in a model equation with parameters fit to the data, which will be used to judge whether parameter estimates from the developed grey box models are reasonable. Most importantly, the performance of the models developed in this thesis will be compared to that of the model from the ISO standards.

As presented in Section 3.1.1, the ISO quasi-dynamic method has heat balance equation

$$\begin{aligned} c_f Q_f (T_o - T_i) = & F'(\tau\alpha)_{en} [K_{\tau\alpha,b}(\theta)G_b + K_{\tau\alpha,d}G_d] - F'U_0(T_f - T_a) \\ & - F'U_1(T_f - T_a)^2 - F'U_u u'(T_f - T_a) + F'\epsilon(E_L - \sigma T_a^4) \\ & - (mC)_e \frac{dT_f}{dt} - a_6 u'G - a_7 u'(E_L - \sigma T_a^4) - a_8 (T_f - T_a)^4 \end{aligned} \quad (11.1)$$

where  $c_f$  is the *known* fluid heat capacity,  $Q_f$  is flow per unit aperture area of the collectors,  $T_o$ ,  $T_f$  and  $T_i$  are the outlet, average, and inlet temperature,  $T_a$  is the ambient temperature,  $E_L$  is long wave irradiance, and  $u$  is the wind velocity. All other variables are parameters to be estimated as described in the following.

For flat plate collectors, the ISO standard requires the inclusion of parameters  $F'(\tau\alpha)_{en}$ ,  $K_{\tau\alpha,b}(\theta)$ ,  $b_0$ ,  $K_d$ ,  $(mC)_e$ , and the heat loss coefficients  $F'U_0$  and  $F'U_1$ . This means that these parameters *must* be included in the estimation procedure, while the others may be set to zero.

We are restricted by available measurements, and therefore obtain the simplified expression

$$c_f Q_f (T_o - T_i) = F'(\tau\alpha)_{en} G - F'U_0(T_f - T_a) - F'U_1(T_f - T_a)^2 - (mC)_e \frac{dT_f}{dt} \quad (11.2)$$

If we use split global radiation, we replace  $G = [K_{\tau\alpha,b}(\theta)G_b + K_{\tau\alpha,d}G_d]$  and thus obtain estimates of those parameters as well.

The model selection according to ISO standards is based on multiple linear regression (MLR) modelling. The procedure is carried out iteratively. First, all required parameters are included in the model, and estimated from the data by least square fitting. If this results in parameters with negative value (which is physically unreasonable) or insignificant parameters (based on t-test as described in

Section 7.6), the corresponding parameter is set to 0 and the procedure is repeated with the reduced model equation. This is repeated until all fitted parameters are positive and significant. The method is an example of a *backward* model selection procedure. The parameter  $(mC)_e$  may however not be set to zero: if it is negative or insignificant it should be estimated separately from the weighted sum of the panel components' heat capacities.

The data for estimation should cover 4 – 5 days, containing enough variation in weather conditions and inlet temperature. It is recommended to use  $\geq 5$  minutes average values. Here, 5 minute average have been used.

In the following sections, we will follow the ISO procedure to fit two models. The first model will include only total radiation, whereas the second will work with the split radiation ( $G_{\text{direct}}$  and  $G_{\text{diffuse}}$ ). The final section will discuss and compare their performance.

## 11.1 Results without split radiation: model *MLR1*

The least square fit of parameters in Equation (11.2) based on only total radiation first results in the parameter estimates as given in Table 11.1. The negative value of  $F'U1$  implies that the procedure should be repeated with this parameter set to zero.

	Estimate	Std. Error	t value	Pr(>  t )
Fta	0.76	0.01	82.41	0.00
FU0	8.56	0.24	-35.16	0.00
FU1	-0.10	0.01	17.85	0.00
mC	4181.80	331.73	-12.61	0.00

Table 11.1: Parameter estimates first iteration, for total radiation MLR model.

The least square estimates for the reduced model are given in Table 11.2. All parameters now have the right sign and are significant, which marks the end of the MLR model selection procedure.

	Estimate	Std. Error	t value	Pr(>  t )
Fta	0.86	0.01	95.77	0.00
FU0	4.60	0.12	-38.62	0.00
mC	3885.14	392.95	-9.89	0.00

Table 11.2: Parameter estimates second iteration of total radiation MLR model (with  $FU1$  set to zero)

## 11.2 Results with split radiation, model *MLR2*

As a first step, the least square parameter estimates are obtained for the full equation. The resulting estimates are given in Table 11.3. The parameters  $b_0$  and  $K_d$  are computed as  $Ftab0/Fta$  and  $FtaKd/Fta$  respectively, and therefore their statistics (p-value, variance, etc.) cannot be given in a simple manner.

The parameter  $FU1$  is negative, and is therefore set to zero in the next iteration.

	Estimate	Std. Error	t value	Pr(>  t )
Fta	0.67	0.01	77.46	0.00
Ftab0	0.29	0.01	31.87	0.00
FtaKd	0.22	0.03	7.93	0.00
FU0	0.73	0.32	-2.28	0.02
FU1	-0.01	0.00	2.61	0.01
mC	2541.12	226.06	-11.24	0.00
b0	0.44			
Kd	0.33			

Table 11.3: Parameter estimates first iteration, for split radiation *MLR2* model.

The next iteration results in an insignificant parameter *FU0* (see Table 11.4), which is therefore set to zero for the next estimation step.

	Estimate	Std. Error	t value	Pr(>  t )
Fta	0.67	0.01	77.37	0.00
Ftab0	0.30	0.01	40.39	0.00
FtaKd	0.20	0.03	7.47	0.00
FU0	0.01	0.17	-0.09	0.93
mC	2428.98	222.77	-10.90	0.00
b0	0.46			
Kd	0.29			

Table 11.4: Parameter estimates second iteration (without *FU1*), for split radiation *MLR2* model.

The final estimation result is obtained in the third and last iteration, which estimates only positive and significant parameters as shown in Table 11.5.

	Estimate	Std. Error	t value	Pr(>  t )
Fta	0.67	0.00	301.16	0.00
Ftab0	0.30	0.01	45.55	0.00
FtaKd	0.19	0.01	28.06	0.00
mC	2422.51	210.05	-11.53	0.00
b0	0.46			
Kd	0.29			

Table 11.5: Parameter estimates final iteration (without *FU0* & *FU1*), for split radiation *MLR2* model

### 11.3 Discussion and comparison of model *MLR1* and *MLR2*

The obtained parameter estimates will first be discussed. Next, we will explore the predictive performance of the two models.

Note that the parameter estimates vary considerably between the models. The introduction of split radiation has a large influence on the parameter estimates. This shows that even the state of the art efficiency parameter estimation method does not necessarily reveal any ‘true’ parameters:

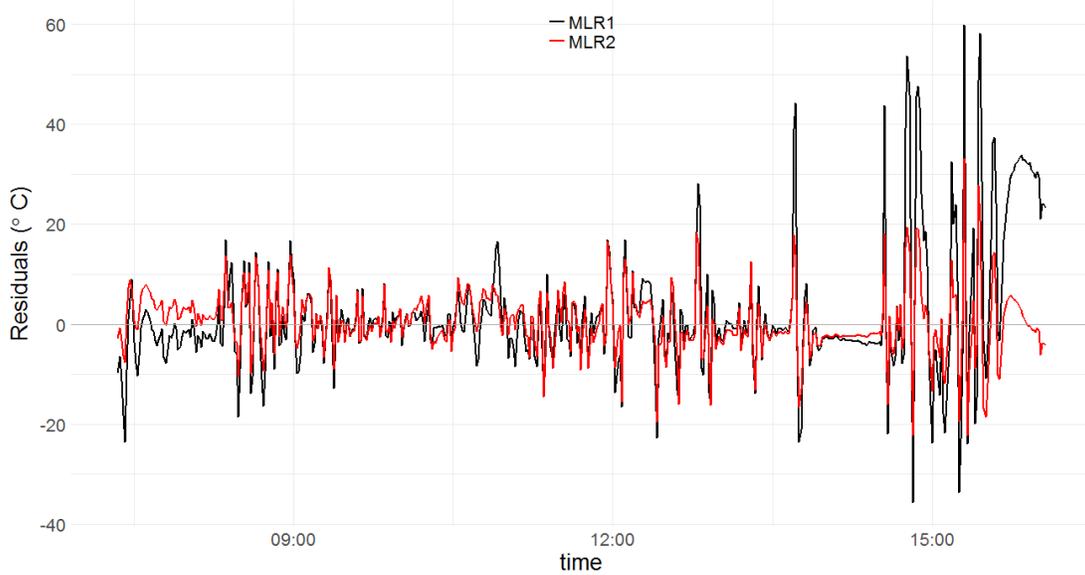


Figure 11.1: MLR residuals on  $D_i$

the parameters depend on the model equation. An especially surprising difference between  $MLR1$  and  $MLR2$  is that the first has a large heat loss coefficient, whereas the heat loss parameters were insignificant and thus set to zero in the second model.

The estimated  $mC_e$  differs considerably between the fits: 3885 for the model  $MLR1$ , and 2423 for  $MLR2$ . Both values are reasonable, although the former resembles estimates as obtained by Bacher et al. more [14]. The efficiency parameter  $Fta$  is 0.86 in  $MLR1$  and 0.67 in  $MLR2$ . Comparing this to the value in the panel specifications ( 0.857 for single glass and 0.814 for double glass) shows that the value for  $MLR1$  matches well, whereas the value for  $MLR2$  is on the low side. In  $MLR1$ , the value of  $FU0 = 4.60$  is very high compared to the panel specifications (3.083 for single glass and 2.102 for double glass). In model  $MLR2$  this parameter was set to zero. It should further be noted that the parameter  $b_0$  is very large compared to usually obtained parameters. This can be due to shading effects: the panels on the plant are shaded by trees and other panels in the afternoon, which may have been captured in the  $b_0$  parameter. Finally, the diffuse radiation IAM constant  $K_d$  (only present in  $MLR2$  is of an extremely low value around 0.3, where it is normally expected to be around 0.9. It is likely that this parameter compensates for the heat loss term being set to zero.

Next, we use these models to predict the outlet temperature given the measured inputs and compute the corresponding residuals. Note that in these linear regression models,  $T_o(t_n)$  is not needed to compute the estimated  $T_o(t_{n+1})$ . Although the MLR models were fit using 5-minute average data, for prediction the original 1-minute average version of the data is used. This will facilitate comparison to the grey box models developed hereafter.

The in this way obtained residuals on cloudy day  $D_i$  and sunny day  $D_s$  are plotted in Figure 11.1 and 11.2, for both model  $MLR1$  and  $MLR2$ . Between 9 AM and 2 PM, the models perform almost equally good, and the residuals are in the range of  $(-20, 20)$  on  $D_i$  and in  $(-5, 5)$  on  $D_s$ . In the afternoon, the model  $MLR2$  with split radiation performs considerably better in terms of residual size. The residuals of fit  $MLR1$  extend from values in the range  $(-100, 50)$ , whereas the second model remains within  $(-50, 50)$ . For both models the residuals are quite correlated, especially on the sunny day.

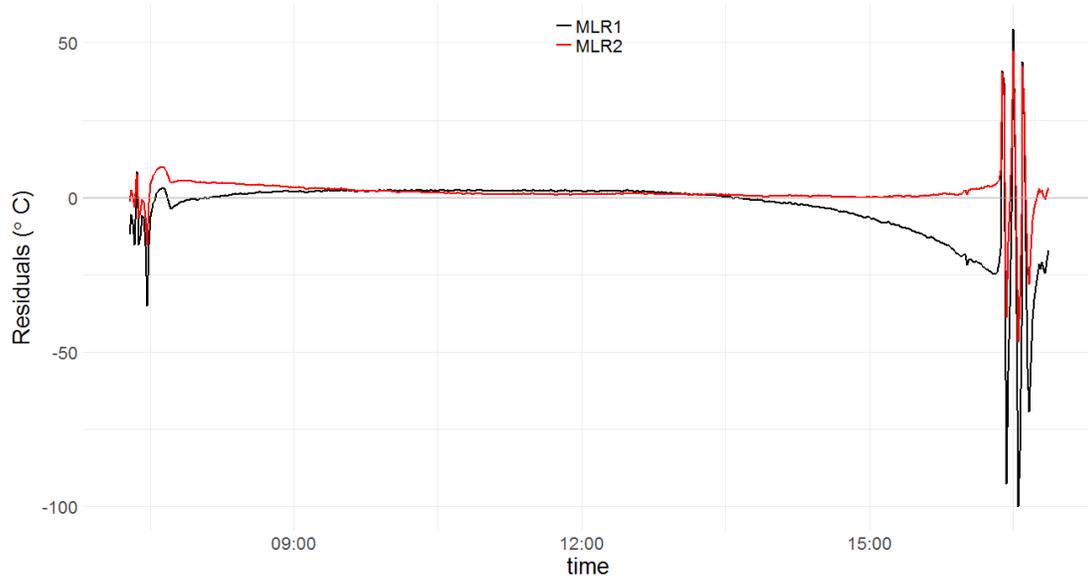


Figure 11.2: MLR residuals on  $D_s$

The developed MLR models will in the following chapter be used for comparison with the grey-box models.

# Chapter 12

## Model selection results

This chapter describes the modelling sequence that was executed as the main research in this thesis. The applied model selection method was described in Chapter 9. The first section of the current chapter described the base model selection. The following sections each discuss one extension step. The last section summarizes the sequence and compares the different extended models.

All models were fit on data set  $D$ , that was introduced in Section 8.3.

### 12.1 Base model selection

To start the model extension procedure, an initial model needs to be selected. This model should match the data reasonably well, so that applied statistical tests like the t-test and likelihood ratio test will be valid. The first trial with model  $Tf$  is presented here, which was slightly modified to model  $To$  in order to satisfy this requirement. The initial trial  $Tf$  is retained in this text, as this attempt revealed a strong disadvantage of certain model formulations worth remembering for future research.

#### 12.1.1 Model $Tf$

Table 12.1 shows the results of the parameter estimation for model  $Tf$  based on full 7-day data set  $D$ , as well as single days  $D_s$  and  $D_i$ . The latter is only done for this first model, to investigate the influence of the weather type on the parameter estimates. The Table shows that the parameter estimates in this first model are indeed affected by the weather type in the data set.

The optimization in the estimation procedure finished well for all three fits. The parameters were significantly different from zero, as computed by the t-test (described in Section 7.6). The uncertainties are generally higher for parameter estimates based on the intermittent day than on the sunny day. The variance of the observation error is much higher for fit  $Tf_i$  than the other two fits. This is expected, as clouds are not included in the model and therefore introduce larger errors into the model.

The values of the parameters are reasonable, but quite far from what they are expected to be based on ISO test values. The parameters from fit  $Tf_i$  are closest to such values. We note that the

	Tf	Tf_s	Tf_i
Tf0	66.3 ( 0.51 )	66.9 ( 0.31 )	66.3 ( 0.62 )
Fta	1.0 ( 0.0059 )	1.0 ( 0.0097 )	0.9 ( 0.0095 )
FU0	6.8 ( 0.02 )	7.0 ( 0.037 )	5.3 ( 0.11 )
mC	21408.3 ( 568 )	29932.7 ( 176 )	10766.5 ( 385 )
sigma_f	-2.6 ( 0.012 )	-3.2 ( 0.062 )	-2.8 ( 0.032 )
sigma_obs	-21.9 ( 0.27 )	-29.6 ( 0.31 )	-12.6 ( 0.15 )

Table 12.1: Parameter estimates model  $Tf$

parameter estimates are likely to be influenced by the model equations, an effect that was also observed in the MLR models of the previous Chapter. When terms are in- or excluded, the physical meaning of parameters may change. In the final section of this chapter, which summarizes the model sequence, the evolution of parameter estimates over the models will be discussed.

Furthermore, we emphasize that the parameters from the panel specifications are determined using the ISO test, and thus correspond to a different model equation than the form used here. As the current equation is derived from the same heat balance, it is desirable that the parameters stay within reasonable limits from the ISO parameters. However, it cannot be said that one of the two methods computes a ‘true’ parameter value.

Model  $Tf_s$  has high values in the correlation table, Table 12.2. There seems to be too little variation in the solar radiation to distinguish the influence of radiation and that of in- and outflow of hot water.

	Tf0	Fta	FU0	mC	sigma_f	sigma_obs
Tf0	1.00	-0.04	0.09	-0.09	-0.08	-0.08
Fta	-0.04	1.00	-0.43	0.46	0.39	0.38
FU0	0.09	-0.43	1.00	<b>-1.00</b>	<b>-0.79</b>	<b>-1.00</b>
mC	-0.09	0.46	<b>-1.00</b>	1.00	<b>0.79</b>	<b>0.99</b>
sigma_f	-0.08	0.39	<b>-0.79</b>	<b>0.79</b>	1.00	<b>0.78</b>
sigma_obs	-0.08	0.38	<b>-1.00</b>	<b>0.99</b>	<b>0.78</b>	1.00

Table 12.2: Correlation table of parameters model  $Tf_s$

## Simulation and prediction

To evaluate the performance of model  $Tf$ , we further investigate simulation and prediction of fit  $Tf$ . The upper panel in Figure 12.1a shows the simulated and measured outlet temperature  $T_o$ ; the lower panel shows the (measured) inlet temperature  $T_i$ . The simulation shows that the model is a reasonable start. The temperature level in the steady state (mid-day) situation matches the true steady state temperature.

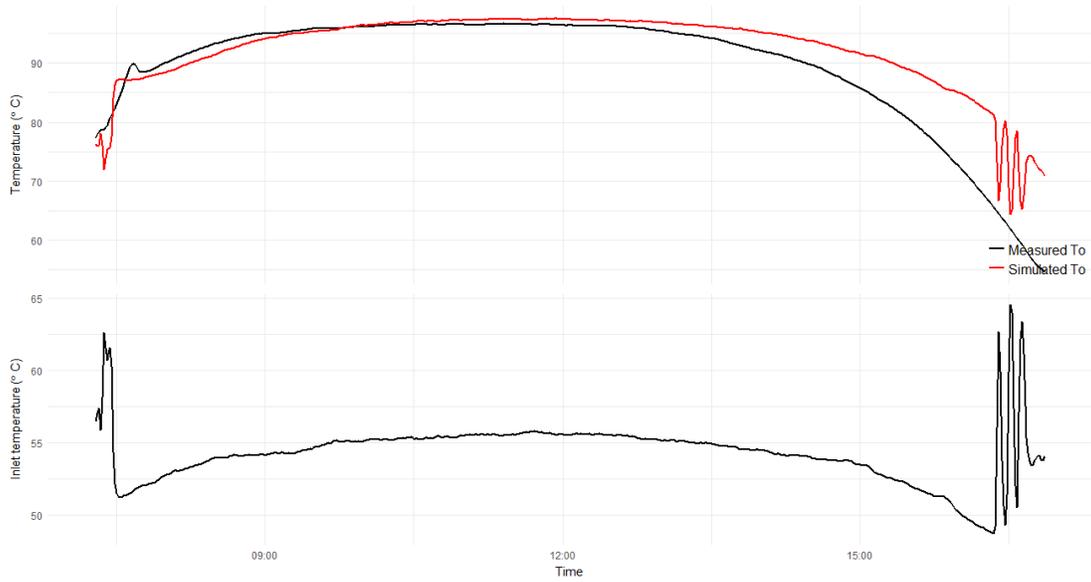
The graphs show that the simulated outlet temperature is very sensitive to changes in the inlet temperature, which shows large fluctuations in the morning and afternoon. This effect was also present in simulations for cloudy days, but less easily distinguishable from other fluctuations in the inlet temperature and therefore omitted here. The measured outlet temperature is not visibly affected by these fluctuations, which is quite remarkable.

The same effect is seen when the model with fitted parameters is used to predict on the same sunny day. The residuals (residual =  $T_{o,\text{measured}} - T_{o,\text{predicted}}$ ) are shown in the upper panel of Figure

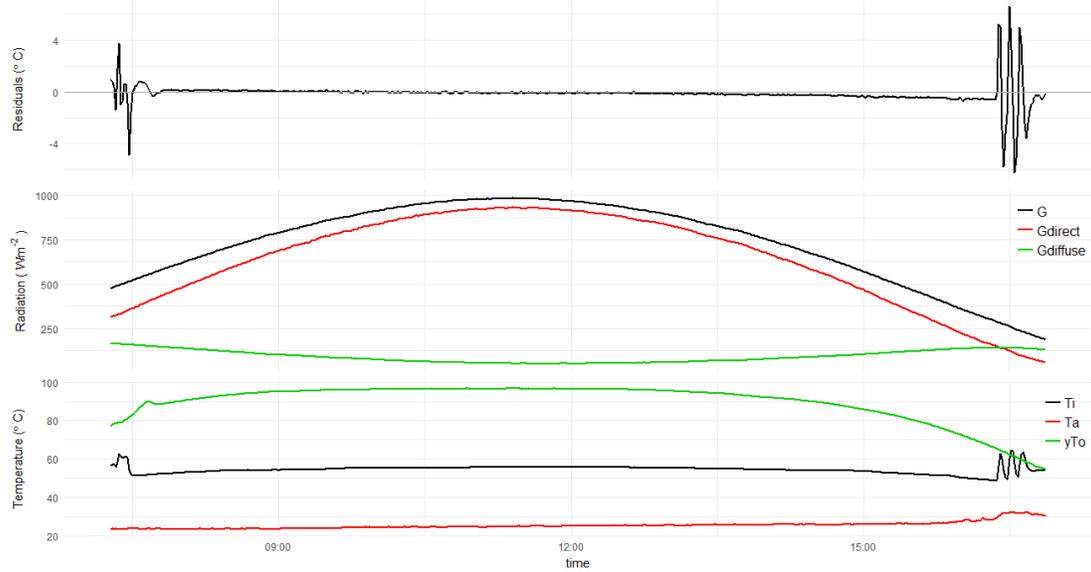
12.1b.

Plots of the individual heat flows (omitted here) showed that none of the terms in the system equation causes these peaks in the predicted and simulated outlet temperature. The measurement equation must therefore be the source of the problem. As the inlet temperature is included in the measurement equation, it directly influences the outlet temperature, without interacting with other heat flows first. This is not physically reasonable and harms model performance, as seen in the residuals and simulation using this model.

It can be concluded that input measurements, here the inlet temperature, cannot directly be in the measurement equation. For this reason, further models should take  $To$  as the system state, so that input variables will not need to be present in the measurement equation.



(a) Simulated and measured  $T_o$ . Lower panel shows measured  $T_i$

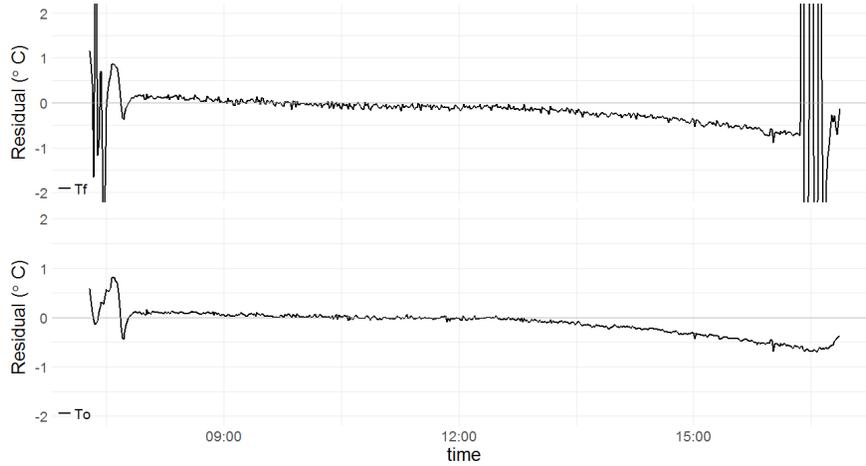


(b) Predicted (1 ahead) and measured  $T_o$ , compared to various inputs.

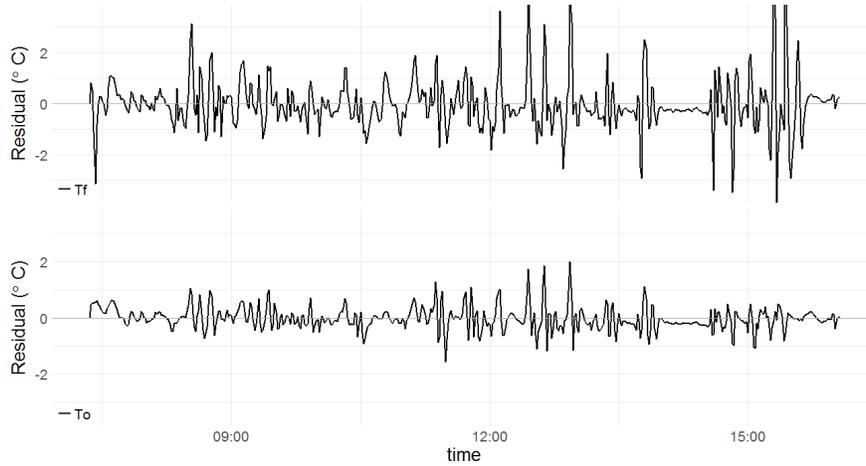
Figure 12.1: Simulation and prediction of model  $Tf$

### 12.1.2 Model $To$

The 1-ahead prediction residuals of model  $To$  and  $Tf$  on sunny day  $D_s$  are shown in Figure 12.2a. It is clear that model  $To$  performs better, mainly because it is less sensitive to changes in  $T_i$ . The residuals on cloudy day  $D_i$  are clearly reduced as well, but this plot is omitted here for brevity. The model  $To$  will therefore be chosen as the start for the model selection procedure.



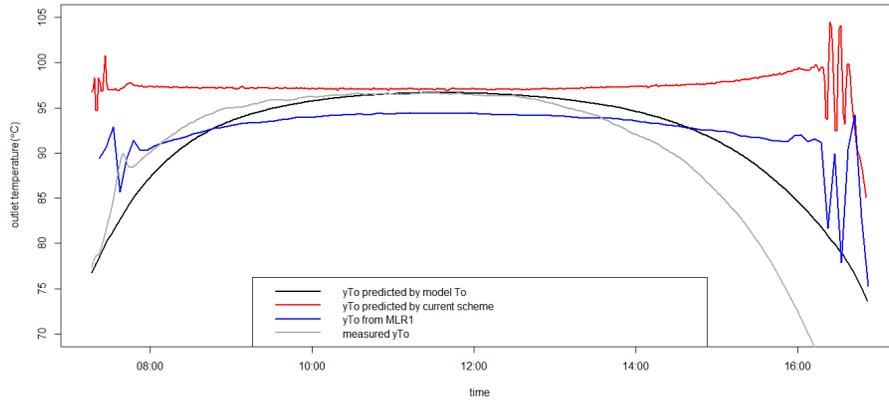
(a) Clear sunny day  $D_s$



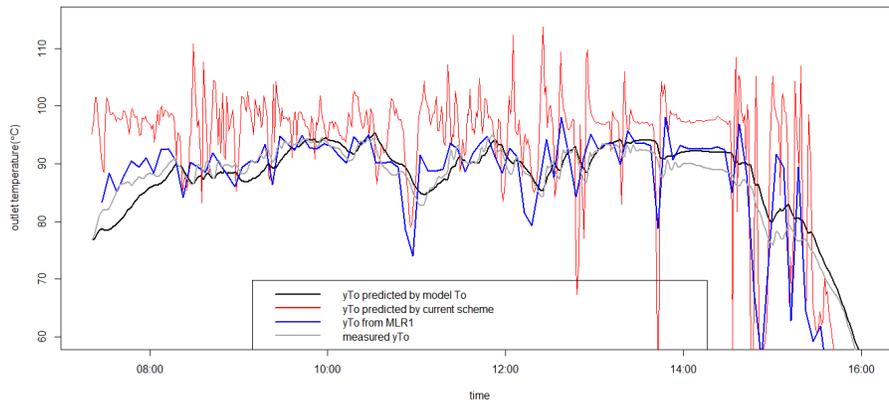
(b) Intermittently cloudy day  $D_i$

Figure 12.2: Residuals (1-ahead) of model  $Tf$  and  $To$

The result of simulations with model  $To$ , versus simulation of the currently used prediction equation and the prediction of fit  $MLR1$  is shown in Figure 12.3a and 12.3b. The current scheme overestimates the output temperature during a large part of the day, as expected from the observed control behaviour of the plant. On the cloudy day, it is visible that the current scheme shows large fluctuations. It seems therefore that even this simple grey-box model captures the plant dynamics better than the currently used equation. This shows that allowing the outlet temperature at  $t_{n+1}$  to depend on the outlet temperature at  $t_n$  improves model performance.



(a) Clear sunny day  $D_s$ .



(b) On intermittent day  $D_i$ .

Figure 12.3: True and simulated outlet temperatures using model  $T_o$ , model  $MLR1$ , and current control equation

### 12.1.3 Motivation for model extensions

Comparing the prediction residual time series of model  $T_o$  versus input variables reveals several opportunities for model extension. On sunny day  $D_s$  (Figure 12.4), the prediction residuals show the following model flaws

1. Structural underestimation of  $T_o$  in the first half of the day, thereafter structural overestimation. Especially in the afternoon the residual is very large. Therefore it is expected that the shading of the panels is causing the pattern. This is also visible on cloudy days, although less distinguished due to more fluctuations in all variables.
2. Some high residual peaks in the morning, seemingly related to fluctuations in the inlet temperature. This is verified by plotting the CCF for the first 35 minutes of  $D_s$ , see Figure 12.5a. Lag  $k$  in this plot corresponds to  $\text{cor}(T_{i,t+k}, \text{resid}_t)$ , so the plot shows that past values of the inlet temperature are positively correlated to the current residual

- The residual ACF on sunny days is very significant for lags up to 2.5 hours, see Figure 12.5b. This is not the case on cloudy days. This means that longer periods of sun are not modelled well.

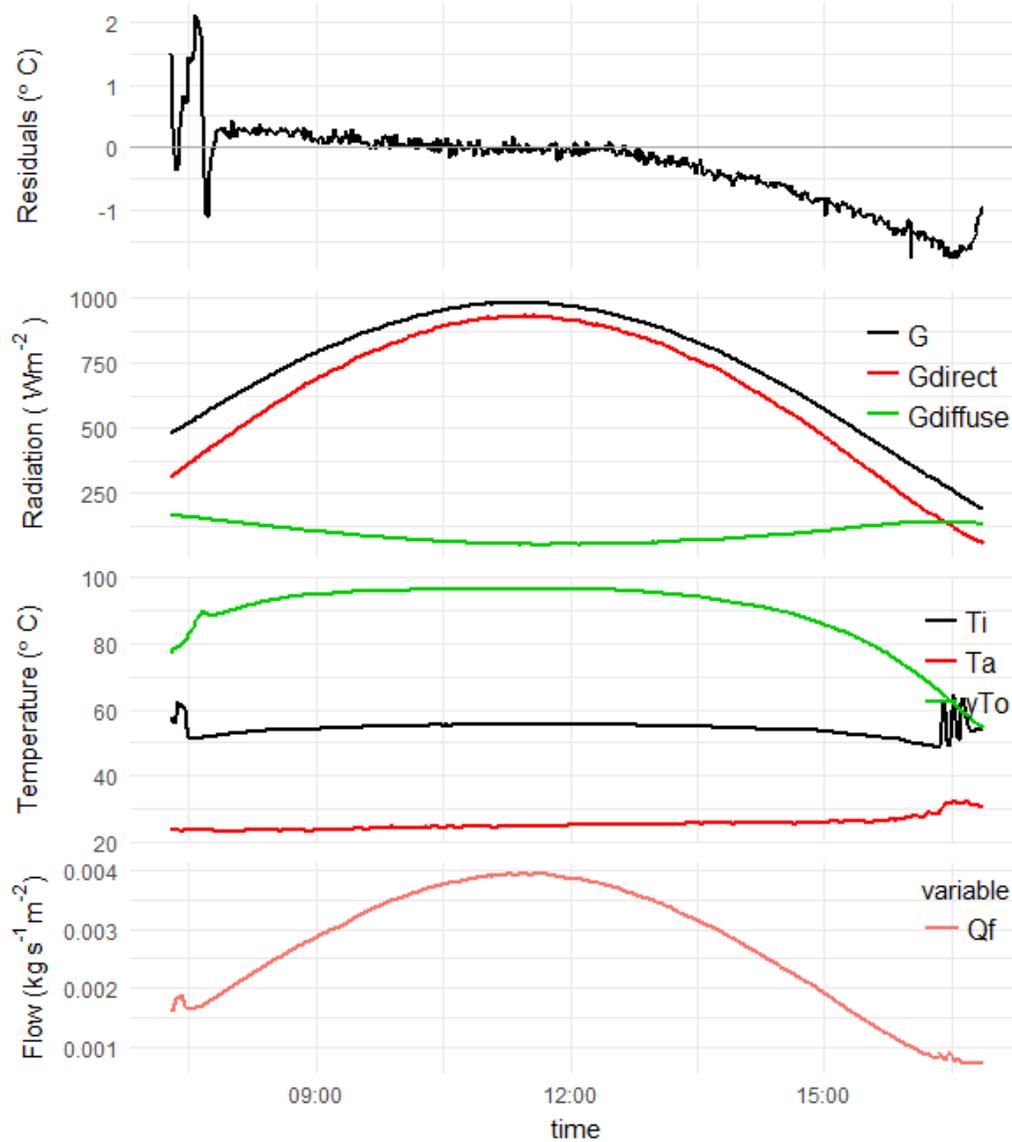
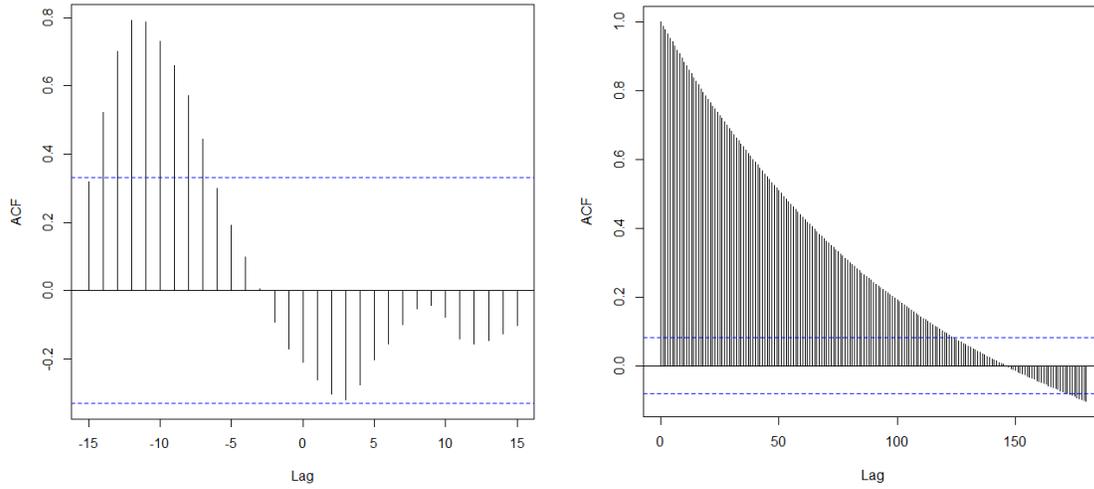


Figure 12.4: Model  $T_o$  1-ahead residual along with input time series on clear sunny day  $D_s$

To improve on these aspects, several extensions are proposed here. For the first, we will use splines for the non-parametric modelling of a curve, which can describe the changing  $F'(\tau\alpha)_{en}$  over time. It will also be tested whether the introduction of an IAM function could improve the modelled influence of solar radiation.

Two extensions will be introduced in an attempt to reduce the model errors due to inlet temperature fluctuations. First of all, introducing more compartments in the flow direction may delay the effect of  $T_i$  on  $T_o$ . Secondly, it is possible to manipulate which measurement of  $T_i$  should affect the outlet temperature at time  $t$ , by replacing the time series of  $T_i$  by a delayed version of it.



(a)  $CCF(T_i, \text{residuals})$  for first 35 minutes of  $D_s$

(b) ACF 1-ahead residuals

Figure 12.5: Correlation plots for  $D_s$ , using model  $T_0$

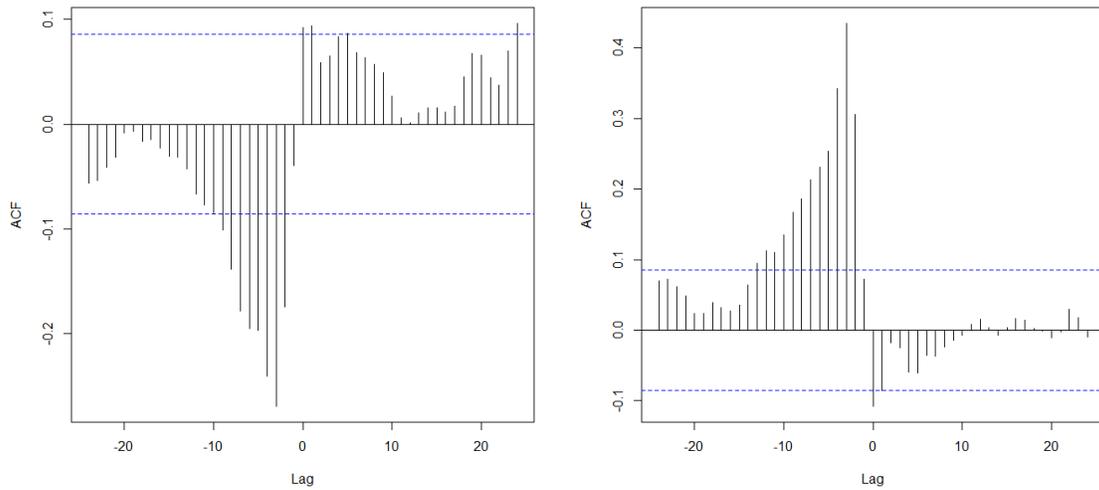
It will be investigated whether more detailed modelling of heat loss can improve the observed high auto-correlation for sunny day residuals. This will be done by including a second order heat loss term (i.e. temperature dependence of the heat loss). Including this term would also allow for a more straightforward comparison between the currently developed model and the ISO standard model. The introduction of a metal compartment could also improve the modelled heat loss.

On cloudy day  $D_i$ , it is more difficult to see in the time series plot which fluxes influence the residual, as all inputs fluctuate simultaneously. We can however infer several aspects that could be improved, using different plotting techniques.

1. A plot of the Cross-Correlation Function (CCF) between total radiation and the residuals shows that the residuals are negatively correlated to past radiation measurements (Figure 12.6a). This means that the increased outlet temperature due to radiation input is overestimated and perhaps too instantaneous.
2. The CCF between estimated diffuse radiation and the residuals (Figure 12.6b) shows a positive correlation between past diffuse radiation and current residual, indicating that a high diffuse component results in small residuals (i.e. overestimated outlet temperature)
3. A time series plot shows overestimation of the outlet temperature in periods of low solar radiation during the day, corresponding to cloudy moments.

To overcome the first, a metal component may help to delay the effect of solar radiation on outlet temperature. The separate modelling of diffuse and direct radiation effects are expected to improve the second and third aspect.

In the following paragraphs, these extensions will be applied.



(a) CCF between total radiation and residuals      (b) CCF between diffuse radiation and residuals

Figure 12.6: Correlation plots for  $D_i$ , using model  $To$

## 12.2 First extension

The extensions for model  $To$  that have been considered in this first step are the following models:

- Model  $ToGsplitted$  uses the split radiation as explained in Section 10.4
- Model  $ToIAM$  includes an IAM function instead of  $Fta$
- Model  $ToTm$  has a metal component, as introduced in Section 10.7
- Model  $To_{2C}$  has two compartments in the flow direction
- Model  $To_{splined}$  uses splines to divide  $G_{total}$  into several components, which each have their own efficiency parameter. This method was introduced in Section 10.3

For all extended models, the parameter estimation procedure ended successfully. The loglikelihoods of the resulting fits are displayed in Table 12.3

We will first decide on the number of splines to be used per day for the  $To_{splined}$  model. This choice will be based on the loglikelihoods, simulations, and predictive performance of the models. Models with 2 up to 6 splines have been fitted. Based on loglikelihood only, the model with 6 splines performs best. It remains to be checked with a likelihood ratio test if the extension from each of the splined models to  $To_{6Splined}$  is significant. The LRT gives p-value  $< 0.01$  for each of these different tests. Therefore the 6 splined model should be preferred based on loglikelihood.

The simulation and prediction of this model is however barely different from the model with 4 splines. The model with 4 and 5 splines do not differ at all, and this extension is not significant based on the LRT. The parameter estimates for  $FU0$ ,  $mC_e$ , and  $To0$  are very similar the 4- and 6- spline model as well (values omitted here). As we would like to pick the best model that is as simple as possible, based on these observations the model with 4 splines is chosen here for further analysis.

	nr.of.par.	Loglikelihood	LRT p-value
To	6	-2125.08	
ToGsplit	7	-2076.56	0.00
ToTm	10	-2215.09	1.00
To_IAM	7	-1928.25	0.00
To2C	8	-533.36	0.00
To_2Splined	7	-1873.28	0.00
To_3Splined	8	-1677.78	0.00
To_4Splined	9	-1589.95	0.00
To_5Splined	10	-1589.91	0.00
To_6Splined	11	-1583.77	0.00

Table 12.3: Overview of models in first extension step: number of estimated parameters, loglikelihood, p-value of LRT compared to base model  $To$ .

Now we will consider this splined model and the other extended models combined. All extensions except the introduction of a metal compartment in  $ToTm$  are very significant. The model  $To2C$  with more compartments has the lowest loglikelihood. It does well in 1-step ahead prediction, but performs badly in simulation as shown in Figure 12.7. The model has a low loglikelihood because the extra (hidden) state absorbs unreasonably much of the (yet un-modelled) fluctuations, indicating an overfitted model. This nonphysical behaviour becomes clear when plotting the filtered states, as shown in Figure 12.8. It may be too early in the model sequence to introduce these compartments with hidden states.

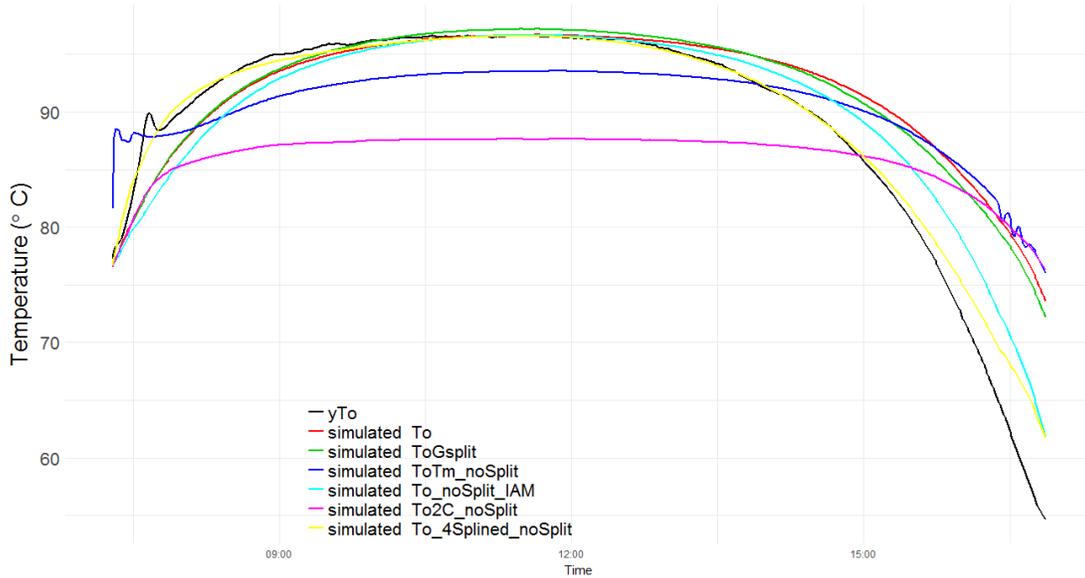


Figure 12.7: Simulated outlet temperature on  $D_s$  for models in first extension step.

For this reason, we will select the model with the next largest loglikelihood, which is one of the splined models. By the previous discussion, we take model  $To_4splined$ . For comparison, we plot the simulated outlet temperature of the extensions in Figure 12.7. The yellow line, for the chosen model  $To_4splined$ , is closest to the true temperature. This is mainly clear in the afternoon, where shading of the panels shading occurs. The same holds for prediction, for which the plot is omitted for brevity.

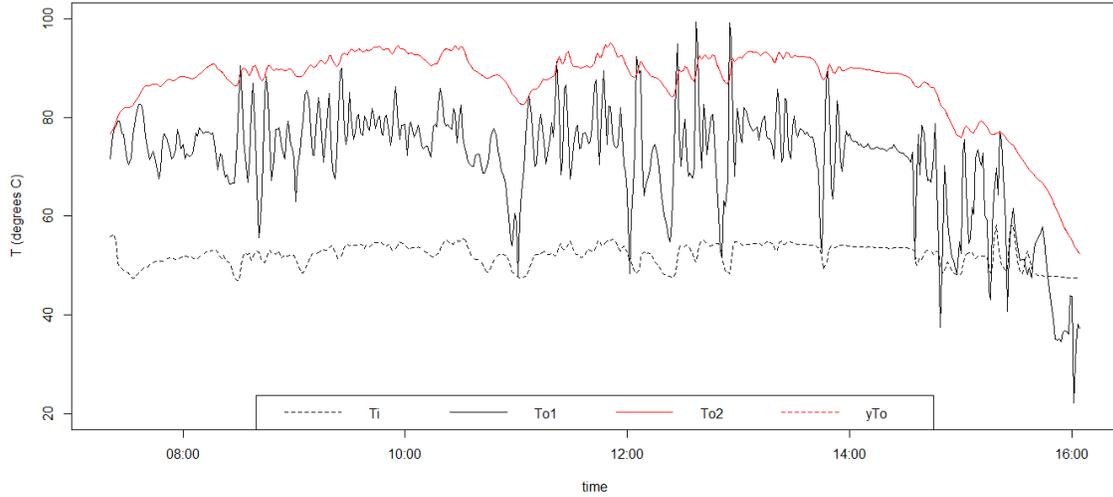


Figure 12.8: Filtered states for model  $To2C$  on  $D_i$ .

The 1- and 10-ahead residuals of the chosen model are compared an overview plot in Section 12.7, Figures 12.22, 12.23, 12.24, and 12.25. These Figures show the residuals of all chosen models in the model sequence are plotted. The negative residuals in the afternoon on  $D_s$  are reduced by the splines, but it can be seen that the splines do impose a new pattern in the residuals. However, the residual sizes are largely reduced and the prediction error remains within a range of only  $0.5^\circ C$  for most of the day. In addition, the sunny afternoon's residuals are reduced considerably. The cloudy day prediction is improved as well, as the residual sizes are reduced over the whole day.

The fit  $To_4Splined$  has high correlations between the efficiency parameters of neighbouring splined radiation components. The other parameters do not have high correlations. It is sensible that the spline coefficients are correlated, as the different components of the solar radiation overlap, and therefore an increase of one coefficient can be combined with a decrease in the coefficient of the spline next to it. This may in turn lead to an increase in the coefficient two steps away (for example an increase in  $s_1$  would imply an increase in  $s_4$ ). We would therefore expect negative correlations between  $s_i$  and  $s_{i\pm 1}$  and positive between  $s_i$  and  $s_{i\pm 2}$ . This is indeed the case for the estimates in  $To_4splined$ .

	To0	FU0	mC	s1	s2	s3	s4	sigma_o	sigma_obs
To0	1.00	0.02	0.04	-0.07	0.07	-0.03	0.04	-0.00	0.08
FU0	0.02	1.00	0.42	0.61	-0.27	0.58	0.05	-0.06	-0.31
mC	0.04	0.42	1.00	0.44	-0.40	0.60	-0.51	-0.06	-0.40
s1	-0.07	0.61	0.44	1.00	<b>-0.88</b>	<b>0.89</b>	-0.52	-0.04	<b>-0.70</b>
s2	0.07	-0.27	-0.40	<b>-0.88</b>	1.00	<b>-0.89</b>	<b>0.75</b>	0.02	0.70
s3	-0.03	0.58	0.60	<b>0.89</b>	<b>-0.89</b>	1.00	<b>-0.72</b>	-0.06	-0.61
s4	0.04	0.05	-0.51	-0.52	<b>0.75</b>	<b>-0.72</b>	1.00	0.05	0.37
sigma_o	-0.00	-0.06	-0.06	-0.04	0.02	-0.06	0.05	1.00	0.01
sigma_obs	0.08	-0.31	-0.40	<b>-0.70</b>	0.70	-0.61	0.37	0.01	1.00

Table 12.4: Correlation table for the parameters in model  $To_4splined$

In this splined model, the parameters  $s_1$  up to  $s_4$  determine the estimated effective radiation during

the day. When dividing this effective radiation by the measured radiation, the panel efficiency  $Fta$  can be computed as a function of time. This is plotted in Figure 12.9b. It shows that until the afternoon, the panel efficiency parameter  $F'(\tau\alpha)_{en}$  is quite constant around a value of 0.9. The parameter is in standard models intended as a constant. Here, it is shown that unmodelled effects in the afternoon result in different efficiency: the panels are increasingly shaded by one another and by the trees West of the field. This can be seen in the fitted efficiency, which drops drastically in the afternoon.

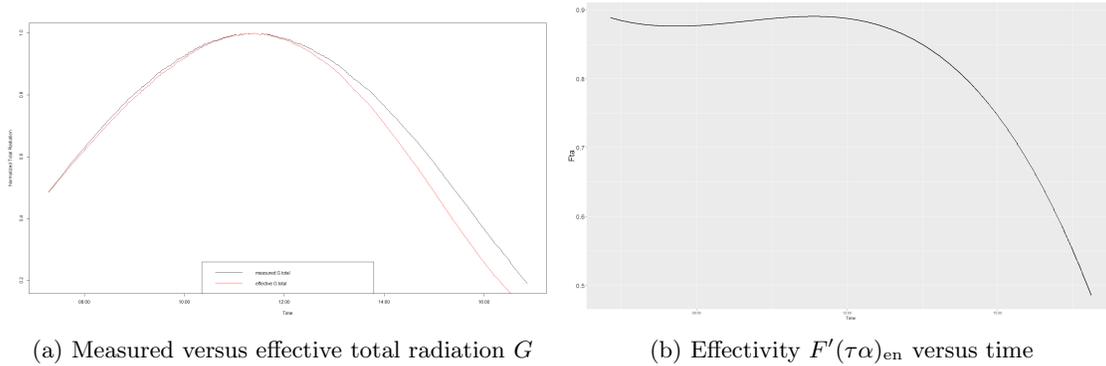


Figure 12.9: Visualizing effective total radiation for model *To\_4Splined* (first extension)

The solar radiation is the factor of highest influence on the outlet temperature. Therefore it is reasonable that the introduction of splines, accounting for shading effects and thereby improving the solar part of the model, is the first most significant extension.

## 12.3 Second extension

The extensions for model *To\_4Splined* that have been considered in this second step are the following models:

- Model *ToGsplit\_splined*, which adds the split radiation
- Model *ToIAM\_splined*, which includes an IAM for  $Fta$
- Model *ToTm\_splined* that has a metal component
- Model *To\_2C\_splined*, which has two compartments in the flow direction

Of each model type, we will again choose the number of splines that fits the data best, based first on loglikelihood and thereafter checked using simulation and prediction performance. Next, we will compare the different model types and select the best model extension.

	Nr. of par.	Loglikelihood	LRT p-value
To_4Splined	9	-1589.95	
ToIAM_4Splined	10	-1589.93	0.85
ToIAM_5Splined	11	-1670.71	1.00
ToIAM_6Splined	12	-1583.89	0.01
ToTm_4Splined	13	-1334.10	0.00
ToTm_5Splined	14	-1993.66	1.00
ToTm_6Splined	15	-1328.39	0.00
To2C_4Splined	11	-556.08	0.00
To2C_5Splined	12	-442.99	0.00
To2C_6Splined	13	-678.31	0.00
ToGsplit_4Spline	10	-1473.82	0.00
ToGsplit_5Spline	11	-1472.98	0.00
ToGsplit_6Spline	12	-1453.25	0.00

Table 12.5: Overview of models in second extension step: number of estimated parameters, loglikelihood, p-value of LRT compared to base model *To\_4Splined*.

## Number of splines for each model type

Again, the loglikelihood of the multiple compartment model in the flow direction is the highest, but for the same reasons as previously this model is not considered suitable. These models in addition have large values of  $dF/dPar$  after the estimation procedure, which indicates that the true optimum has not been found. These models are therefore not suitable, and none will be chosen as the next extended model.

The IAM models with 4 and 6 splines have an estimated value of  $b_0$  smaller than 0.001. This corresponds to an IAM of value close to 1 for all angles, so the function does not alter the outcome of the model much. For this reason, these models are not preferred over the previous model. The 5-splined version of the IAM model has a more reasonable value of  $b_0 = 0.12$  (still a little low), but the estimation procedure ends with high values of  $dF/dPar$ , which means that the true optimum may not have been found. This is also reflected in the large loglikelihood of this model. It is concluded that the IAM models are not suitable in this extension step.

For the models with split radiation, the likelihood ratio test favours 6 splines over 5 splines and 4 splines with p-values close to zero. Therefore, *ToGsplit\_6splined* is selected.

The metal models with 4 and 6 splines have a low loglikelihood. They perform well in simulation and prediction. The extension from 4 to 6 splines has p-value = 0.003, therefore 6 splines model *ToTm\_6Splined* is selected.

It is noted that two models have a lower loglikelihood than the previous model, while it is known that generalizations of a model should always have a loglikelihood higher or equal than the original. However, the models in question have 5 splines whereas the previous model had 6. Therefore, they are in the strictest sense not a generalization of the previous model, and thus it can happen their loglikelihood is lower. In addition, the estimation procedure of model *ToIAM\_5spline* did not find the true optimum (judged by a large value of  $dF/dPar$ ).

## Selection of extended model

It remains to choose between the metal model and the split radiation model. Model *ToTm\_4Splined* has the highest loglikelihood, closely followed by *ToGSplit\_6splined*. The performance of these models on plots is very similar, although the metal model is slightly better on sunny days and the split radiation is better on cloudy days. It is concluded that the models are equally suitable in this step. Therefore, the simpler model *ToGsplit\_6Spline* is selected in this step. Figure 12.22 to 12.25 show the improvement in residuals of the selected model compared to the previous model, which are mainly clear in the afternoon of  $D_i$ . As discussed in Section 12.1.3, we intended this extension to improve the CCF between estimated diffuse radiation and the residuals. Comparing the new CCF in Figure 12.10 to the initial one in Figure 12.6b, we indeed see reduced correlation. There are however still several significant correlation bars.

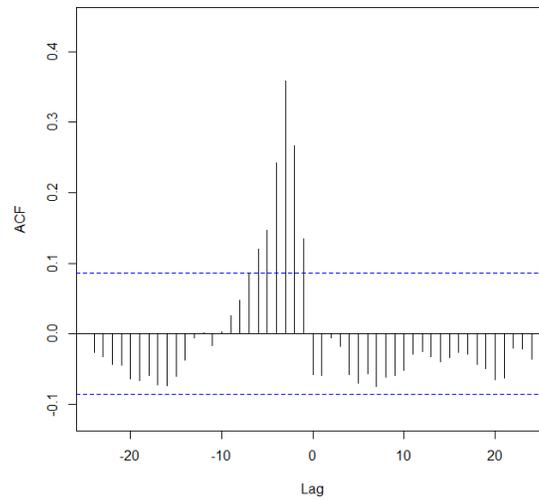


Figure 12.10: CCF between diffuse radiation and residuals on  $D_i$ , using model *ToGSplit\_6splined*

It is thus found that the radiation split into direct and diffuse components is more significant than the introduction of an IAM on total radiation  $G$ . This is an interesting result, because many solar heat plants do not measure the two components separately. This result indicates it could be useful to do so, although the ‘theoretical’ splitting as done now is perhaps good enough.

Figure 12.11a and 12.11b respectively show the effective radiation and the change in efficiency over the day. This effective  $F'(\tau\alpha)_{en}$  is again computed as  $\sum_i s_i G_{dir,i} / G_{b,measured}$ . The model does not include an IAM as of yet. Therefore, the incidence angle effects are also modelled by this splined parameter, which should therefore be interpreted as the proportion of measured radiation that is taken up by the fluid (as a function of time).

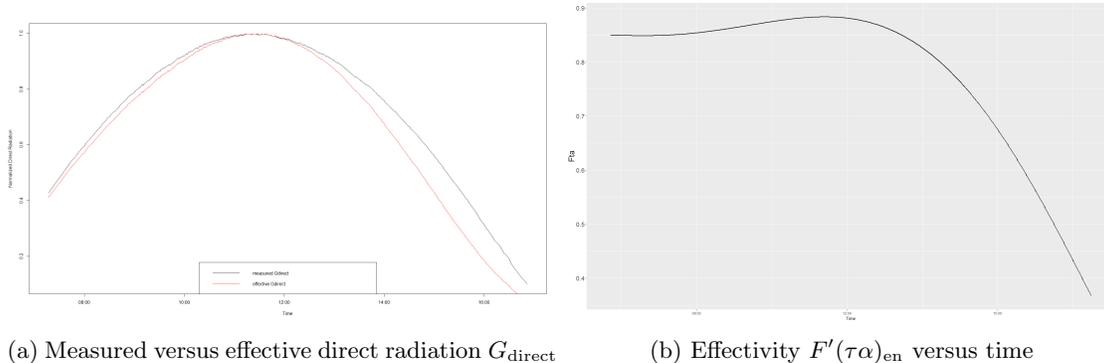


Figure 12.11: Visualizing effective direct radiation for model *To\_Gsplit\_6Spline* (second extension)

## 12.4 Third extension

The extensions for model *ToGsplit\_6splined* that have been considered in this third step are the following models:

- Model *ToIAMGsplit\_splined*, which introduces an IAM function
- Model *ToTmGsplit\_splined*, with a metal component
- Model *To2CGsplit\_splined*, with 2 compartments in the flow direction
- Model *ToGsplit\_splined\_FU1*, which includes a second order heat loss term

An overview of all obtained loglikelihoods is given in Table 12.6.

	Nr. of par.	Loglikelihood	LRT p-value
ToGsplit_6Spline	12	-1453.25	
ToIAM_Gsplit_4Splined	11	-1426.22	
ToIAM_Gsplit_5Splined	12	-1424.95	
ToIAM_Gsplit_6Splined	13	-1412.30	0.00
ToTmGsplit_4Splined	14	-1188.73	0.00
ToTmGsplit_5Splined	15	-2275.32	1.00
ToTmGsplit_6Splined	16	-2118.36	1.00
To2CGsplit_4Splined	12	-628.97	0.00
To2CGsplit_5Splined	13	-613.24	0.00
To2CGsplit_6Splined	14	-340.66	0.00
ToGsplit_4Splined_FU1	11	-1473.67	
ToGsplit_5Splined_FU1	12	-1472.64	1.00
ToGsplit_6Splined_FU1	13	-1452.96	0.45

Table 12.6: Overview of models in third extension step: number of estimated parameters, loglikelihood, p-value of LRT compared to base model *ToGsplit\_6Spline*.

## Number of splines for each model type

For the *ToIAMGsplit\_splined* models, the LRT of the extension from 4 to 6 splines favours the 6 spline model with a p-value  $\approx 0$ . Therefore, model *ToIAM\_Gsplit\_6Splined* is selected for further analysis.

The metal models were difficult to fit; only for 4 splines the estimation procedure finished with good results, which is therefore selected for further analysis. This version also has the highest loglikelihood of this model type.

In the multi-compartment model, *To2CGsplit\_6Splined* is selected.

For the second order heat loss model *ToGsplit\_Splined\_FU1*, the extension from 4 or 5 to 6 splines is very significant with a p-value of  $\approx 0$ . Therefore the best of these models is *ToGsplit\_6Splined\_FU1*.

## Selection of extended model

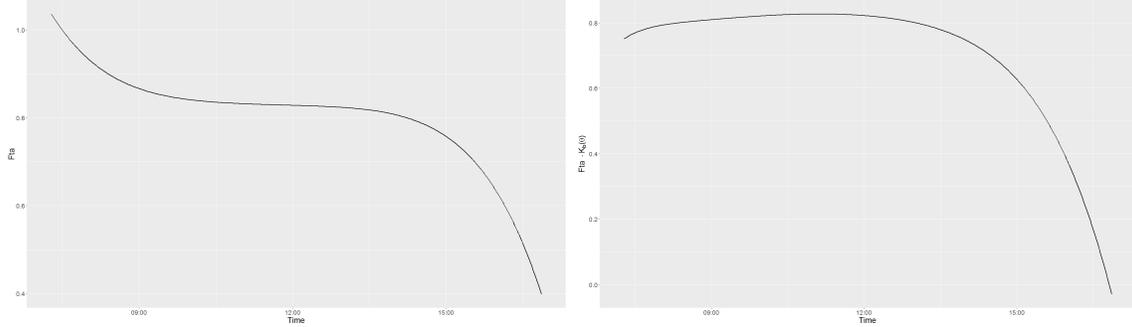
The lowest loglikelihood is once again obtained by the multi-compartment model. For the same reason as before, this model is not suitable. The extension to the second order heat loss model is not significant. The metal model and the IAM model remain as suitable extensions.

The LRT p-value is smaller than floating point precision for each of the extensions, so that the choice can not be based on this test. Model *ToTmGsplit\_4Splined* has the highest likelihood, but includes an extra parameter compared to *ToIAM\_Gsplit\_6Splined*. The models are very close in performance. The metal model is slightly better on cloudy days, which is to be expected because the clouds cause fluctuations in radiation. The effect of these fluctuations are buffered by the metal, which is modelled better with the metal component. The IAM model does better on sunny days, when the incidence angle effects matter most.

However, the metal model starts to perform worse even on cloudy days when the prediction is 10 minutes ahead or more. Another disadvantage of this model is the longer estimation time, and the additional parameter. Therefore, the IAM model *ToIAM\_Gsplit\_6Splined* is selected in this iteration.

The main visible effect of the added IAM is reduced afternoon residuals on  $D_s$ , see Figure 12.22 and 12.24. This was previously also achieved by the splines, however, in the current case the residuals on the cloudy afternoon remain unchanged (see Figure 12.25). This shows that the model's representation of the direct beam solar component has improved.

The plots in Figure 12.12 show the effectivity parameters of this selected model as a function of time. The upper panel shows the single parameter  $F'(\tau\alpha)_{\text{en}}$ , which is computed from the parameters  $s_1$  to  $s_6$  as before. This parameter no longer includes incidence angle effects, as these have now been explicitly included with an IAM, and therefore only includes shading effects. From this plot, it should be concluded that there is some shading in mid day already, but this is not the case at the time of this data set (May). This indicates that the effectivity parameter can no longer be interpreted alone, but should instead be analyzed combined with the IAM  $K_d$ , as these two are fitted simultaneously. This plot is shown in the lower panel. The splines and IAM function together allow the effective radiation to take on a shape that is quite different from the previous two selected models (as previously shown in Figure 12.9 and 12.11b).



(a) Effectivity  $F'(\tau\alpha)_{\text{en}}$  versus time      (b) Effectivity and IAM combined  $F'(\tau\alpha)_{\text{en}} \cdot K_b(\theta)$

Figure 12.12: Visualizing effective direct radiation for model *ToIAM\_Gsplit\_6Splined* (third extension)

## 12.5 Fourth extension

The extensions for mode *ToIAM\_Gsplit\_6Splined* that have been considered in this fourth step are the following:

- Model *ToIAM\_Gsplit\_Splined\_FU1*
- Model *ToTmIAM\_Gsplit\_Splined*, which contains a metal component
- Model *ToIAM\_Gsplit\_2C\_splined*, with an additional compartment in the flow direction

### Number of splines for each model type

The loglikelihoods of the fits are given in Table 12.7.

The metal models were difficult to fit: only the estimation procedure of the 4-spline version *ToTmIAM\_Gsplit\_4splined* finished successfully. The 5- and 6-spline versions had high values of  $dF/dPar$ , which indicates that the optimum was not found.

The same holds for the multi-compartment models with 5 and 6 splines. Although the likelihood of the 5-splined model *To2C\_IAMGsplit\_5splined* is considerably higher than the model that was to be extended, the optimum found for this model is not a good fit based on simulation and prediction plots. The 4 spline version of this model is therefore selected for further analysis.

Of the FU1 models, the 6 spline version is the only one of higher loglikelihood than the model that was to be extended.

	nr. of par.	Loglikelihood	LRT p-value
ToIAM_Gsplit_6Splined	13	-1412.30	
ToIAM_Gsplit_4splined_FU1	12	-1624.75	
ToIAM_Gsplit_5splined_FU1	13	-1415.14	1.00
ToIAM_Gsplit_6splined_FU1	14	-1404.41	0.00
ToTmIAM_Gsplit_4splined	15	-1118.03	0.00
ToTmIAM_Gsplit_5splined	16	-1767.74	1.00
ToTmIAM_Gsplit_6splined	17	-1756.01	1.00
To2C_IAMGsplit_4splined	13	-363.45	0.00
To2C_IAMGsplit_5splined	14	-642.40	0.00
To2C_IAMGsplit_6splined	15	-1129.97	0.00

Table 12.7: Overview of models in fourth extension step: number of estimated parameters, loglikelihood, p-value of LRT compared to base model *ToIAM\_Gsplit\_6Splined*.

## Selection of extended model

As in all previous steps, the two-compartment model has the highest loglikelihood, and is therefore examined for suitability first. The one-step ahead prediction of this model has smaller residuals than all other models. However, in further-ahead prediction it soon performs worse than the other models, except at times where  $T_i$  shows fluctuations. This indicates that the model is over-fitted to the data yet again.

The other two extensions both perform well. They predict and simulate very similar to one another *and* the previous model on both sunny and cloudy days. The predictions of the metal models become slightly worse for further ahead prediction, as shown in Figure 12.13.

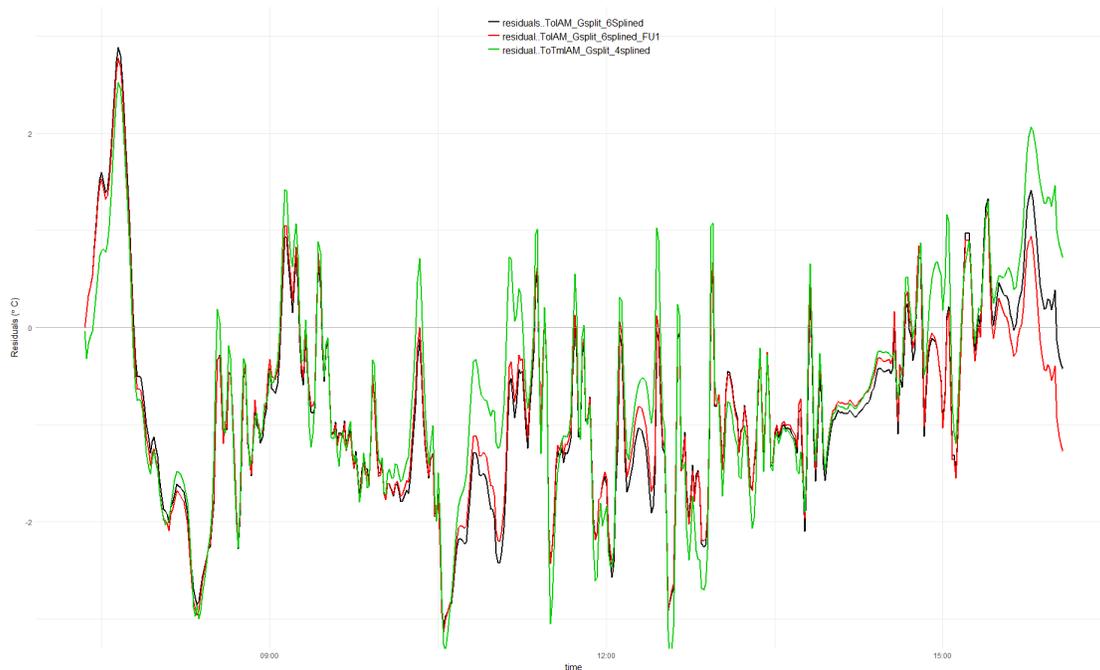


Figure 12.13: Residuals of 40-ahead prediction with the metal and FU1-model on  $D_s$ .

The estimated parameters are shown in Table 12.8. Their estimates of common parameters  $b_0$

and  $K_d$  are similar. The parameter  $(mC)_e$  has different meaning in the different models: in the metal model it captures only the heat capacity of the fluid, whereas in the single component model it represents the total module's heat capacity. Therefore it is expected to be higher in model *ToIAM\_Gsplit\_6splined\_FU1*, which is indeed the case.

parameter	<i>ToIAM_Gsplit_6splined_FU1</i>	<i>ToTm_IAM_Gsplit_6splined_FU1</i>
To0	76.75 ( 2.8e-01 )	76.79 ( 8.3e-02 )
Tm10		82.99 ( 2.9e-01 )
$b_0$	0.27 ( 8.9e-05 )	0.27 ( 1.8e-04 )
Cm		351 ( 1.5e+01 )
$F'U_0$	2.64 ( 2.3e-02 )	
$F'U_1$	0.012 ( 6e-04 )	
$K_d$	0.656 ( 5.7e-03 )	0.650 ( 2.8e-03 )
$(mC)_e$	7158 ( 8.7e+01 )	5288 ( 7.9e+01 )
$U_{fm}$		20.18 ( 6.1e-01 )
$U_{ma}$		2.985 ( 9.8e-03 )
$\sigma_m$		-0.199 ( 2.6e-02 )
$\sigma_o$	-3.16 ( 1.1e-02 )	-25.55 ( 5.4e-02 )
$\sigma_{obs}$	-29.3 ( 3.1e-02 )	-18.5 ( 5.4e-02 )

Table 12.8: Parameter estimates (standard deviation in brackets) for model *ToIAM\_Gsplit\_6splined\_FU1* and *ToTm\_IAM\_Gsplit\_6splined\_FU1*

Note that the metal model has a large value of  $\sigma_m$ , which implies large variation in this state. This is confirmed when plotting the filtered states over the intermittent day  $D_s$ , as shown in Figure 12.14. The hidden metal state, like the hidden flow compartment state, absorbs much of the un-modelled effect of clouds. This makes the metal model prone to over-fitting as well.

For practical reasons, the single state model *ToIAM\_Gsplit\_6splined\_FU1* could also be preferred over the metal model. It is easier to find a good fit for this model. It further is slightly simpler, as it contains 2 parameters and 1 state equation less. This model is therefore selected in this step.

The estimated effective solar radiation based on the selected model is as depicted in Figure 12.15a. This effective direct radiation curve includes the influence of the IAM function, with  $b_0 = 0.27$  as estimated. The estimated  $F'(\tau\alpha)_{en}K_d$  over time is plotted in Figure 12.15b.

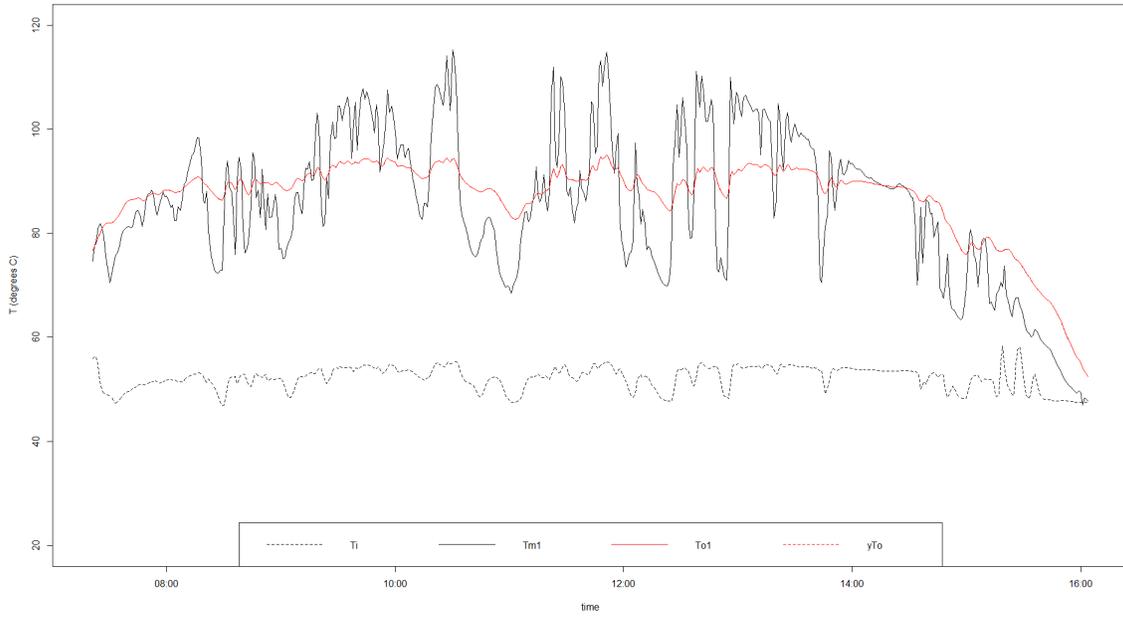


Figure 12.14: Filtered states for model *ToTm\_IAM\_Gsplit\_6splined\_FU1* on  $D_i$ .

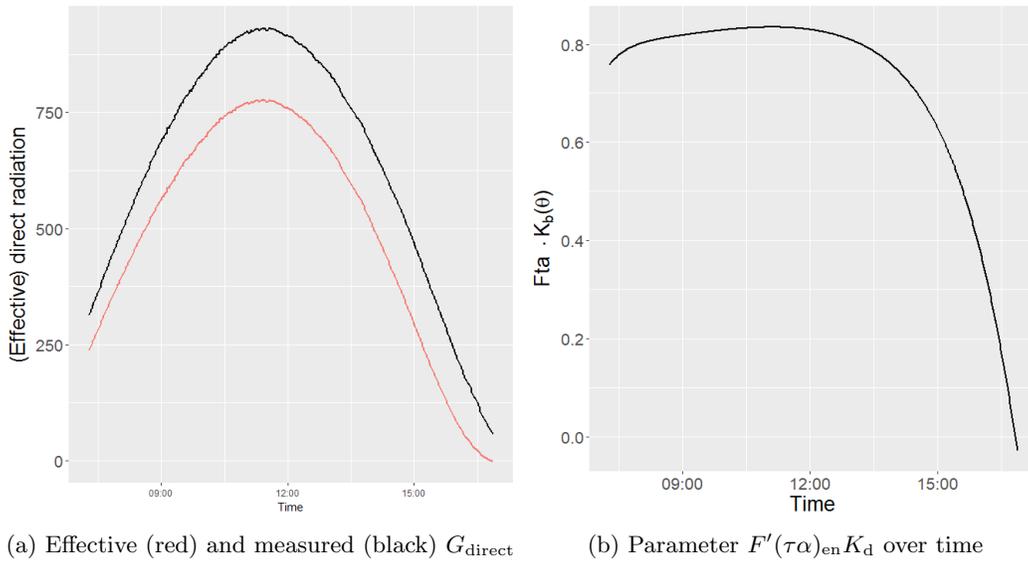


Figure 12.15: Visualizing effective direct radiation for model on  $D_s$  of model *ToIAM\_Gsplit\_6Splined\_FU1* (4<sup>th</sup> extension)

## End of model sequence

As an extension from *ToIAM\_Gsplit\_6splined\_FU1*, an attempt was made to combine the FU1 and metal model, to *ToTm\_IAM\_Gsplit\_6splined\_FU1*. The estimation of this model however turned out not to be feasible in CTSM. The same was true for the introduction of more compartments in the flow direction. Therefore, the fourth extension step was the final step in this sequence.

In the next section, two small improvements will be introduced into this model. Thereafter, an overview of the model sequence and the final model performance will be provided.

## 12.6 Fine-tuning the final model

Finally, we will consider two adaptations to the final model that are not strictly extensions, and therefore not treated as previously in the model sequence.

First of all, the way the splines are constructed will be adapted, mostly for practical reasons. As the data set used for fitting does not contain time steps in which  $0 < \cos \theta < 0.2$ , the splined radiation situated in such periods of day does not add to the quality of the model. Therefore, boundary knots will be chosen at the moments in time where  $\cos \theta = 0.2$ , so that the fitting data set  $D$  entirely contains all splined components of the radiation. In addition, the spline functions will now be constructed using a variable amount of equally spaced (in time) inner knots, while spline order will be cubic no matter how many splines are constructed. As a result, the `bs()` spline function in R constructs the splines independent of the properties of the given data, which is useful when using the fitted model on other data sets. Moreover, the cubic splines have less overlap, which is desired in our case as the morning efficiency parameter should for example not influence the efficiency in late afternoon.

As a second adaptation, the value of the inlet temperature  $T_i$  that corresponds to outlet temperature  $yTo(t_k)$  at a certain point in time will be altered, in order to account for the time that the water resides in the collector rows. This will be done in two ways: first by assuming an average delay of 10 minutes, thereafter by estimating the delay for each data point individually. The delays are not constant over time, because the flow is adapted during the day.

### 12.6.1 Adapted splines

The final model *ToIAM\_Gsplit\_splined\_FU1* was fit with the new spline boundary knots for 4 up to 6 splines. The loglikelihoods of the resulting fits (named *final\_model\_nk\_Xspline* where nk stands for “New Knots”) and the corresponding fits with the previous spline method are given in Table 12.9. For each number of splines, the loglikelihood is higher for the new splines.

The prediction and simulation of the 6-spline model were compared in plots. The only visible improvement of new spline method is observed in the later afternoon of clear sunny day  $D_s$ , both in simulation and prediction. To illustrate this, the 10-ahead prediction residuals on this day are shown in Figure 12.16.

	nr. of par.	Loglikelihood
ToIAM_Gsplit_4splined_FU1	12	-1624.75
final_model_nk_4spline	12	-1419.37
ToIAM_Gsplit_5splined_FU1	13	-1415.14
final_model_nk_5spline	13	-1408.00
ToIAM_Gsplit_6splined_FU1	14	-1404.41
final_model_nk_6spline	14	-1390.89

Table 12.9: Comparison between the previous and new spline methods for the final model *ToIAM\_Gsplit\_splined\_FU1*.

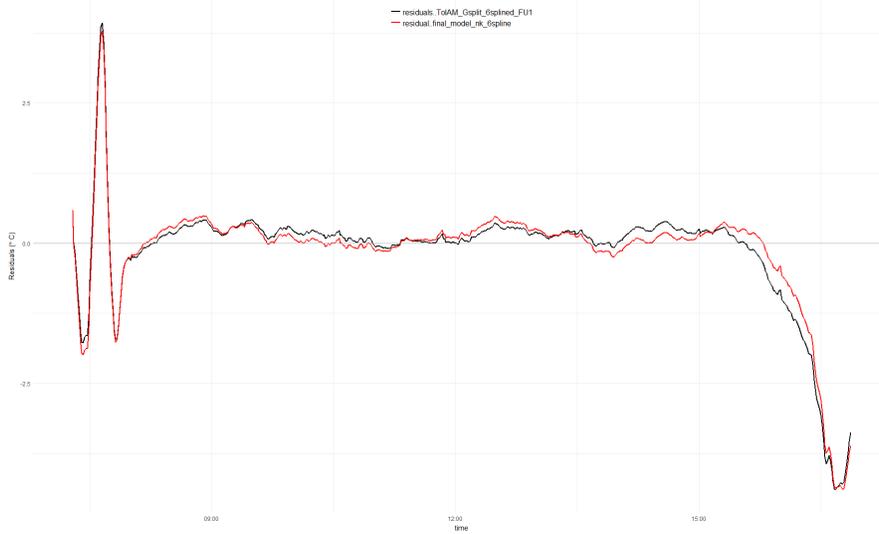


Figure 12.16: Residuals on  $D_s$  for the final model with old and new spline method

## 12.6.2 Delayed inlet temperature

The models in all previous sections related the system equation increment  $dT_o$  at time  $t_k$  to the inlet temperature  $T_i$  at the same time point  $t_k$ . However, as the water resides within the collector rows for 3 up to as long as 20 minutes, this instantaneous effect of the inlet on the outlet temperature is not very accurate.

To solve this, two ways of delaying the inlet temperature have been implemented. First, all  $T_i$  measurements were shifted by 10 minutes (roughly corresponding to the delay for average flow in the data set). The second method computes the delay for each outlet temperature measurement  $yT_o$ , based on the flow of the preceding minutes. In this way, it is possible to approximate when the water that flows out at  $t_k$  entered the collector row. In this way, the temperature of a column of out-flowing water is related to the temperature of this same column of water when it flowed into the row.

These two methods have been applied individually, and combined with the new spline method. The resulting loglikelihoods of these fits are shown in Table 12.10. The abbreviations *nk*, *dnk*, and *dsnk* are for Delayed, Smart Delayed, and New Knots, in different combinations.

The effect of the delayed  $T_i$  is evident on intermittent days and shortly in the morning and afternoon of the clear sunny day, which are the moments that the inlet temperature fluctuates most. The

	Loglikelihood
ToIAM_Gsplit_6splined_FU1	-1404.41
final_model_nk_6spline	-1390.89
final_model_Delayed	-1231.17
final_model_SmartDelay	-1142.87
final_model_dnk_6spline	-1222.25
final_model_dsnk_6spline	-1138.77

Table 12.10: Loglikelihoods of adapted *ToIAM\_Gsplit\_6splined\_FU1* versions

morning and afternoon peaks are reduced on  $D_s$  in 1-ahead prediction for both delay methods, but only shown for smart delay in Figure 12.17 (inspect more closely in Figure 12.17b).

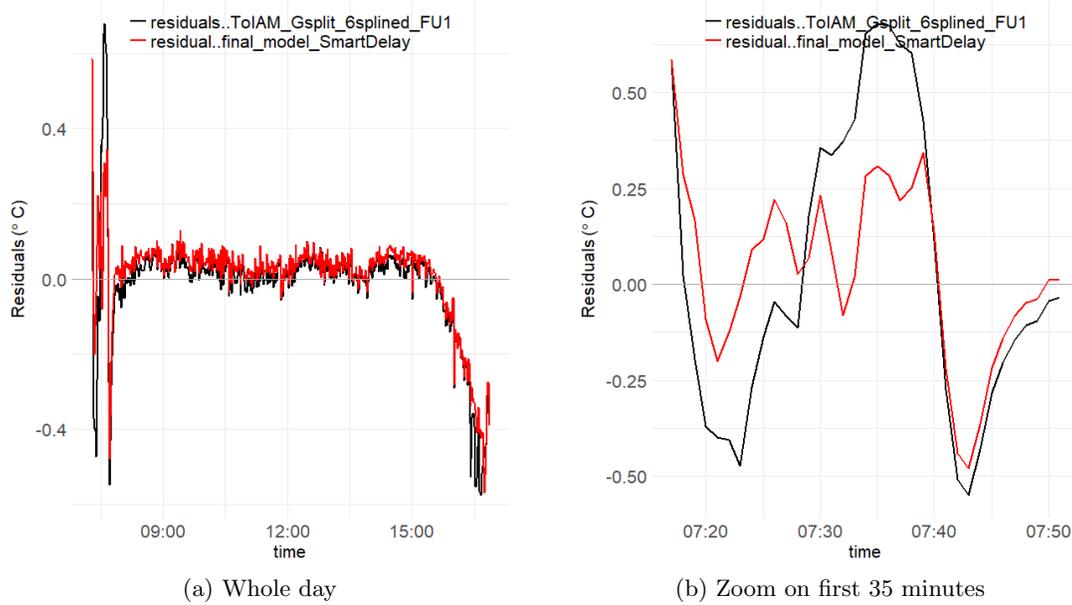
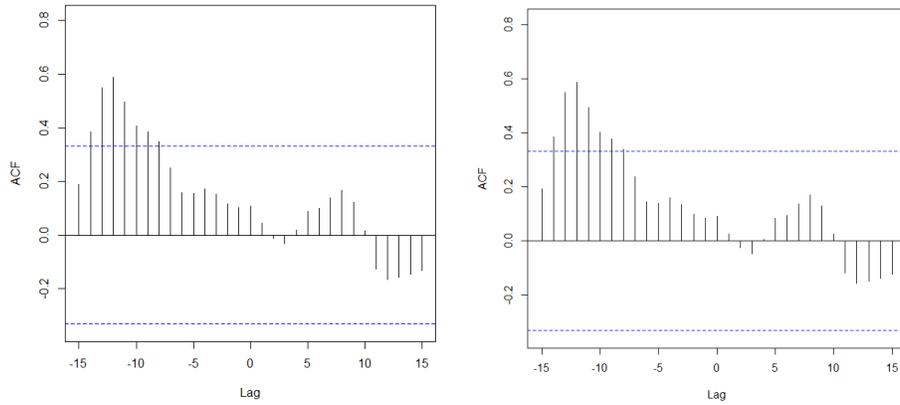


Figure 12.17: Residuals of final model and smart delayed model on  $D_s$

The improved CCF plot between inlet temperature and sunny morning residual is shown in Figure 12.18a (compare to model  $T_o$ , Figure 12.5a).

The inlet temperature delay also reduced some residual peaks on cloudy day  $D_i$ , indicating that some of the residual there was also a result of inlet temperature fluctuation. Figure 12.19 shows the new residuals compared to those of *ToIAM\_Gsplit\_splined\_FU1* on this day.



(a) Using model *final\_model\_SmartDelay*    (b) Using model *final\_model\_dsnk\_6spline*

Figure 12.18: CCF between residuals and  $T_i$  on first 35 minutes of  $D_s$ . Negative lag corresponds to the relation between past inlet temperature and a later residual value.

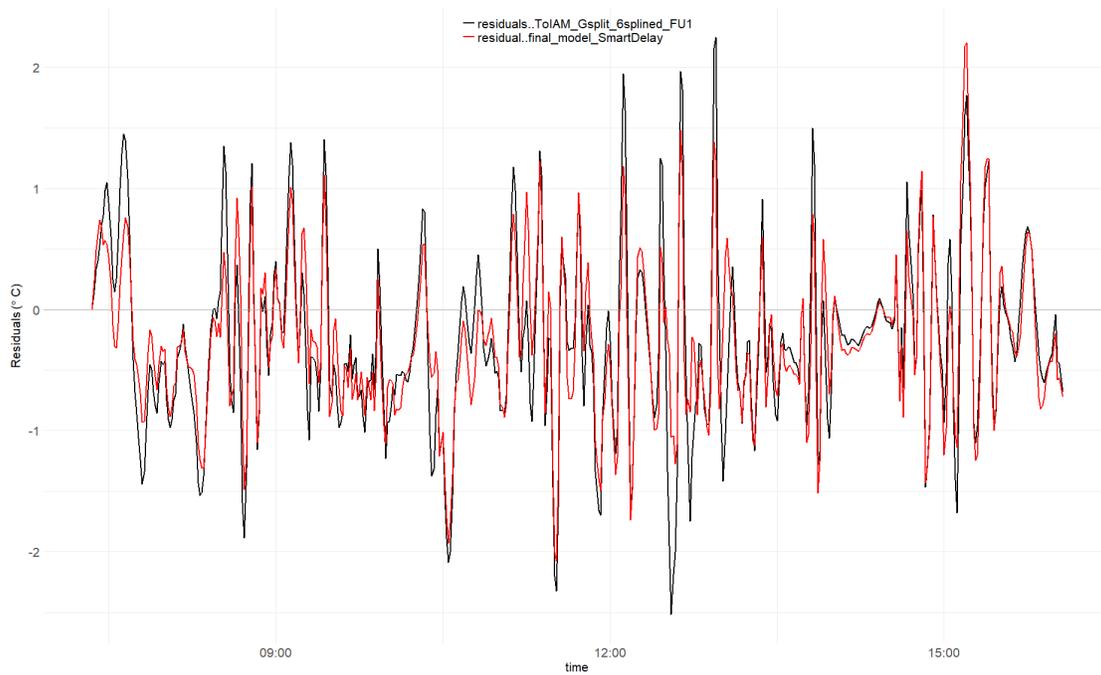


Figure 12.19: Residuals of 5-ahead prediction on  $D_i$  for model *ToIAM\_Gsplit\_splined\_FU1* and the adapted versions with smart delay

### 12.6.3 Combined effect

The models with both the new spline method and the (fixed or smart) delayed inlet temperature combine the improvements of the two different adaptations. The combined ‘smart delay & new knots’ method has the highest loglikelihood (Table 12.10). This model *final\_model\_dsnk\_6spline* retains the CCF improvement between residuals and inlet temperature in the morning of  $D_s$ , as shown in Figure 12.18b, and the time series on the sunny day shows the same reduced morning peaks (plot omitted here). The same holds for the reduced residual peaks on the cloudy day, see Figure 12.20.

Model *final\_model\_dsnk\_6spline*, is thus performing better than *ToIAM\_Gsplit\_6splined\_FU1*, and this ‘fine-tuned’ model will therefore be accepted.

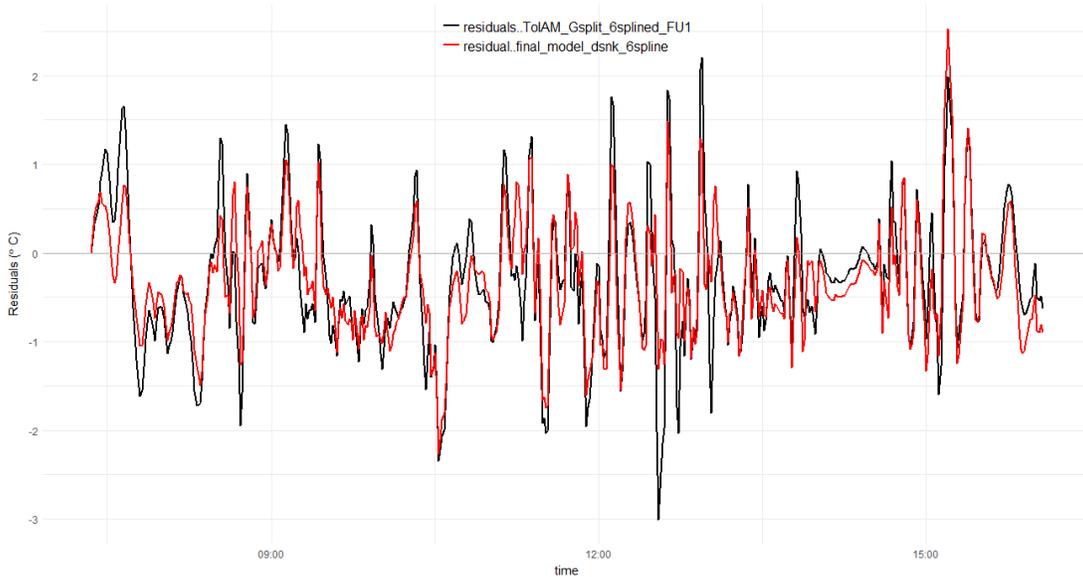


Figure 12.20: Residuals of 5-ahead prediction on  $D_i$  for model *ToIAM\_Gsplit\_splined\_FU1* and the adapted version with smart delay and new spline method

## 12.7 Summary model sequence

This section will provide an overview of the model sequence, as well as a comparison of the different models. It further intends to highlight how each of the extensions improved the model performance.

Table 12.11 shows for each chosen model the number of parameters, loglikelihood, and the LRT p-value of the extension wrt the previous model. Over the sequence, the number of parameters increased from 6 to 14. Of the latter, 6 parameters are spline efficiency coefficients. The loglikelihood increased from  $-6586$  to  $-1134$ . All extensions were very significant based on the LRT. The final model is not strictly an extension of the previous, and therefore the LRT is not applied.

Model	nr .of par.	Loglikelihood	LRT p-value
Tf	6	-6585.86	
To	6	-2125.08	0.00
To_4Splined	9	-1589.95	0.00
fit_ToGsplit_Gsplit_4Spline	10	-1473.82	0.00
ToIAM_Gsplit_6Splined	13	-1412.30	0.00
ToIAM_Gsplit_6splined_FU1	14	-1404.41	0.00
ToIAM_Gsplit_6splined_FU1_sdnk	14	-1133.96	

Table 12.11: Summary of Loglikelihood and p-value of LRT for all selected models in sequence

## Parameter estimates

The parameter estimates have changed over the selected models in the sequence. Table 12.12 summarizes these values. As noted before, in this research we are not aiming to find some ‘true’ parameter values. Instead we are interested in a model with parameters that fit the data well when used for prediction and simulation. As the models were derived from the ISO standard heat balance however, the parameters should be reasonably close to their expected values from such a test. It is useful for the parameters to keep their original interpretation, a.o. for monitoring plant performance.

Although the model selection procedure did not take the parameter values into account as a selection criterion, it turns out that the parameter estimates over the sequence moved towards the values expected from ISO tests. Especially  $(mC)_e$  showed a steady decrease from a very high ( $> 20.000$ ) towards a final value of 6540. This is more in the range as found by Bacher et al. [14]. The same happened for first order heat loss parameter  $F'U_0$ , which in the panel specifications has value 2.1 for double glass and 3.1 for single glass. Diffuse radiation parameter  $K_d$  decreased from 0.71 to 0.58 over the last four selected models. As this parameter replaces the factor  $F'(\tau\alpha)_{en} \cdot K_d$  in the ISO model, it is expected to be of a value around  $0.81 \cdot 0.9 \approx 0.73$ .

The second order heat loss coefficient  $F'U_1$  has assumed values very comparable to the panel specifications 0.013 for single and 0.016 for double glass. The IAM parameter  $b_0$  has a rather high value in all models, but in these models it cannot be interpreted without considering the influence of the splines. It will therefore be discussed using efficiency plots in the following.

par.	..._FU1_sdnk	..._FU1_	ToIAMGsplit_6spl	
To0	76.75	76.75	76.75	
b0	0.27	0.27	0.27	
FU0	2.20	2.64	3.21	
FU1	0.010	0.012		
Kd	0.58	0.66	0.66	
mC	6540	7156	7063	
s1	1.42	1.93	1.95	
s2	0.26	0.0666	0.010	
s3	1.43	1.493	1.544	
s4	0.14	0.2015	0.139	
s5	1.46	1.521	1.546	
s6	0.12	$10^{-5}$	$2 \cdot 10^{-5}$	
sigma_o	-3.22	-3.16	-3.16	
sigma_obs	-23.69	-29.28	-29.85	
par.	ToGsplit_4Spline	To_4Spline	To	Tf
Tf0				66.3
To0	76.75	76.75	76.75	
Fta			0.98	1.0
FU0	3.62	4.52	6.24	6.8
Kd	0.71			
mC	7558	8642	13294	21408
s1	0.99	1.00		
s2	0.48	0.63		
s3	1.47	1.30		
s4	$8.8 \cdot 10^{-5}$	0.31		
sigma_o	-3.14	-3.12	-2.97	-2.6
sigma_obs	-29.76	-13.34	-13.17	-21.9

Table 12.12: Parameter estimates of all selected models.

## Correlation between parameters

The correlation table improved significantly from the initial to the final model. The correlation between the performance parameters  $b_0$ ,  $F'U_0$ ,  $F'U_1$ ,  $K_d$ , and  $(mC)_e$  remains below an absolute value of 0.41, except for the correlation between the heat loss coefficients. The latter is expected to occur, as the terms together need to balance to model the heat loss best.

The occurrence of extremely high correlations of, say, higher than 0.8 can be an indication of too little variation in the data set, so that the effect of different parameters can hardly be distinguished. The current model parameter correlations take acceptable values, indicating that the fitting data set was appropriate for this model.

	To0	b0	FU0	FU1	Kd	mC	$\sigma_o$	$\sigma_{obs}$
To0	1.00	0.24	0.03	-0.14	-0.08	-0.10	0.08	-0.12
b0	0.24	1.00	0.14	-0.38	0.07	-0.40	0.19	-0.72
FU0	0.03	0.14	1.00	-0.75	0.41	-0.08	-0.02	-0.10
FU1	-0.14	-0.38	-0.75	1.00	0.14	0.21	-0.10	0.02
Kd	-0.08	0.07	0.41	0.14	1.00	-0.19	-0.08	-0.34
mC	-0.10	-0.40	-0.08	0.21	-0.19	1.00	-0.14	0.15
$\sigma_o$	0.08	0.19	-0.02	-0.10	-0.08	-0.14	1.00	-0.14
$\sigma_{obs}$	-0.12	-0.72	-0.10	0.02	-0.34	0.15	-0.14	1.00

Table 12.13: Model *ToIAM.Gsplit.6splined.FU1.sdnk*, correlation table. Removed the efficiency parameters *s1* to *s6*.

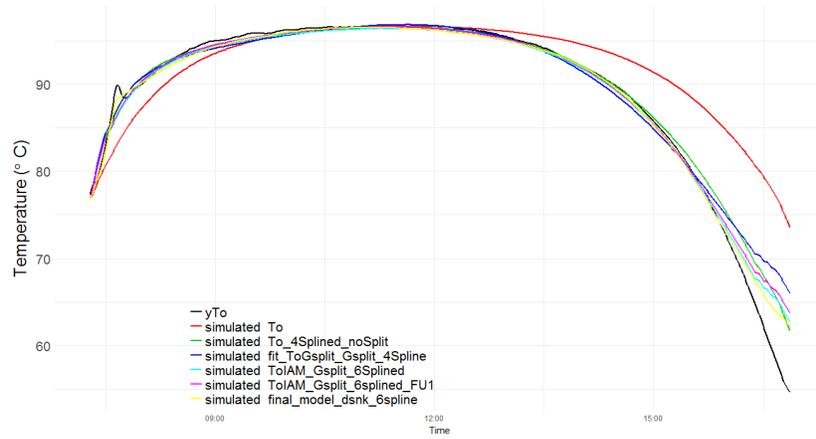
## Efficiency of radiation absorption

The radiation absorption was first modelled as a constant, then extended to a function of Time of Day using splines. Finally, the constant was both a function of Time of Day and of the incidence angle  $\theta$ . This led to changing efficiency curves over the modelling sequence, as plotted in Figures 12.9, 12.11b, and 12.12b.

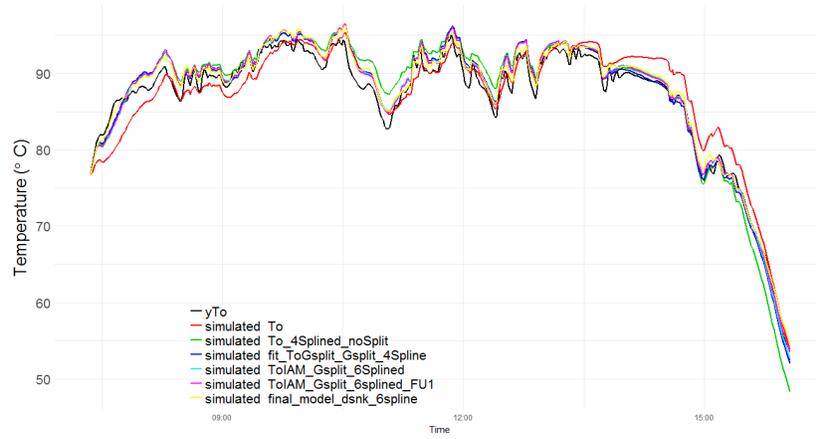
## Simulation and prediction

The simulation capabilities of the models steadily improved, although the size of the improvement decreased over the sequence. For the clear sunny day  $D_s$  and intermittent day  $D_i$ , the simulations are plotted in Figure 12.21a and 12.21b. This plot mainly intends to show that all models were reasonable for simulation purposes, *given perfect predictions of inputs*.

The 1-step ahead prediction residuals have evolved as displayed in Figures 12.22 and 12.23. On sunny days, the main improvement is seen in the early morning and late afternoon. On the cloudy day, the small improvements in 1-ahead prediction residuals can hardly be seen through all fluctuations. The improvement in predictive performance of the models becomes more apparent when plotting further-ahead prediction (Figure 12.24 and 12.25).



(a)  $D_s$



(b) on  $D_i$

Figure 12.21: Simulated outlet temperature for the models in the sequence

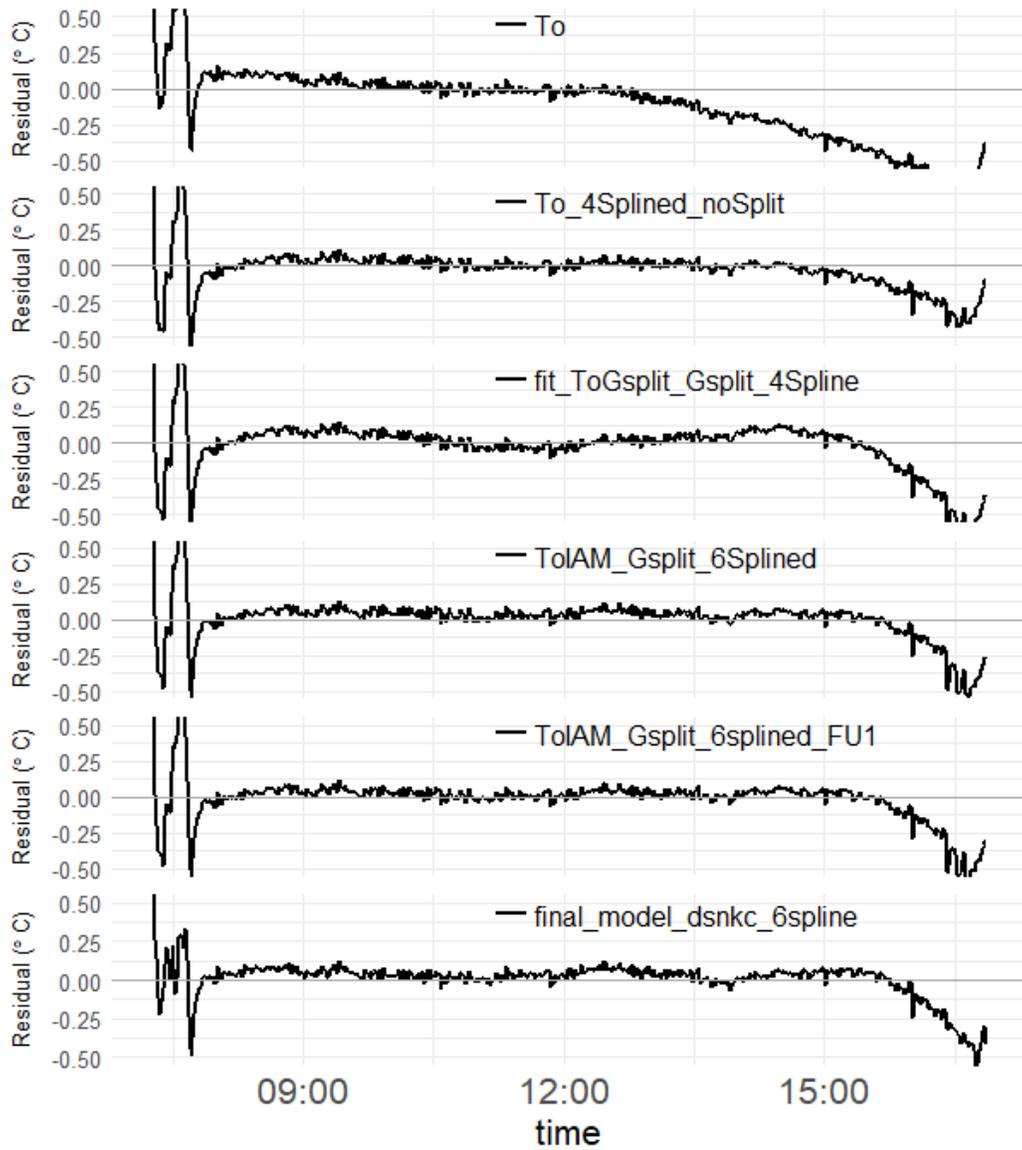


Figure 12.22: Prediction (1-ahead) residuals on  $D_s$  for the models selected in the sequence

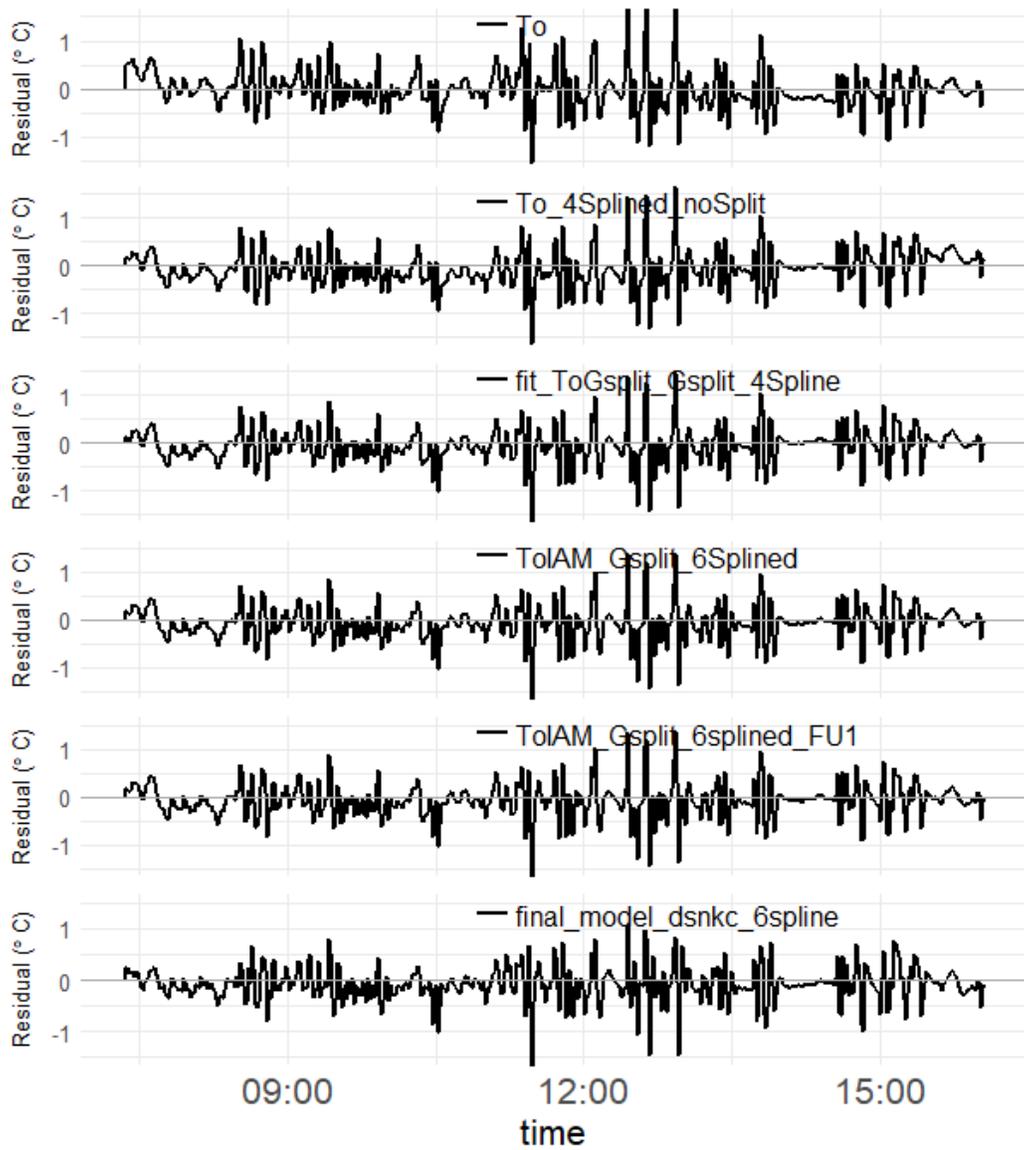


Figure 12.23: Prediction (1-ahead) residuals on  $D_i$  for the models selected in the sequence

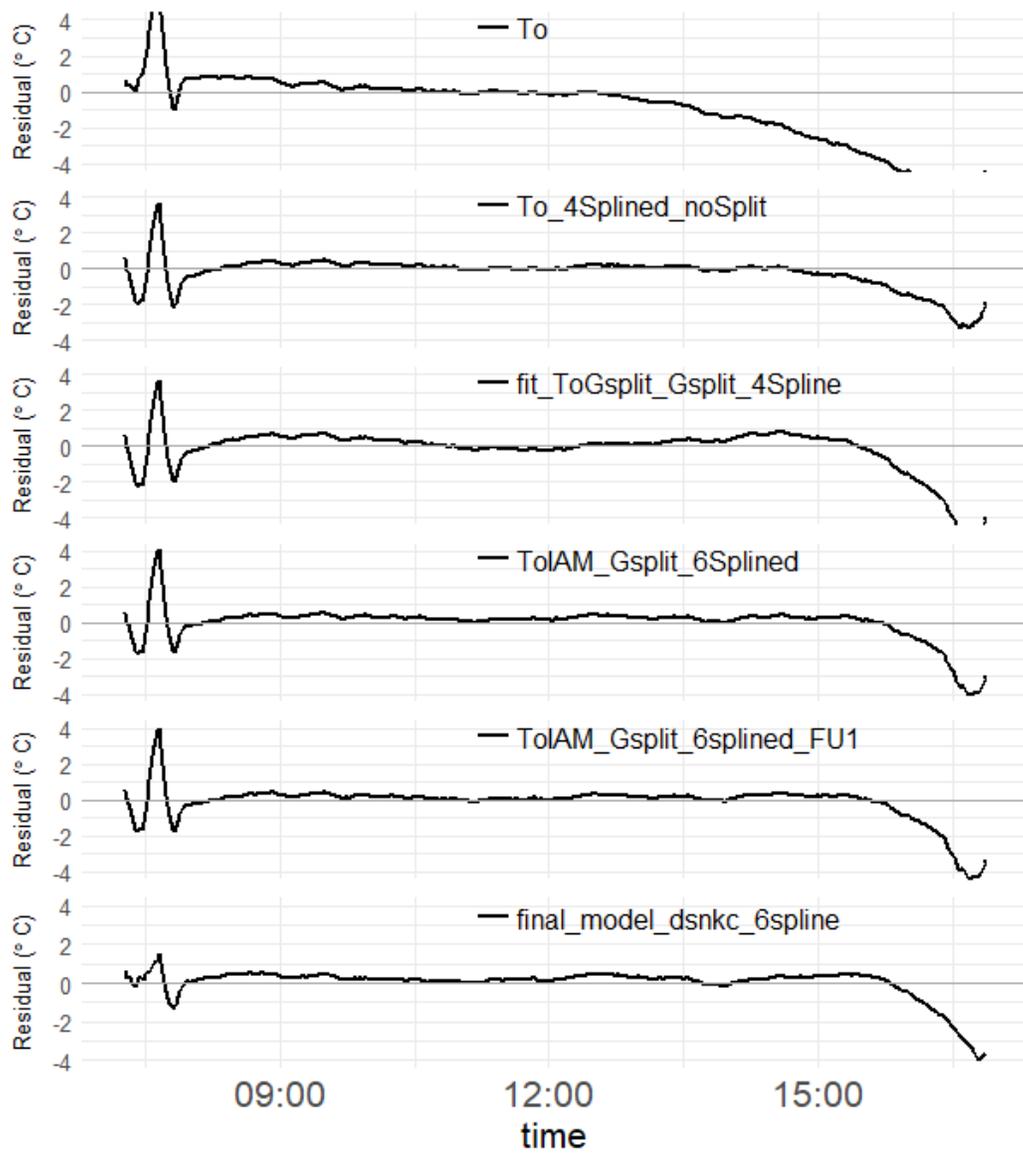


Figure 12.24: Prediction (10-ahead) residuals on  $D_s$  for the models selected in the sequence

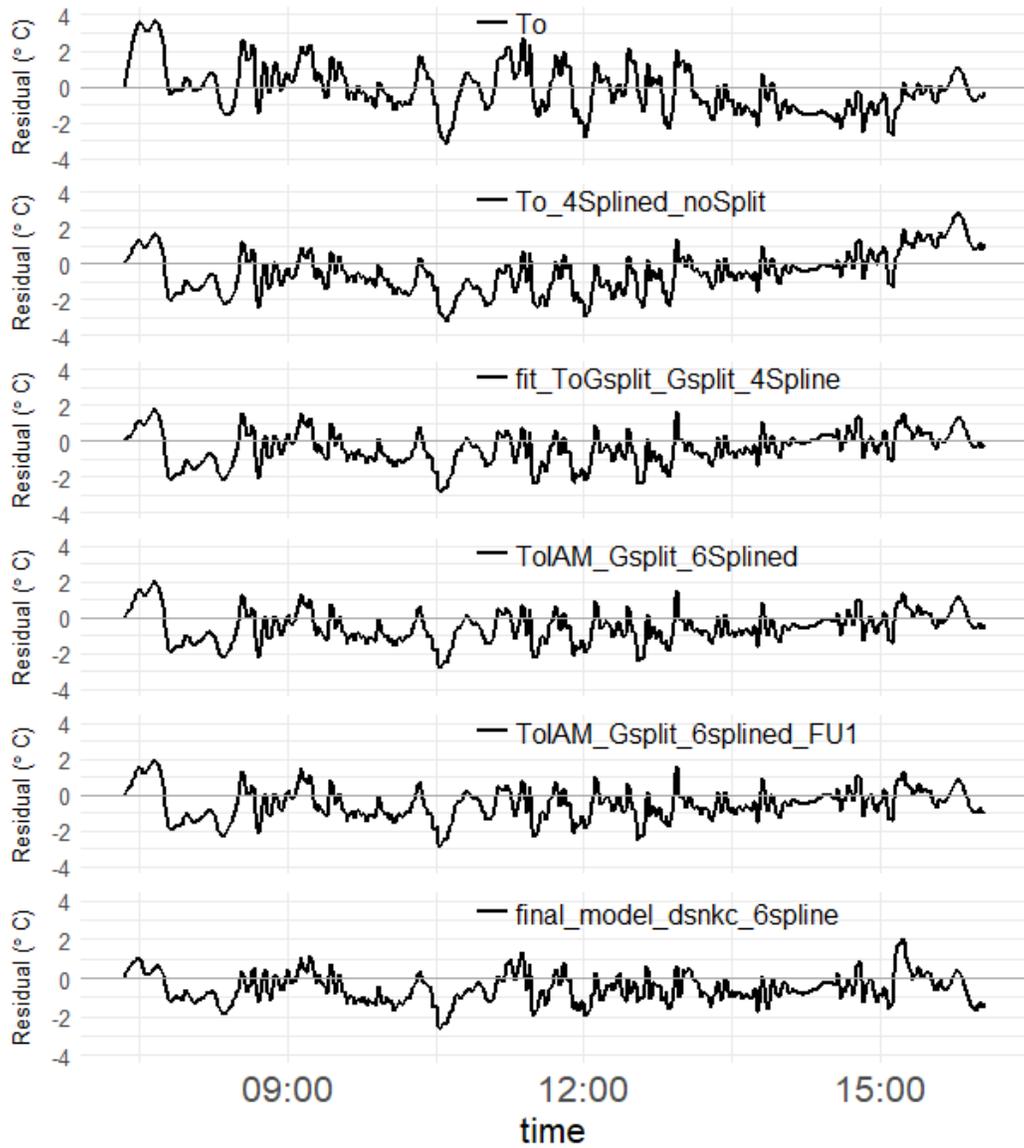


Figure 12.25: Prediction (10-ahead) residuals on  $D_i$  for the models selected in the sequence

## Prediction residuals ACF

Figure 12.26 shows the ACF plot of 1-ahead prediction residuals based on the full data set  $D$ . It can be seen that the white noise properties have improved over the model sequence, as the correlation bars are reduced in size and less frequently exceed the confidence band. The largest improvements are seen in the first two model extensions. Little change is observed from model 4 to 6, although the size of lag-1 correlation is reduced. The lag-1 autocorrelation is however significant for all model versions.

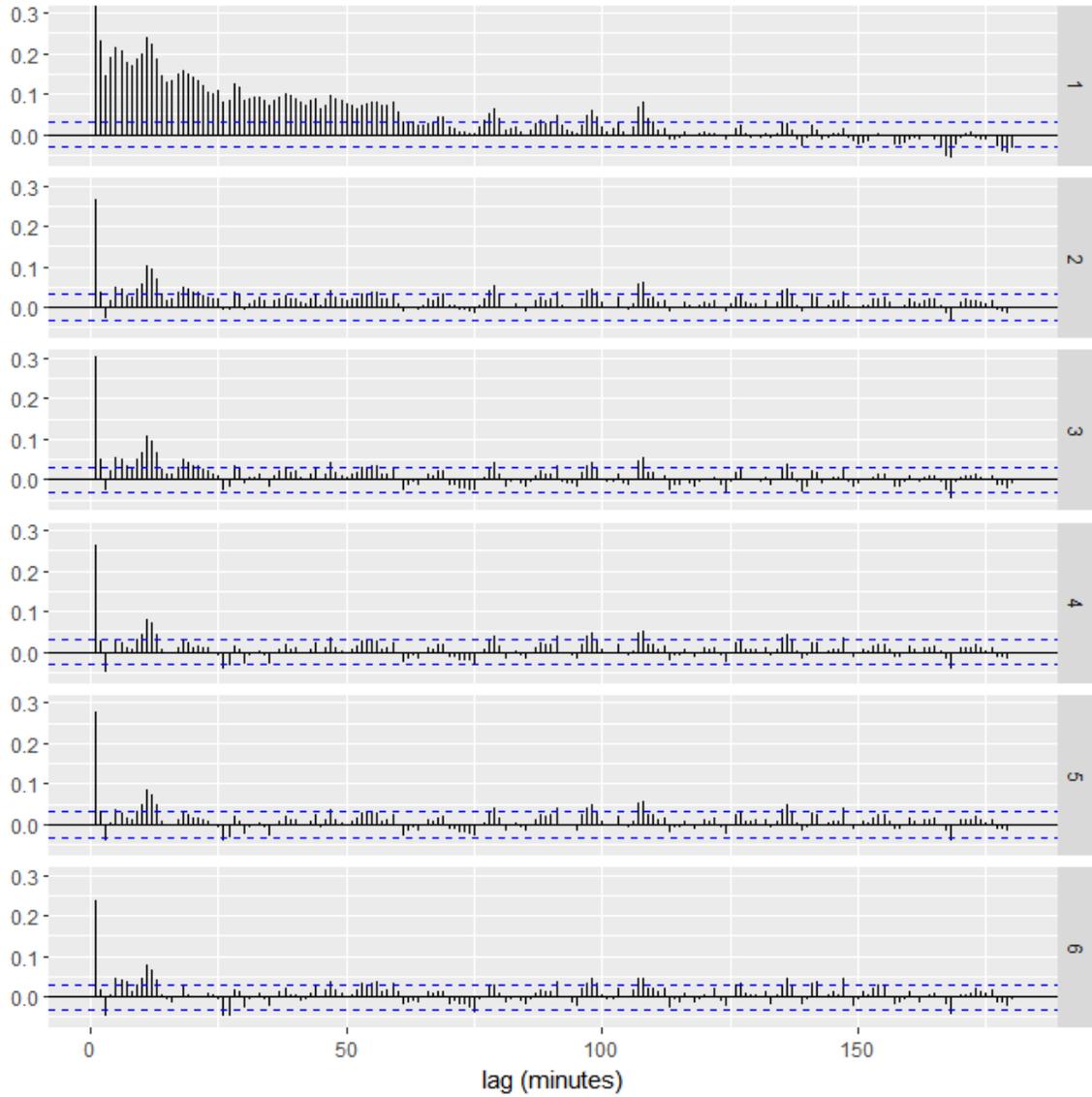


Figure 12.26: Overview of normalized residuals' ACF over the model sequence, for lag 1 up to lag 180 minutes. The numbers on the right hand side denote the number of the model in the sequence, where 1 corresponds to model *To* and 6 to final fine-tuned model *ToIAM\_Gsplit\_6splined\_FU1\_sdnk*. Note that lag 0 is not displayed, as this value is 1 by definition.

## Prediction residuals CCF

The residual time series and past values of input variables should be independent. This is checked using a CCF plot between the residuals and a particular input variable.

As we were interested in reducing the inlet temperature dependence of the model, the CCF for the residuals with this variable is plotted in Figure 12.27. The main improvements are seen in the extension from model *To* to *To\_splined*, and after fine-tuning (model 6). The final model still has significant correlation at lag 0.

We have further investigated the prediction residuals' dependence on radiation, where the effects of both diffuse and total radiation were of interest.

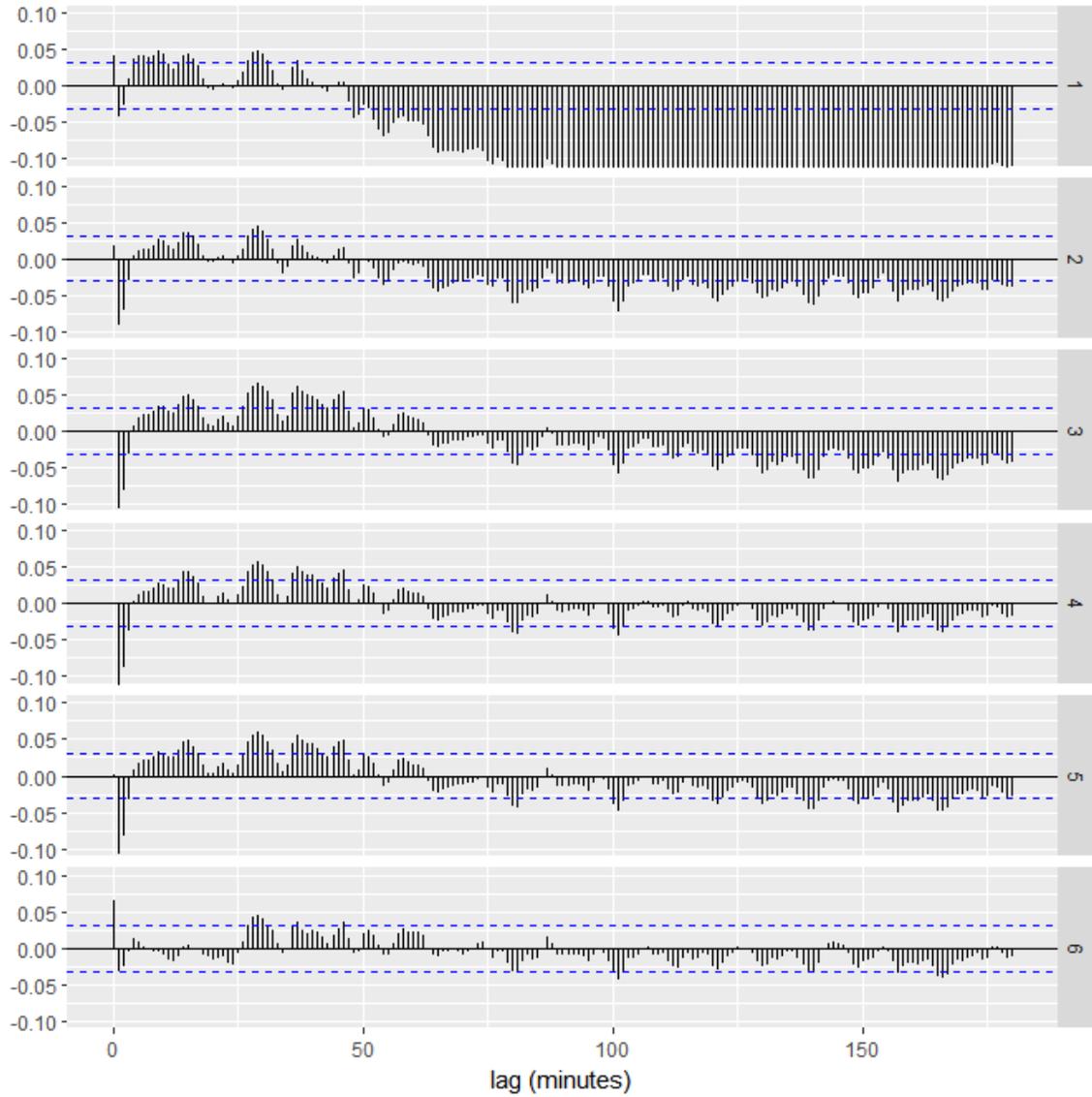


Figure 12.27: The CCF between normalized residuals and inlet temperature over the model sequence, for lag 1 up to lag 180 minutes. The numbers on the right hand side denote the number of the model in the sequence, where 1 corresponds to model *To* and 6 to final fine-tuned model *ToIAM\_Gsplit\_6splined\_FU1\_sdnk*. Note that lag 0 is now displayed.

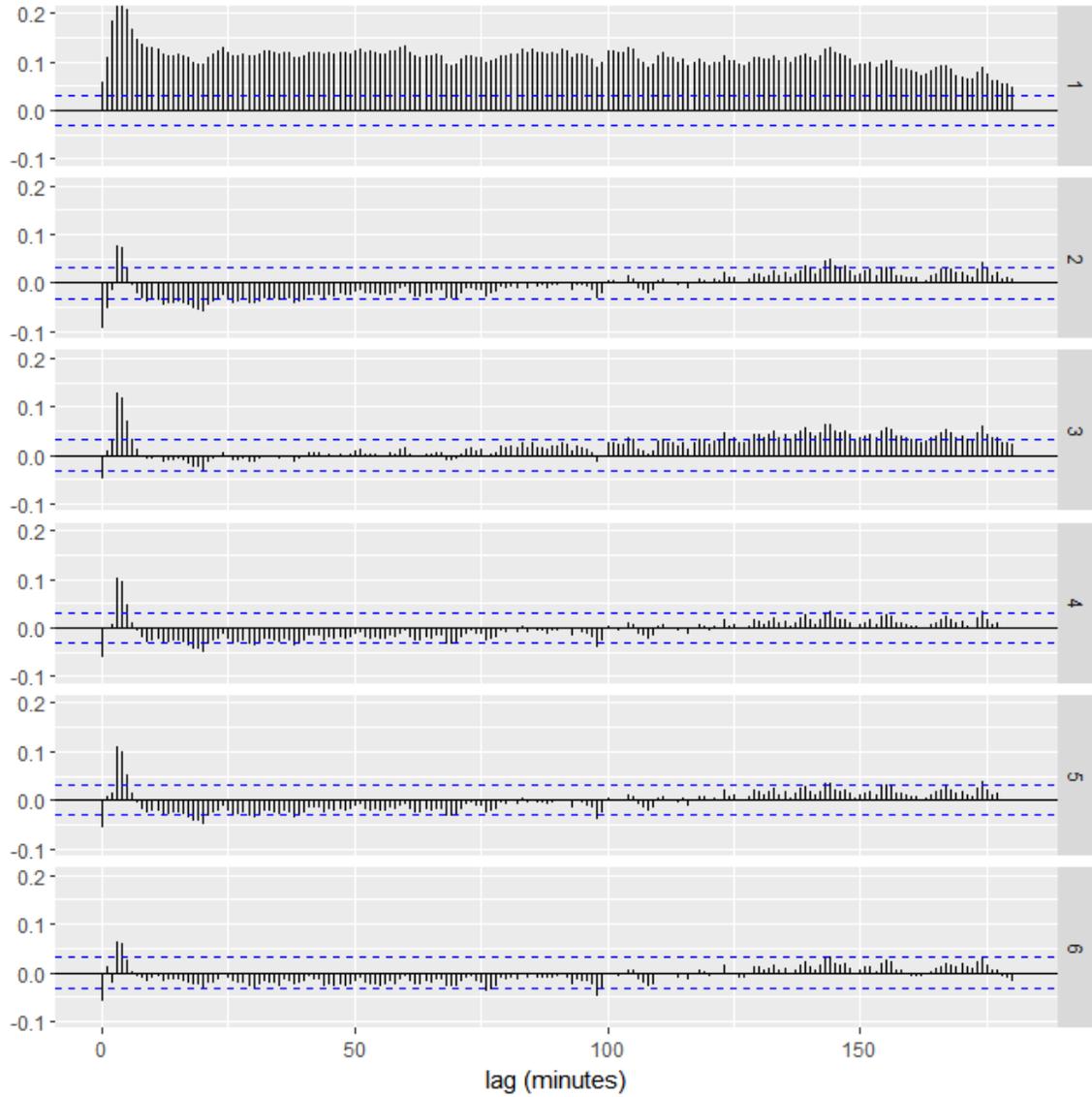


Figure 12.28: The CCF between normalized residuals and diffuse radiation over the model sequence, for lag 1 up to lag 180 minutes. The numbers on the right hand side denote the number of the model in the sequence, where 1 corresponds to model *To* and 6 to final fine-tuned model *ToIAM\_Gsplit\_6splined\_FU1\_sdnk*. Note that lag 0 is now displayed.

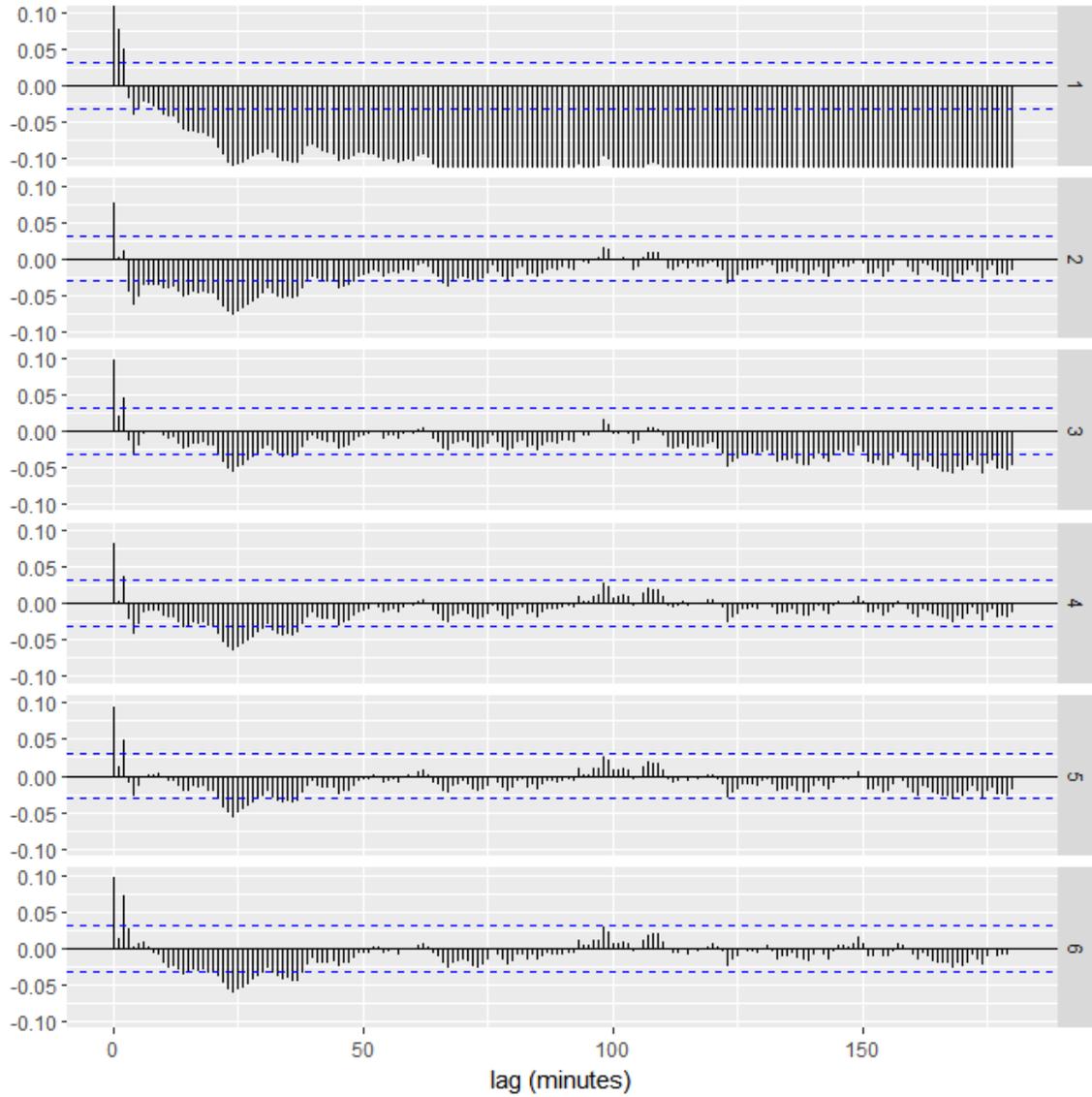


Figure 12.29: The CCF between normalized residuals and total radiation over the model sequence, for lag 1 up to lag 180 minutes. The numbers on the right hand side denote the number of the model in the sequence, where 1 corresponds to model  $T_0$  and 6 to final fine-tuned model  $ToIAM\_Gsplitted\_6splined\_FU1\_sdnk$ . Note that lag 0 is now displayed.

# Chapter 13

## Final model performance

This chapter intends to give insight in the performance of the final model, regarding both prediction and simulation. This will include a discussion of model performance on data from other months than the fitting data.

In the first section, the predictive performance of the final model on the fitting data  $D$  is analyzed in more detail, and prediction errors are quantified. Next, in Section 13.2, the simulation performance of the final model will be compared to the *MLR2* model and the model currently used for control in Solrød. Finally, Section 13.3 shows the results of a cross-validation on data sets from other times of year.

### 13.1 Final model predictive performance on fitting data

This section will provide a short overview of final model performance in prediction. We will quantify the prediction error, and show how it evolves when the prediction horizon is increased. The reader is reminded that in  $n$ -ahead prediction, we assume perfect knowledge or prediction of future inputs, but no knowledge of the state estimate within the  $n$ -step prediction period.

The tables below show the mean, maximum, and minimum of the absolute prediction error, for several values of the prediction horizon  $n$  (where ‘sim’ means full day prediction). A distinction is made between the sunny and cloudy day, as they have quite different sizes of the error. Comparing Table 13.1 and 13.2, we see that the average error is doubled on  $D_i$  compared to  $D_s$  for all prediction horizons  $n$ . However, the maximum error on  $D_s$  blows up for larger prediction horizons. Plots showed that this blow up occurs in the last 35 minutes of the sunny day, where we have in the model sequence seen structural overestimation of the outlet temperature.

n ahead	1	5	10	60	300	sim
mean	0.24	0.59	0.76	1.05	1.06	1.06
max	1.68	2.56	2.67	2.85	2.85	2.85
min	1.82e-04	1.82e-04	1.82e-04	1.82e-04	1.82e-04	1.82e-04

Table 13.1: Absolute error (mean, max, min) of final model on intermittent day  $D_i$ , various step ahead prediction

n ahead	1	5	10	60	300	sim
mean	0.07	0.29	0.45	0.67	0.66	0.66
max	0.59	2.37	3.95	7.77	7.76	7.76
min	4.70e-05	5.27e-04	5.19e-03	3.25e-05	1.79e-04	1.79e-04

Table 13.2: Absolute error (mean, max, min) of final model on sunny day  $D_s$ , various step ahead prediction

For better overview of the rest of the sunny day's performance, Table 13.3 shows the error measures for that day with removal of the last 35 minutes. It shows that on the rest of  $D_s$ , the error remains below 2°C even for the full simulation.

n ahead	1	5	10	60	300	sim
mean	0.06	0.20	0.30	0.38	0.38	0.38
max	0.59	1.50	1.47	2.00	1.98	1.98
min	4.70e-05	5.27e-04	5.19e-03	3.25e-05	1.79e-04	1.79e-04

Table 13.3: Absolute error (mean, max, min) of final model on sunny day  $D_s$  with last 35 minutes removed, various step ahead prediction

The evolution of the maximum absolute error for the different days is plotted in Figure 13.1. The maximum error does not increase further when  $n$  is raised above 60 minutes. The sunny afternoon reaches a maximum absolute error of as large as 8°C, whereas the rest of the day it remains approximately below 2°C. On the cloudy day the absolute error remains below 3°C for all prediction horizons.

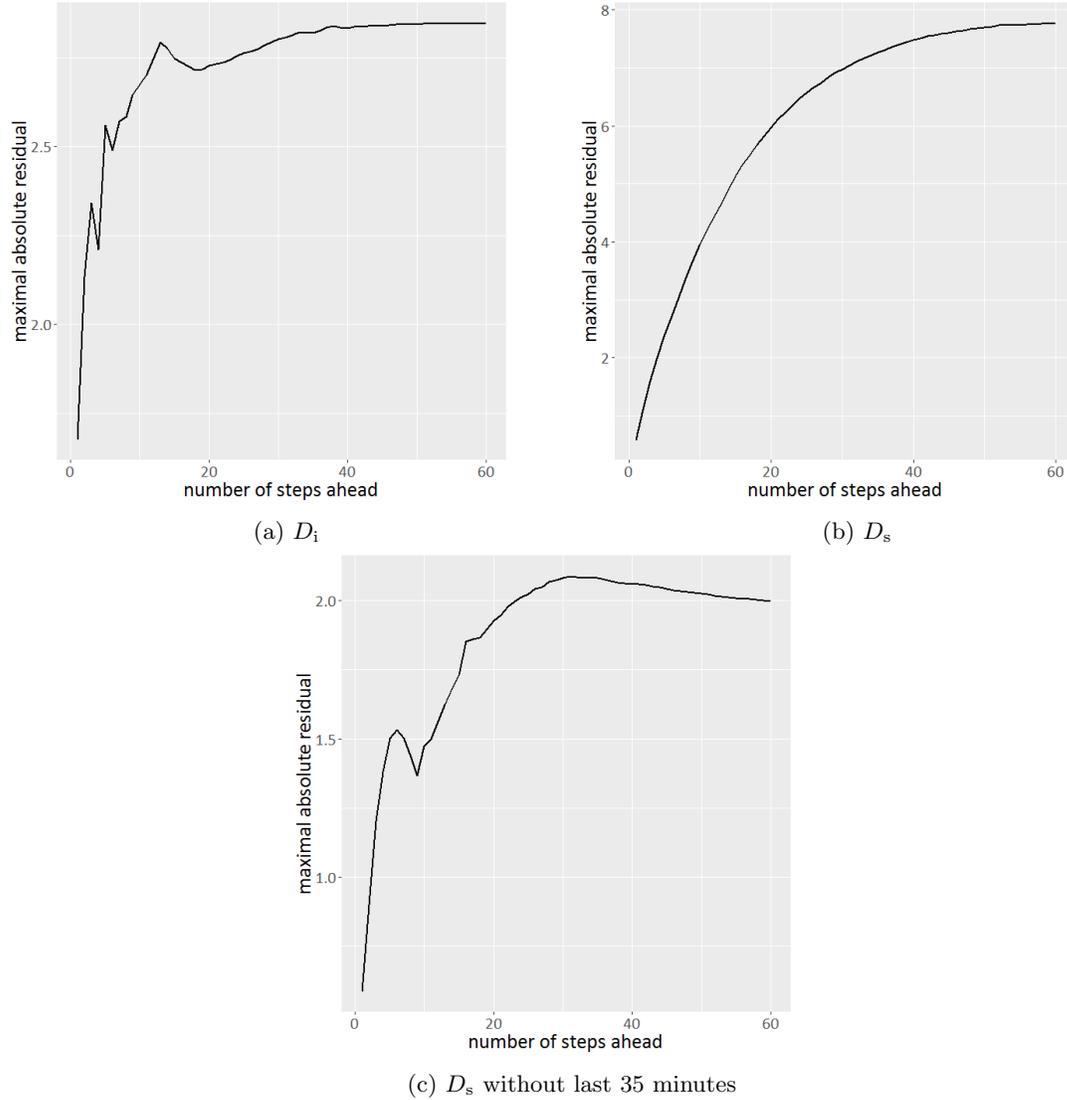


Figure 13.1: Evolution of maximum absolute error when increasing prediction horizon

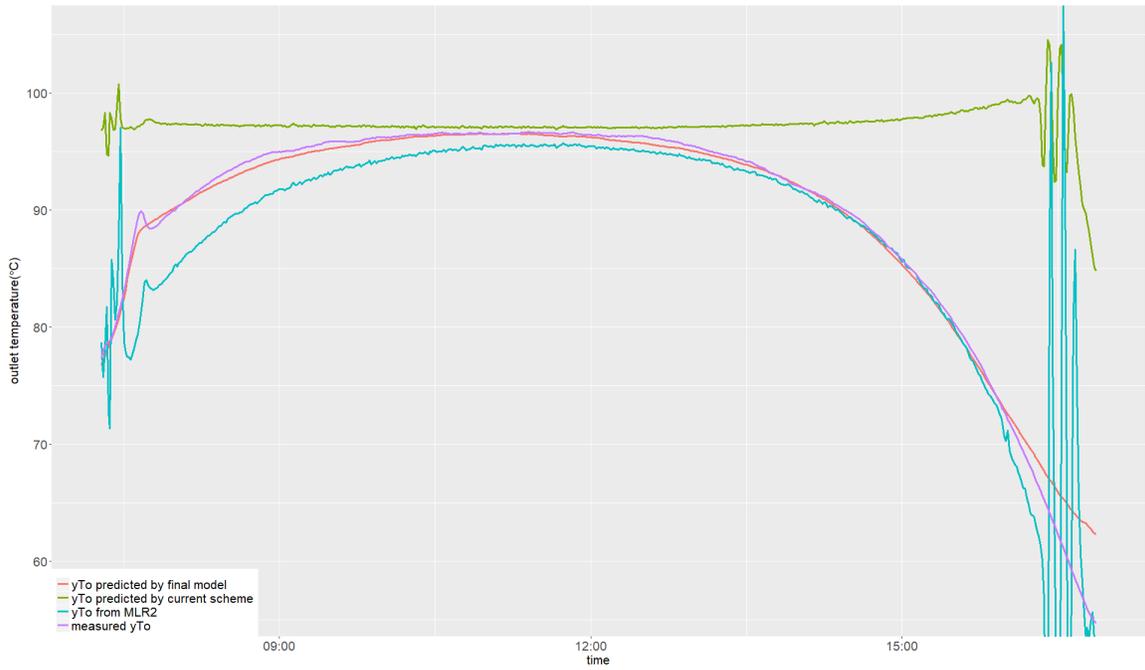
## 13.2 Comparison to MLR2 and current Solrød scheme

In order to evaluate the performance of the newly developed model, we compare its simulation to the simulated values from model *MLR2* and those simulated using the equation as currently used in Solrød. These simulations are plotted in Figure 13.2a and 13.2b for  $D_i$  and  $D_s$ .

The plots show that the developed model follows the measured outlet temperature better than the current scheme simulation as well as the MLR fit. The latter two curves show large fluctuations when  $T_i$  changes. The direct presence of  $T_i$  in the computation of  $T_o$  for these models is a contributor to this effect, as well as the fact that these models do not consider any influence of  $T_o(t_k)$  on  $T_o(t_{k+1})$ .



(a) On cloudy day  $D_i$



(b) On sunny day  $D_s$

Figure 13.2: Simulated and true outlet temperature using 3 different models

### 13.3 Cross-validation on other data sets

In order to ensure that the developed model is suitable for all normal circumstances of the plant, the model is cross-validated on data sets from other times of year. This is done in two ways. First, the model fit from May data is used for prediction in other months. Second, the model parameters are re-estimated on the data from other months.

#### 13.3.1 Without re-estimation of the parameters

For days in June it is checked whether the parameters that were found for the week in May are still valid 3-4 weeks later. The comparison of model performance on different data sets is not trivial, as the prediction error is dependent on the weather type. The residuals are proportional to the size of fluctuations in solar radiation intensity which occur in partially cloudy conditions. In order to make a fair comparison, it is therefore decided to only compare plots of (non-standardized) residuals on days with comparable radiation fluctuations.

Such a plot is shown in Figure 13.3. The middle panel shows that both data sets contain radiation drops as large as  $600-700 \text{ W/m}^2$ . The residuals are of similar size, staying in the interval  $(-4, 4)$  at all times. This result shows that the model is still valid with the same estimated parameters almost a month after the parameters were estimated.

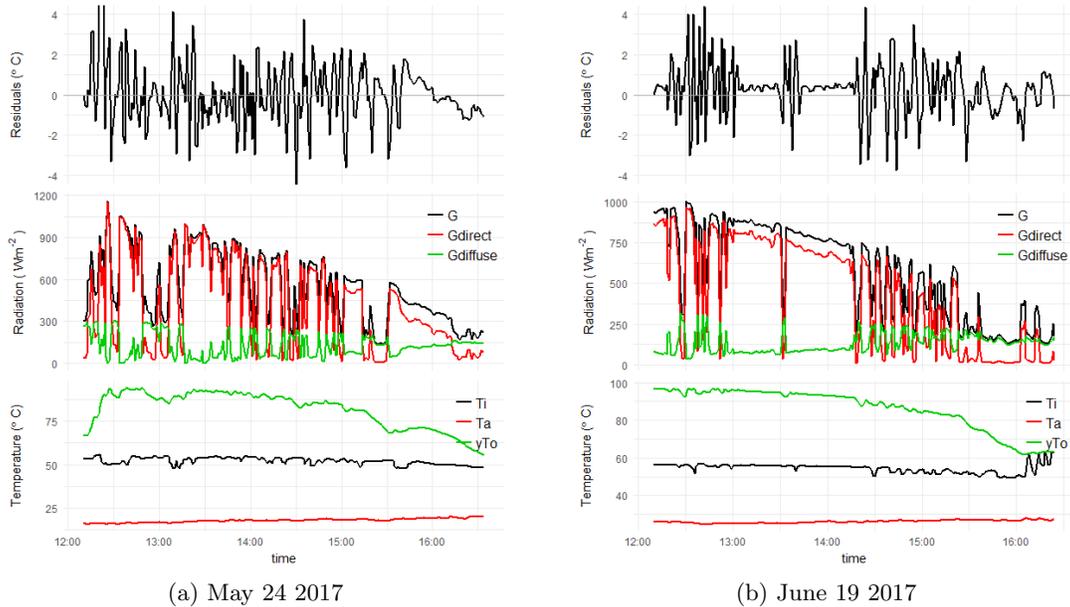


Figure 13.3: Prediction (1 ahead) residuals using final model with parameter estimates obtained in May

The model with parameters as fitted in May is used for 1-ahead prediction on all other data sets. The quality of the prediction is assessed using the NMBE, RMSE and RRSE on the residuals, which

are defined in the following.

$$\begin{aligned}
 NMBE &= \frac{\sum_{k=1}^n x_k - \hat{x}_k}{\sum_{k=1}^n x_k} \\
 RMSE &= \sqrt{\frac{1}{n} \sum_{k=1}^n (x_k - \hat{x}_k)^2} \\
 RRSE &= \sqrt{\frac{\sum_{k=1}^n (x_k - \hat{x}_k)^2}{\sum_{k=1}^n (x_k - \bar{x})^2}}
 \end{aligned} \tag{13.1}$$

where  $x_k$  is the measurement at time  $t_k$ ,  $\hat{x}_k$  the prediction at that time, and  $\bar{x}$  is the average of all measured values in the sample of size  $n$ . The latter measure weights the variance of the residuals with the variance in the measurements.

These measures are computed for each day separately and then combined into an average for each data set. For fairness of the comparison of residual sizes, the two completely clear sunny days from the May data set are excluded from the comparison, as the other data sets do not contain such clear days which come with small residuals. Table 13.4 summarizes the different statistics. The values remain fairly constant from May to June (are even slightly reduced). The RMSE and NMBE for September are considerably higher, especially the measure of bias which shows a 10-fold increase. This indicates that the model should probably be re-estimated for September. All three statistics show a clear increase between October and November.

	05/2017	06/2017	09/2017	10/2017	11/2017
NMBE	-2.735e-04	2.370e-04	-2.484e-03	-2.620e-03	-8.973e-03
RMSE	0.483	0.436	0.655	0.645	0.851
RRSE	0.056	0.046	0.056	0.051	0.091

Table 13.4: Different measures of residual size using the May-fit model on several data sets

### 13.3.2 Re-estimation of parameters

This section contains an investigation of the extent to which the parameter estimates change when the model is fit using data from other times of year. In addition, we will compare the prediction residuals of the May-parameter fit to those of fits in other months. Each data set was prepared for parameter estimation in the way described in Section 8.3.

It can be expected that the model with the parameter fit from May does not predict well in months further than June. The reason for this is mainly the spline method, which makes the shading fraction a function of Time of Day. As the path of the sun through the sky changes over the year, so does the relation between Time of Day and the fraction of shaded panels. Therefore, the parameters will have to be re-estimated about once every 1-2 months in order to have optimal model performance.

Firstly, Table 13.5 shows the parameter estimates for all data sets considered: May (which was used for model development), June, August, September, October, and November. The parameter estimates are mostly in their expected ranges, with the exception of  $b_0$  for the August fit and  $K_d$  for the September and October fit. In addition, the observation error variance varies significantly between the models.

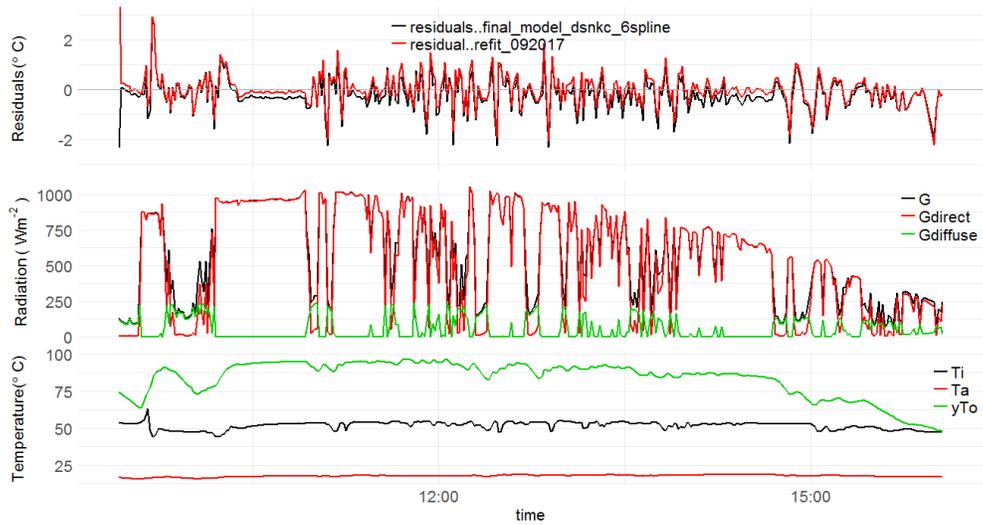
For the fit in September, only one day in the week of data obtained was suitable for inclusion in the estimation. The parameters were thus exclusively determined to match this day, which means that

	May	June	Aug	Sept	Oct	Nov
To0	76.75	71.06	78.90	70.19	80.34	78.49
b0	0.27	0.21	0.0015	0.20	0.27	0.22
FU0	2.24	1.87	2.22	3.15	1.57	2.49
FU1	0.01	0.02	0.0118	0.03	0.0348	0.01
Kd	0.58	0.60	0.61	0.94	0.66	0.38
mC	6533	7288	6721	8065	7193	7466
sigma_o	-3.22	-3.00	-2.70	-2.61	-2.62	-2.73
sigma_obs	-28.89	-10.53	-11.24	-6.59	-24.90	-8.63

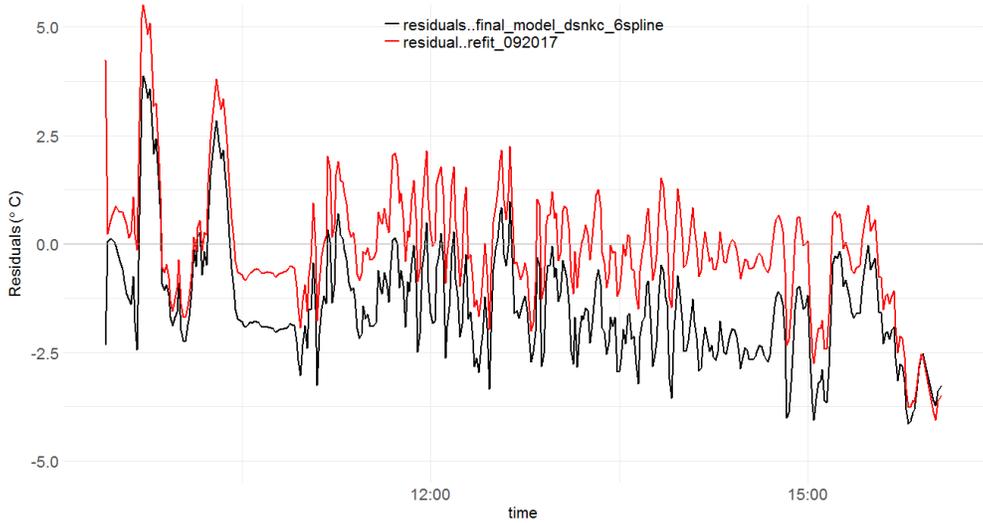
Table 13.5: Parameter fits for the final model on different data sets

the resulting fit is optimal for this day for the given model. We compare the September-fit with the performance of the May-fit on this September day in Figure 13.4. The upper panel shows that for 1 step ahead prediction, the different parameter sets lead to quite similar residuals: the size of the May fit residuals is never more than 1°C larger. For 10-ahead prediction residuals on this same day, as shown in Figure 13.4b, the difference enlarges to around 2°C. Still, the May parameter fit does quite well on this intermittent day, with absolute residual size for 10-ahead prediction not exceeding 4°C.

Figure 13.5 shows the same plots for November, where the residual measures showed a large increase in bias and variance. Both the 1-ahead residuals show a structural underestimation of the outlet temperature when using the May parameters. This bias is enlarged in the 10-ahead prediction, where residual sizes are mostly in the interval  $(-8, -4)$ . This image does however show that the model structure is still suitable for the situation in November, as the prediction using newly fitted parameters (red curve in figure 13.5) has 1-ahead residuals mostly in the interval  $(-1, 1)$  and 10-ahead in the interval  $(-2, 2)$ .

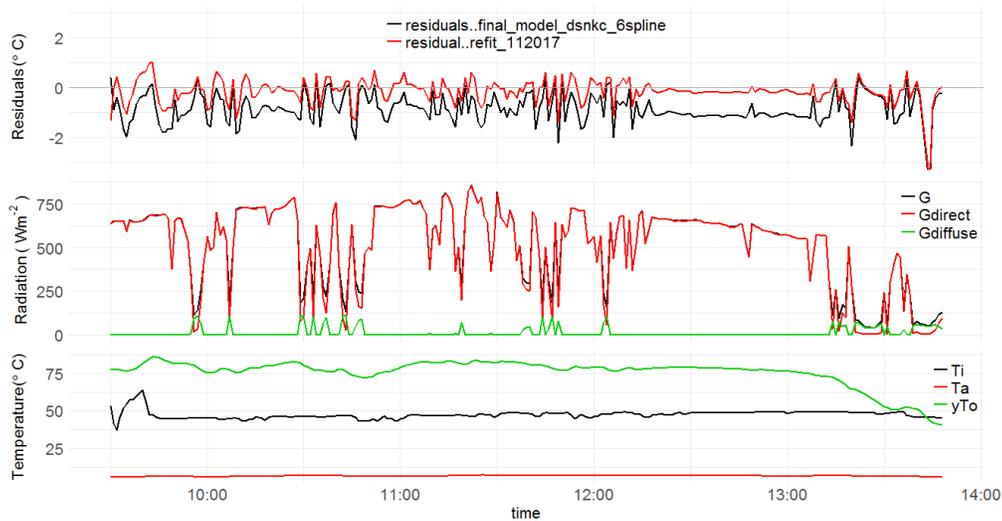


(a) 1 ahead prediction

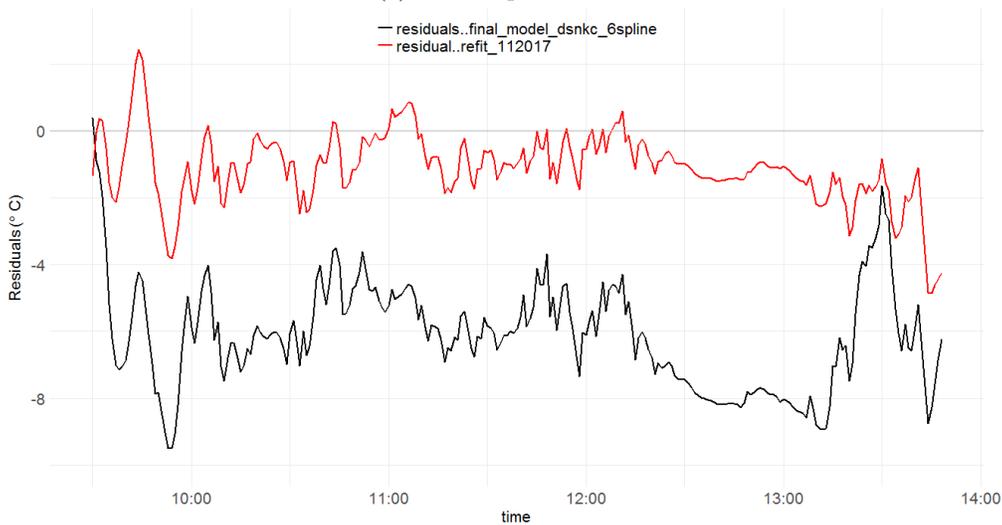


(b) 10 ahead prediction

Figure 13.4: Residuals using May parameters (black) versus September parameters (red), for September 5<sup>th</sup>



(a) 1 ahead prediction



(b) 10 ahead prediction

Figure 13.5: Residuals using May parameters (black) versus November parameters (red), for November 11<sup>th</sup>

## Chapter 14

# Conclusion, Discussion, & Recommendations

This Chapter will discuss the results and insights from the previous chapters, as well as weaknesses and possible improvements of the developed models and the research method. Finally, Section 14.3.2 provides directions for the practical application of the developed model and makes suggestions for future research.

### 14.1 Conclusion

This thesis aimed to develop continuous-discrete stochastic state space models of solar heat plant dynamics, that can be applied for very short or short term prediction (or both) needed for control and heat output prediction of a SDHP. This model type is of a grey-box nature, combining physical knowledge of a system with information that can be extracted from data using statistical methods. Retaining a physically meaningful model formulation was beneficial for model development, as analysis of residuals and outputs provided information on suitable model extensions. The used data set contains measurements from a solar district heating plant in the municipality of Solrød, dating from May 2017. A subset of the data was selected based on operation mode, solar incidence angle, and flow rate.

The research in this thesis has resulted in one such model suitable for both purposes. The model was developed using a forward selection method, starting out with a simple model that was recursively extended. The employed extensions were inspired by the prediction performance of the base model, which showed elements of the model that could be improved on. The Likelihood Ratio Test was applied to identify the most significant model extension in each step, as an objective model selection criterion. The performance of the selected model was successively judged from simulation and predictive properties. In several cases this led to the selection of a different extension than the best LRT extension. The extension was continued until no significant extensions could be found. The then obtained model was modified in the way it deals with inlet temperature measurements, and the way the splines (for the modelling of shading) were constructed.

## The final model

The final model has the following system and observation equation

$$\begin{aligned}
 dT_o &= \frac{1}{(mC)_e} \left( F'U_0(T_a - T_f) + F'U_1(T_f - T_a)^2 + c_f Q_f(T_i - T_o) \right. \\
 &\quad \left. + F'(\tau\alpha)_{en} (K_b(\theta)G_b + K_d G_d) \right) dt + \sigma_o d\omega_o \\
 yT_{o,t_k} &= T_o(t_k) + \epsilon_{obs,t_k}
 \end{aligned} \tag{14.1}$$

where  $\mathbb{V}[\epsilon_{obs}] = \sigma_{obs}$ . The model distinguishes between  $T_o$ , the ‘true’ outlet temperature, and  $yT_o$ , the measured outlet temperature. This allows for separate modelling of process noise and measurement noise. The parameter  $F'(\tau\alpha)_{en}$  is set to 1 for structural identifiability, which is necessary due to the 6-splined direct radiation term

$$G_{b,eff}(t) = \sum_{i=1}^6 s_i \cdot G_{b,i}(t) \tag{14.2}$$

which introduces six additional parameters to the model, which are directly multiplied by  $F'(\tau\alpha)_{en}$ . The spline functions are cubic base splines based on time of day, with boundary knots at  $\cos \theta = 0.2$  and inner knots equally spaced in time between the boundary knots. The inlet temperature is delayed in the introduced adaptive manner, so that the temperature of out-flowing water is approximately related to the temperature of the same water ‘drops’ when flowing in.

Due to the combination of physical knowledge and information from data, the model parameters still have a physical interpretation and therefore provide information about the system. A low parameter estimate for the efficiency may for example indicate that the panels are dusty, whereas an exceptionally high efficiency may indicate dusty sensors.

The model includes several aspects that have not been found in solar heat plant models from the scientific literature, the most notable being the non-parametric modelling of shading and the split of total radiation into diffuse and direct components. This new way of parameterizing the direct radiation term lead to a different understanding of the efficiency parameter: it can now be plotted as a function of time, and this function can be asymmetric (in contrast to the  $K_b(\theta)$  function that is used in many other models). These plots give insight into efficiency changes over the day, as seen in for example Figure 12.15.

We have studied the model performance for n-ahead prediction assuming perfect input prediction. On both  $D_s$  and  $D_i$ , the average absolute residual size is below 0.3°C for 1-ahead prediction, and the maximum absolute residual does not exceed 0.6°C and 1.7°C for the sunny and cloudy day respectively. When using the model for flow control, predictions up to 1 hour may be used. For this time span, the error does not exceed 2°C on the largest part of sunny day  $D_s$ , whereas the maximum absolute error on  $D_i$  equals 2.85°C. However, on the last 30 minutes of  $D_s$  the residuals blow up in many-ahead prediction, leading to errors as large as 7.8°C in 60 minute-ahead prediction. Outside this time period, it is concluded that the developed model is suitable for the intended purposes.

The model was crossvalidated on data sets from other months, both with and without re-estimation of the parameters. The model can still predict well almost a month after the parameters were estimated, as tested on a week in June. Re-estimation was successful for all months the model was fit on, and the resulting fitted models performed well in prediction and simulation for the corresponding data sets.

A comparison of the developed model to the ISO-standard MLR model and the model currently used in Solrød reveals clear improvements of simulated outlet temperature. Particularly when the system is not in steady state, the new model outperformed the others by far.

## Lessons from the model sequence

The iterative model extension procedure provided us not only with a working model, but also with several insights worth remembering for future modelling of solar heat plants. First of all, it was shown that the direct inclusion of input variables in the observation equation is *not* a good idea. In model *Tf*, this led to unreasonably strong dependence of the particular input variable on the output variable predictions.

Interestingly, the first most significant extension was the non-parametric model of shading effects, using the splined radiation. This new approach thus has the potential to improve existing models considerably. In the next step, the split of direct and diffuse radiation improved the model most. This is an important result, because many solar heat plants do not invest in the costly diffuse radiation measurement. From this study we conclude that it may not be necessary to do so, as the applied post-processing of the total radiation into its diffuse and direct components lead to satisfactory results.

The Solrød plant currently includes neither of these effects in the equation used for control. Even without adopting the grey-box model developed here, the plant could improve its control considerably by including the shading and split radiation in their model. Further, the benefits of including the current outlet temperature in the prediction of the outlet temperature of the following time step have become clear, as even the simplest model applied here was much more accurate than the current Solrød model and the MLR based model while using the same input variables.

We further found that certain extensions were unable to improve model performance for the intended purposes of simulation and control. This was the case for the additional compartments in the flow direction of the system, and the added metal component. Both extensions add a *hidden* state to the model. In all extensions steps, the loglikelihood of the multi-compartment model was by far the highest. By plotting the filtered states and hidden states however, it was shown that these high loglikelihoods were a result of overfitting effects: the hidden states showed unreasonable fluctuations that matched the data well, leading to good 1-ahead predictions. Especially in periods with high inlet temperature fluctuations, this reduced the residuals significantly. In  $\geq 2$ -ahead prediction and simulations however, the multi-compartment models (in the flow direction) performed worse than even the simplest model.

The metal component model extensions showed the same overfitting behaviour earlier in the model sequence, but resulted in good simulation and prediction when introduced into more complex models later on in the sequence. It was however deemed inferior to other extensions due to its complexity: the metal component adds a state and several parameters to the model. This made the fitting procedure slower and sometimes lead to failed parameter estimation, while the model did not perform visibly better than another extension (the addition of first order heat loss).

## 14.2 Discussion

This Section will discuss aspects of the research that could be improved, and difficulties that were encountered.

## Discussion of final model

The final model residuals still show some patterns. In some periods, the residuals are structurally negative/positive. The clearest example of this is the late afternoon on clear sunny days. The residual pattern here is likely to be a result of the splines, which require fitting a single efficiency parameter to each time interval. The parameter  $s_6$  for the splines radiation covering the last period of the day is too low for the start of the interval, and too high for the end of the interval. This could perhaps be fixed by placing the knots for the splines differently, so that the splined radiation in the late afternoon covers shorter time periods.

Another way to improve the non-parametric modelling of shading could be to make the splines based on the solar azimuth or zenith angle, or even both. These angles together determine the amount of shading on the panels. A further advantage of doing so, is that the splines may be more generally applied for the whole year. For the current method, the time at which  $\cos(\theta) = 0.2$  needed to be found for each new data set, and afterwards the boundary knots for the splines needed to be changed manually.

Large residuals when predicting with the final model are also caused by fluctuations in the inlet temperature. The delay method for the inlet temperature as introduced in the last step of model development decreased this effect, but it is still present.

It is further expected that the modelling of incidence angle effects can be improved. In most selected models, a very high value of  $b_0$  fit the data best. This is problematic in the IAM function used, as large values of  $b_0$  lead to negative IAM values for small incidence angles. It was chosen to deal with this by discarding data with  $\cos(\theta) < 0.2$ . The plant however also operates outside these ranges of the incidence angle, and needs to be modelled for those period too. It would therefore be beneficial to be able to relax the incidence angle selection criterion.

The ACF and CCF plots of 1-ahead prediction residuals have improved largely over the model sequence. However, the lag-1 autocorrelation is still significant in the final model. Furthermore, the CCF plots of residuals versus both total radiation and inlet temperature reveal significant lag-0 correlations. This means that the way the model deals with these input variables may still be improved.

## Data quality and selection

The obtained data sets used for this research describe the behaviour of the plant *under the current control scheme*. This means that the variables relate to one another in specific ways, not only as determined by the laws of physics. The existing control puts constraints on the variability of the data. Especially the flow and total radiation are extremely correlated over all days, with a correlation of 0.97 for the 1-minute resolution of the data. After splitting into direct and diffuse components, the correlation between  $G_d$  and  $Q_f$  remains at the high value of 0.95. In the case of the flow and the radiation however, the model equations only include the variable  $Q_f$  multiplied by the temperature difference between inlet and outlet temperature. The correlation of this entire term to the direct radiation is also high, at  $-0.97$ .

The question is whether the parameters related to the highly correlated variables (in our case  $(mC)_e$ , the spline factors  $s_1$  to  $s_6$ , and  $b_0$ ) remain practically identifiable. However, the parameter related to  $Q_f(T_i - T_o)$  is  $(mC)_e$ , which also multiplies all other parameters. This may be the reason that the high correlation in the data seems not to affect parameter estimates much, as judged by the correlation Table 12.13. However, the model should be tested on data where no control scheme is

yet implemented. The effect of the implemented control on the parameter estimates can then be determined.

The first obtained data set from May was used for model development. This set was selected by the commissioner for a reason: it covers a week with unusually good weather during which the plant was operating for many hours. Two out of ten were completely clear sunny days. It could therefore be questioned whether the data set was representative for normal plant dynamics. If not, the question is whether this has influenced the choices in model development. From the crossvalidation on measurements from other months, the model however seems to work as good as on the initial data set.

As said, data points with  $\cos(\theta) < 0.2$  were filtered out before fitting the parameters. Due to the non-parametric shade modelling, the fitted model can only be used on times of day that are actually present in the data set used for fitting. To extend the model to lower incidence angles, an assumption on the effect of direct radiation needs to be made.

Another disadvantage of the applied data selection method, is that it is too selective for months outside of summer, where larger portions of the data need to be removed due to the incidence angle criterion. Some other months (with more periods of low radiation intensity) additionally show more periods in pre-heating phase, so that an even smaller data set remains to be used for fitting.

## Model fitting / Parameter estimation

The developed model is intended for prediction on time scales from seconds to hours. The parameters are selected to maximize the loglikelihoods obtained from 1 minute-ahead prediction. It could therefore be questioned whether the resulting parameters are optimal for the purposes of the model. To improve parameter fits for days-ahead prediction, it may be useful to develop the model based on data of a lower sampling rate.

In relation to that, a lower sampling rate may also lead to better working multi-compartment models. These were now prone to overfitting, but when the outlet temperature shows larger deviations from one data point to the next, the parameter estimation using these models may lead to better prediction.

The CTSM-R software could in almost all cases successfully estimate the parameters. However, in the multi-state models, the parameter estimation using more than 4 splines failed or lead to fits with bad properties (high values  $dF/dPar$  and  $dPen/dPar$ ). In some other models, the parameter estimation failed using certain initial settings in CTSM, which resulted in an error related to unsolvable ODEs. Resetting the initial guess for parameter  $(mC)_e$  to larger value helped in such cases. This undesirable property of sensitivity to initial conditions was only seen in multi-compartment models.

## 14.3 Recommendations

We will conclude the thesis with recommendations for the future. Section 14.3.1 describes several steps that need to be taken before the developed model is ready for use. Next steps for the implementation of the developed model are described; thereafter, possible improvements and extensions of the model are suggested.

### 14.3.1 Next steps for model implementation

First of all, designed model identification control sequences in the form of pseudorandom binary sequences (PRBS) should be run at the plant. In such sequences, the flow and inlet temperature should be controlled independently of all other inputs. The model should be verified on such sequences, because the resulting measurements will show a wider range of dynamics than the data of the system under the implemented control.

Next, the performance of the model using (imperfect) input predictions from weather forecasts should be tested. In order to use numerical weather predictions (NWP) as an input for the model, the predictions need to be transformed to match measurements at the plant. For example, the NWP of solar radiation intensity should be adapted to account for shadowing and tilt effects. This has for example been done by Bacher et al. in [57] using a local quantile regression model based on a third order polynomial.

If the model performance well on the PRBS data and using NWP, a nonlinear control scheme can be designed. The scheme should build on previous work, including the model predictive control proposed by Andrade et al. [7].

For application of the model in practice, requirements for the data set used for parameter fitting should be determined, both concerning its size as well as the different weather types it contains. It would be useful to know how many days (or hours) of data are required to obtain reliable parameter estimates, which make the model suitable for all different operation conditions.

In the current research, it was chosen to exclude the pre-heating phase from the data set for fitting. Strict requirements were set for the incidence angle of the data points for fitting as well. It should be investigated if these restrictions can be relaxed, so that a larger data set is left for fitting. This is especially relevant for months in which the number of operating hours is restricted, and the incidence angle is larger than  $80^\circ$  during a longer period.

It is further important to know how frequently the model parameters should be re-estimated over the year. As far as seen in this thesis, the model can perform well for at least a month. It would be good to develop an automatic detection method, that will notify the operator once the model parameters need to be re-estimated. The possibility of using a prior when re-estimating the parameters can be explored too.

### 14.3.2 Suggestions for future research

The developed model could be improved and extended in different ways. This section will outline some interesting directions for for future research.

Ways to improve the non-parametric efficiency modelling should be investigated. Most visibly on clear sunny days, the current method leads to a large prediction error in the afternoon. The residual here was structurally negative, which means that the model over-estimates the outlet temperature. The current splines are not flexible enough to fit the steep drop in efficiency that is apparently required for the afternoon. The flexibility of the splines could be increased by adding the position of the last inner knot as a model parameter. It should then automatically be positioned in a position that allows for closer modelling of the efficiency drop. Alternatively, the final inner knot could manually be positioned closer to the time of sunset.

As mentioned in the Discussion, the splines could be made in a general way to fit the whole year, using two-dimensional splines on the solar azimuth and zenith angle.

As the dynamics of the plant as well as the residual sizes differ considerably between clear and cloudy conditions, a model with different regimes for each of these cases could improve performance. Such regimes would be relevant for control schemes as well, because of the difference in residual sizes. The regimes could be versions of the same model equation with different parameter fits, or even different models. A fish-eye camera could be used for determining the regime the model should be in. These cameras are cheap relative to for example diffuse radiation sensors.

Such fish-eye cameras have also been used for sub-hourly local solar irradiance forecasting, for example using transfer functions [58] and artificial neural networks [59]. Such forecasts are also interesting to combine with the model developed here, in order to predict sudden radiation drops and corresponding outlet temperature drops, and in this way smoothen the control of the plant.

In this thesis, we have not devoted much attention to modelling the preheating phase of the plant. It may be useful to develop a separate regime for the preheating phase as well.

In addition, it could be investigated whether multi-state models in flow directions could improve the predictions. Instead of using hidden states (as done in this thesis), an additional temperature measurement in the middle of a collector row could help to make these multi-state models work.

The model could further be made more detailed in space using more of the individual temperature measurements that are available from the plant. The four rows for which inlet and outlet temperature are measured could be modelled separately, leading to the prediction of four different output temperatures. This may be important for the prevention of boiling in the collectors, as some rows are shaded less than the average in the plant, and therefore may be at higher temperatures than described by the current model.

# Bibliography

1. IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2013).
2. UNEP. *Bridging the Emissions Gap* (United Nations Environment Programme (UNEP), 2011).
3. IEA. *World Energy Outlook 2017* (Organisation for Economic Co-operation and Development, OECD, 2017).
4. Lund, H., Möller, B., Mathiesen, B. V. & Dyrelund, A. The role of district heating in future renewable energy systems. *Energy* **35**, 1381–1390 (2010).
5. Connolly, D. *et al.* Heat Roadmap Europe: Combining district heating with heat savings to decarbonise the EU energy system. *Energy Policy* **65**, 475–489 (2014).
6. Werner, S. International review of district heating and cooling. *Energy* (2017).
7. De Andrade, G. A., Álvarez, J. D., Pagano, D. J. & Berenguel, M. Nonlinear controllers for solar thermal plants: A comparative study. *Control Engineering Practice* **43**, 12–20. ISSN: 09670661 (2015).
8. Camacho, E., Rubio, F., Berenguel, M. & Valenzuela, L. A survey on control schemes for distributed solar collector fields. Part I: Modeling and basic control approaches. *Solar Energy* **81**, 1240–1251. ISSN: 0038092X (Oct. 2007).
9. Pasamontes, M., Álvarez, J., Guzmán, J., Lemos, J. & Berenguel, M. A switching control strategy applied to a solar collector field. *Control Engineering Practice* **19**, 135–145 (2011).
10. Pickhardt, R. & Silva, R. N. Application of a nonlinear predictive controller to a solar power plant. *IEEE*, 6–10 (1998).
11. Bacher, P. & Madsen, H. Identifying suitable models for the heat dynamics of buildings. *Energy and Buildings* **43**, 1511–1522. ISSN: 03787788 (2011).
12. Andersen, P. D., Jiménez, M. J., Madsen, H. & Rode, C. *Characterization of heat dynamics of an arctic low-energy house with floor heating* in *Building Simulation* **7** (2014), 595–614.
13. Friling, N., Jiménez, M. J., Bloem, H. & Madsen, H. Modelling the heat dynamics of building integrated and ventilated photovoltaic modules. *Energy and Buildings* **41**, 1051–1057 (2009).
14. Bacher, P., Madsen, H. & Perers, B. Models of the heat dynamics of solar collectors for performance testing. *Proceedings of ISES Solar World Conference 2011* (2011).
15. R Core Team. *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing (Vienna, Austria, 2017). <https://www.R-project.org/>.
16. Juhl, R. *ctsmr: CTSM for R* R package version 0.6.10 (2016).
17. Kristensen, N. R., Madsen, H. & Jørgensen, S. B. Parameter estimation in stochastic grey-box models. *Automatica* **40**, 225–237. ISSN: 00051098 (2004).

18. Kloeden, P. E., Platen, E. & Schurz, H. *Numerical solution of SDE through computer experiments* (Springer, 1994).
19. Zastawniak, T. & Brzezniak, Z. *Basic stochastic processes* (Springer, 1998).
20. Tian, Y. & Zhao, C. Y. A review of solar collectors and thermal energy storage in solar thermal applications. *Applied Energy* **104**, 538–553. ISSN: 03062619 (2013).
21. Lemos, J., Neves-Silva, R. & Igreja, J. *Adaptive Control of Solar Energy Collector Systems* ISBN: 9783319068527. doi:10.1007/978-3-319-06853-4 (Springer, 2014).
22. Perez-Mora, N. *et al.* Solar district heating and cooling: A review. *International Journal of Energy Research* **42**, 1419–1441. ISSN: 1099114X (2018).
23. Dalenback, J. *Ranking List of European Large Scale Solar Heating Plants* <http://solar-district-heating.eu/ServicesTools/Plantdatabase.aspx> (2018).
24. Furbo, S., Perers, B. & Bava, F. *Thermal performance of solar district heating plants in Denmark in EuroSun 2014* (2014).
25. Bava, F. *Modeling of solar collector fields for solar heating plants in district heating systems* PhD thesis (DTU, 2017). ISBN: 9788778774590.
26. Yaïci, W., Longo, M., Entchev, E. & Foiadelli, F. Simulation study on the effect of reduced inputs of artificial neural networks on the predictive performance of the solar energy system. *Sustainability (Switzerland)* **9**. ISSN: 20711050. doi:10.3390/su9081382 (2017).
27. Kishor, N., Das, M. K., Narain, A. & Prakash Ranjan, V. Fuzzy model representation of thermosyphon solar water heating system. *Solar Energy* **84**, 948–955. ISSN: 0038092X (2010).
28. Medinelli Sanino, L. A. & Rojas Reischel, R. A. Modeling and identification of solar energy water heating system incorporating nonlinearities. *Solar Energy* **81**, 570–580. ISSN: 0038092X (2007).
29. European Committee for Standardization (CEN) Technical Committee. *ISO 9806:2017: Solar energy — Solar thermal collectors — Test methods* tech. rep. (International Organization for Standardization, Geneva, Switzerland., 2017).
30. Kicsiny, R. Multiple linear regression based model for solar collectors. *Solar Energy* **110**, 496–506. ISSN: 0038092X (2014).
31. Fischer, S. *et al.* Collector test method under quasi-dynamic conditions according to the European Standard EN 12975-2. *Solar Energy*. ISSN: 0038092X. doi:10.1016/j.solener.2003.07.021 (2004).
32. Perers, B. An improved dynamic solar collector test method for determination of non-linear optical and thermal characteristics with multiple regression. *Solar Energy* **59**, 163–178. ISSN: 0038092X (1997).
33. Badescu, V. Optimal control of flow in solar collector systems with fully mixed water storage tanks. *Energy Conversion and Management* **49**, 169–184. ISSN: 01968904 (2008).
34. Tian, Z., Perers, B., Furbo, S. & Fan, J. Analysis and validation of a quasi-dynamic model for a solar collector field with flat plate collectors and parabolic trough collectors in series for district heating. *Energy* **142**, 130–138. ISSN: 03605442 (2018).
35. Quintana, H. J. & Kummert, M. Optimized control strategies for solar district heating systems. *Journal of Building Performance Simulation* **8**, 79–96 (2015).
36. Kicsiny, R. Improved multiple linear regression based models for solar collectors. *Renewable Energy* **91**, 224–232. ISSN: 18790682 (2016).
37. Buzás, J. & Kicsiny, R. Transfer functions of solar collectors for dynamical analysis and control design. *Renewable Energy* **68**, 146–155. ISSN: 09601481 (2014).
38. Pickhardt, R. Adaptive control of a solar power plant using a multi-model. *IEE* (2000).

39. Guzman, J. L. Hybrid modeling of a solar-thermal heating facility. **97**, 577–590 (2013).
40. Madsen, H. *Time Series Analysis* 1–345. ISBN: 1584883170 (2007).
41. Maybeck, P. *Stochastic Models: Estimation and Control* ISBN: 0125111606. doi:10.1016/S0076-5392(08)62171-2 (Elsevier, 1979).
42. Jazwinski, A. *Stochastic Processes and Filtering Theory* (Academic Press, New York, USA, 1970).
43. Nise, N. S. *Control System Engineering* (John Wiley & Sons, 2007).
44. Nielsen, J. N. & Madsen, H. Applying the EKF to stochastic differential equations with level effects. *Automatica* **37**, 107–112 (2001).
45. Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. *Journal of Fluids Engineering* **82**, 35–45. ISSN: 00219223 (1960).
46. Kristensen, N. R. & Madsen, H. Continuous Time Stochastic Modelling - CTSM Mathematics Guide, 1–32 (2003).
47. Juhl, R., Møller, J. K. & Madsen, H. CTSMR – Continuous Time Stochastic Modeling in R. *The R User Conference 2014*, 1–11 (2014).
48. Huber, P. J. *Robust statistics* 1248–1251 (Springer, 2011).
49. Thornton, C. L. & Bierman, G. J. UDU/T/covariance factorization for Kalman filtering (1980).
50. Van Loan, C. Computing integrals involving the matrix exponential. *IEEE transactions on automatic control* **23**, 395–404 (1978).
51. Sidje, R. B. Expokit: a software package for computing matrix exponentials. *ACM Transactions on Mathematical Software (TOMS)* **24**, 130–156 (1998).
52. Dennis Jr, J. E. & Schnabel, R. B. *Numerical methods for unconstrained optimization and nonlinear equations* (Siam, 1996).
53. Kristensen, N. R., Madsen, H. & Jørgensen, S. B. A method for systematic improvement of stochastic grey-box models. *Computers and Chemical Engineering* **28**, 1431–1449. ISSN: 00981354 (2004).
54. Møller, J. K., Carstensen, J. & Madsen, H. Structural Identification and Validation in Stochastic Differential Equation based Models-With application to a Marine Ecosystem NP-model. *Journal of the Royal Statistical ...* (2011).
55. Bacher, P., Madsen, H. & Perers, B. Models of the heat dynamics of solar collectors for performance testing. *Proceedings of ISES Solar World Conference 2011* (2011).
56. Ruiz-Arias, J., Alsamamra, H., Tovar-Pescador, J. & Pozo-Vázquez, D. Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions. *Energy Conversion and Management* **51**, 881–893 (2010).
57. Bacher, P., Madsen, H., Perers, B. & Nielsen, H. A. A non-parametric method for correction of global radiation observations. *Solar Energy* **88**, 13–22 (2013).
58. Gauchet, C., Blanc, P., Espinar, B., Charbonnier, B. & Demengel, D. *Surface solar irradiance estimation with low-cost fish-eye camera in Workshop on "Remote Sensing Measurements for Renewable Energy"* (2012).
59. Chu, Y. *et al.* A smart image-based cloud detection system for intrahour solar irradiance forecasts. *Journal of Atmospheric and Oceanic Technology* **31**, 1995–2007 (2014).
60. Iqbal, M. *An introduction to solar radiation* (Elsevier, 1983).

# Appendices

# Appendix A

## Derivation of the incidence angle

This section describes the computation of the incidence angle  $\theta$  of the solar rays on the panels. It is determined from the incidence angle of the sun ( $\theta_s$ ) and the orientation of the flat plates have to be known. The latter is fixed, whereas the former depends on time, date, and the latitude. In the computation we shall use the concepts of the azimuth and zenith angle. The azimuth is the angle between a vector's projection in the horizontal plane, and the vector pointing South. The zenith is the angle in between the vector and the normal vector of the horizontal plane. The orientation of the solar collector is given by the azimuth ( $\gamma$ ) and zenith angle ( $\beta$ ) of its normal vector.

The computation of the solar incidence angle will be derived here, using the quantities  $d$ ,  $h$ , and  $l$  as depicted in Figure A.1. We can relate  $d$ , and  $h$  to the day of year  $n$  and the Local Clock Time (LCT);  $l$  is the latitude.

The declination angle  $d$  captures how many degrees above the equator the sun's rays hit the Earth perpendicular to its surface. It is approximately a sine function, given by

$$d = 23.45 \sin\left(\frac{284 + n}{365}\right) \quad (\text{A.1})$$

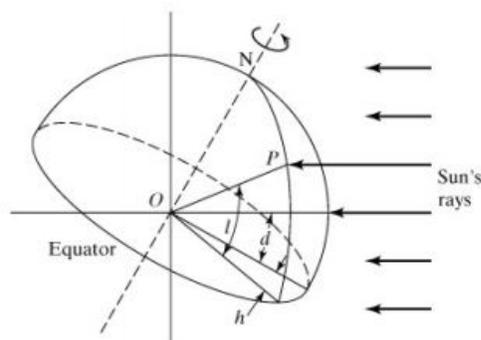


Figure A.1: Defining  $d$ ,  $h$ , and  $l$

The hour angle  $h$  expresses the angle that the earth has rotated since solar noon, the time of day

that the sun was at its highest point. For the computation of this angle, it is important to take into account the difference between a solar day and a calendar day. Whereas the latter always lasts 24 hours, a solar day is defined as the time that it takes for a stationary observer on Earth to go through one solar cycle. the length of a solar day varies slightly with the obliquity and eccentricity of the Earth's orbit around the sun, and some other factors [60].

Each hour, the angle increases by  $15^\circ$ . However, clock time (also Local Civil Time, LCT) and the so-called Local Solar Time do not match completely. Whereas a day on the clock is always 24 hours, The Equation of Time,  $E$ , is used to correct for these differences between LST and LCT.

$$\begin{aligned}
h &= 15 \frac{\pi}{180} (LST - 12) \\
LST &= LCT + \frac{L_{std} - L_{loc}}{15} + E - DT \\
E &= 0.165 \sin(2B) - 0.126 \cos(B) - 0.025 \sin(B) \\
B &= \frac{360}{364} (n - 81)
\end{aligned} \tag{A.2}$$

In these equations  $L_{loc}$  is the longitude of actual location,  $L_{std}$  is the standard meridian for the local time zone (degrees west),  $n$  is the day of the year.

The computation of the sun's direct beam zenith ( $\theta_H$ ) and azimuth ( $\phi$ ) angle are then as follows:

$$\begin{aligned}
\cos(\theta_H) &= \cos(l) \cos(h) \cos(d) + \sin(l) \sin(d) \\
\cos(\phi) &= \frac{1}{0.5\pi - \cos(\theta_H)} (\cos(d) \sin(l) \cos(h) - \sin(d) \cos(l))
\end{aligned} \tag{A.3}$$

Combining these with zenith and azimuth angle of the panels, the incidence angle of solar rays on the panel can be computed.