STOCHASTIC MODELLING OF ENERGY SYSTEMS

Klaus Kaae Andersen

Department of Mathematical Modelling
The Technical University of Denmark
DK-2800 Kgs. Lyngby
Denmark

IMM-PHD-2001-79

IMM

ISSN ****-****
ISBN **-*****-**-*

Bound by Hans Meyer, Technical University of Denmark

The work presented in this thesis is also documented in the following papers:

Paper A: Andersen, K.K. and Poulsen, H. (1999). Building integrated heating systems. *Building Simulation '99, Kyoto, Japan*, **1**, 105–112.

Paper B: Andersen, K.K., Palsson, O.P., Madsen, H., and Knudsen, L.H. (2000). Experimental design and setup for heat exchanger modelling. *International Journal of Heat Exchangers*, Accepted for publication.

Paper C: Andersen, K.K., Lundby, M., Madsen, H., and Paulsen, O. (1999). Identification of continuous time smooth threshold models of physical systems. *Presented at the Joint Statistical Meetings*, Baltimore, 1999.

Paper D: Andersen, K.K., Hansen, L.H., and Madsen, H. (2000). A model for the heat dynamics of a radiator. Submitted.

Paper E: Knop, O., Andersen, K.K., Madsen, H., Gregersen, N.H., and Paulsen, O. (2000). Modeling of a thermostatic valve with hysteresis effects. Submitted.

Paper F: Andersen, K.K., Madsen, H., and Hansen, L.H. (2000). Modelling the heat dynamics of a building using stochastic differential equations. *Energy and Buildings* **31** 1, 13-24.

Paper G: Andersen, K.K., and Reddy, A. (2000). The error in variable (EIV) regression approach as a means of identifying unbiased physical parameter estimates. Submitted.

Preface

This thesis has been prepared at the Department of Mathematical Modelling (IMM), The Technical University of Denmark (DTU), in fulfilment of the requirements for the degree of Ph.D. in engineering.

The thesis is concerned with grey box modelling of heating system components. The main contribution to this field is a discussion on how statistical techniques and physical insight can be combined in building models. As a result, dynamic models of heating system components are presented.

Lyngby, January 3, 2001.

Klaus Kaae Andersen

Acknowledgements

I wish to express my gratitude to all who have contributed to this research. First of all, I want to thank my academic advisor, Professor Henrik Madsen, IMM, DTU, for his help and guidance during this work. Professor Madsen's enthusiasm and insight into the discipline of mathematical modelling has been a source of inspiration and invaluable for this research.

I also wish to thank Associate Professor Agami Reddy at Drexel University in Philadelphia, USA, with whom I have spent 9 month as a Visiting Researcher. During the stay at Drexel I have benefited greatly from Dr. Reddy's strong insight in the analysis and modelling of heating systems.

During my research I have had the pleasure to work with very competent people from the Danish heating industry. I wish to thank the participants in the joint project IT-Energy, for good advises and interesting discussions. I would like to thank the participants from Grundfos A/S, Danfoss A/S, APV systems A/S, the Danish Technological Institute and the Energy Engineering Department at the Technical University of Denmark. Especially, I would like to thank Mr. Henrik Poulsen, chief of the Energy section at the Danish Technological Institute, for pushing me forward at certain times and for giving me the opportunity to participate in various projects.

I am very grateful to Mr. Mads Lundby and Mr. Ole Knop, whom I have been advising and working with during their master's studies. It has been a pleasure to work with Mads and Ole, and their contributions have been of great value to my research.

I am also grateful to Dr. Lars Henrik Hansen, for guiding me through my master's project and later in advising me in my Ph.D. studies, and to Associate Professor Olafur Palsson at the University of Iceland for good discussions.

At IMM I wish to thank past and present members of the time series analysis group and the statistics group. Especially, I would like to thank Mr. Peter Thyregod, who has become a good friend.

Finally, I would like to express my thanks to my family and friends who have supported me during the past three years. Above all, I would like to thank my fiancée, Zorana, for her strong support during the preparation of this thesis.

Summary

In this thesis dynamic models of typical components in Danish heating systems are considered. Emphasis is made on describing and evaluating mathematical methods for identification of such models, and on presentation of component models for practical applications.

The thesis consists of seven research papers (case studies) together with a summary report. Each case study takes it's starting point in typical heating system components and both, the applied mathematical modelling methods and the application aspects, are considered. The summary report gives an introduction to the scope of application and the applied modelling method and summarizes the research papers.

The foundation of the identification process is the grey box modelling method. The grey box modelling method is characterized by using information from measurements in conjunction with physical knowledge. The combination of statistical methods and physical interpretation is exploited in the modelling procedure, from the design of experiments to parameter estimation and model validation. The presented models are mainly formulated as state space models in continuous time with discrete time observation equations. The state equations are expressed in terms of stochastic differential equations. From a theoretical viewpoint the techniques for experimental design, parameter estimation and model validation are considered. From the practical viewpoint emphasis is put on how this methods can be used to construct models adequate for heating system simulations.

Significant parts of the research work have been done in cooperation with leading companies from the Danish heating industry. The presented models have been developed for the purpose of analyzing typical heating system installations. The focal point of the developed models is that the model structure has to be adequate for practical applications, such as system simulation, fault detection and diagnosis, and design of control strategies. This also reflects on the methods used for identification of the component models.

The main result from this research is the identification of component models, such as e.g. heat exchanger and valve models, adequate for system simulations. Furthermore, the thesis demonstrates and discusses the advantages and disadvantages of using statistical methods in conjunction with physical knowledge in establishing adequate component models of heating systems.

Resumé (in Danish)

Nærværende afhandling omhandler dynamiske modeller for komponenter, der typisk indgår i danske varmesystemer. Der er lagt vægt på at beskrive og vurdere matematiske metoder til identifikation af komponent-modeller samt at diskutere anvendeligheden af sådanne modeller ved løsning af praktiske problemstillinger.

Afhandlingen består af syv artikler samt et sammendrag og en perspektivering heraf. Artiklerne tager hver for sig udgangspunkt i en typisk system komponent og herunder er både de anvendte metoder samt anvendelsesområderne behandlet. Sammendraget omfatter en introduktion til anvendelsesområdet og de anvendte metoder, samt en sammenfatning og kort diskussion af de præsenterede artikler.

Udgangspunktet for modelleringsværktøjerne, der er taget i brug, er metoder relateret til grey box modellering. Grey box modellering er karakteriseret ved, at information fra måledata anvendes og knyttes sammen med kendte fysiske lovmæssigheder, hvilket giver fordele ved identifikation af empiriske modeller, hvor den fysiske indsigt er god, men også vigtig for den videre anvendelse. Kombinationen af statistiske metoder og indsigt i den bagvedliggende fysiske struktur er udnyttet i modelopbygningen, fra planlægning af forsøg, til estimation af ukendte parametre og model validering. De foreslåede modeller er overvejende formulerede som tilstands modeller i kontinuert tid med observationer i diskret tid. Tilstandsligningerne er formuleret som stokastiske differential ligninger. Fra et teoretisk synspunk er teknikker for forsøgsplanlægning, parameter estimation og model validering behandlet. Fra et anvendelses orien-

teret synspunkt er hovedvægten lagt på hvordan disse metoder kan udnyttes til at opstille anvendelige modeller til simulering af varme installationer.

Betydelige dele af forskningsarbejdet er blevet udført i samarbejde med danske virksomheder inden for varmesystemer. Det afspejles klart i de præsenterede modeller, der er udviklet med henblik på analyse af typiske system installationer. Der lægges vægt på, at de udviklede modeller er anvendelige for praktiske applikationer, som system simulering, fejl- detektering og -diagnose samt analyse af regulerings strategier. De praktiske problemstiller afspejler sig endvidere i valget af de anvendte metoder.

Hovedresultatet fra forskningsarbejdet er formulering af komponentmodeller, til brug for simulering af varme installationer. Desuden vises og diskuteres fordele og ulemper ved at anvende grey box modellerings teknikker ved modellering af komponent modeller for varme systemer.

Outline of the report

The thesis consists of seven research papers (case studies) together with a summary report and is organized into two separate parts.

Part I is the summary report entitled "Dynamical models of HVAC&R systems". The summary report gives an introduction to the scope of application and the applied modelling method as well as summarizing the research papers. The summary report is intended as a short introduction to the research documented in the subsequent papers, discussing the practical aspects and the applied mathematical framework. Finally, an overview of the included papers is given followed by a conclusion.

The main part of the research work is presented in **Part II** and is entitled 'Included papers'. The part consists of 7 research papers, each discussing different aspects of the modelling method and treatment of different component models. The 7 papers, which are denoted A-G, are:

A: Building integrated heating systems.

B: Experimental design and setup for heat exchanger modelling.

C: Identification of continuous time smooth threshold models of physical systems.

D: A model for the heat dynamics of a radiator.

E: Modelling of a thermostatic valve with hysteresis effects.

F: Modelling the heat dynamics of a building using stochastic differential equations.

G: The error in variable (EIV) regression approach as a means of identifying unbiased physical parameter estimates.

The research papers have been published either as journal papers, conference proceeding papers, or have been submitted for publication. The papers can be read as a stand-alone or as a part of the project. However, some parts of these papers may be similar in nature due to publication in different journals or proceedings.

Contents

I

Part I

DYNAMIC MODELLING OF HVAC&R SYSTEMS

Chapter 1

Background

This thesis deals with how statistical techniques can be applied to identify dynamic models of energy systems in conjunction with the physical interpretation of the system. More specifically, the thesis is concerned with grey box modelling of Heating Ventilation Air-Conditioning and Refrigerating (HVAC&R) components. The relevance of this research work is to apply and develop methods adequate for model identification, and hereby to establish models adequate for analysis, simulation, and applications of HVAC&R components and systems. In the following, the application aspects of the study will be presented followed by an introduction to the modelling framework.

1.1 IT tools within the field of energy

The work presented in this thesis is closely related to the project IT-Energy, a forum for developing Information Technology (IT) tools for energy systems analysis. IT-Energy was established during the fall 1997 and is a joint project between the Danish Technological Institute, the Technical University of Denmark, and leading Danish companies within the field of energy. The participants are:

- APV Systems A/S, which is an international supplier of heat exchangers and units for the district heating sector.
- Danfoss A/S, that possesses a substantial product range within energy saving automatics and equipment for protection of improved environment.
- Grundfos A/S, that produces pumps for a wide range of purposes and dominates the market internationally in several fields.
- The Energy division at the Danish Technological Institute, which is one of Europe's most comprehensive institute units within the field of energy.
- The Department of Energy Engineering at the Technical University of Denmark.
- The Department of Mathematical Modelling at the Technical University of Denmark.

The objective of IT-Energy is to develop competence within the field of energy using IT based tools, and hereby make simulation of components and their interaction in systems possible.[1] The formulation and agreement of this milestone has risen from the recognized need among the participating companies, that in the development of the next generation HVAC system it is important to improve and develop:

(i) The understanding of the mode of operation of each individual component separately and it's interaction with the entire energy system. This is recognized as an important factor for meeting the future competition.

(ii) The use of IT in the communication between the components. This is a major factor for the future product development. An important milestone for IT-Energy is to contribute to this development, testing, and demonstration.

[1] Significant parts of this section follows the official declaration made for the project and can be found at the IT-Energy web site: www.ITEnergy.dk. The contents of this section is written in consent with the participants of IT-Energy.

In this context the field of energy includes industrially produced energy components, involved in the conversion of energy in energy systems. It is desired that the results and experiences from the project will influence a number of development areas, such as increasing the possibility of the companies for using simulation models as tools in their product development, design, planning phase, and marketing activities. Ultimately, this will lead to an increase in the functionality and effectiveness of Danish energy products by adding a large content of knowledge to the products and by using advanced technology as a major factor in maintaining the competitiveness. In summary, an overall goal is to help make Danish energy products more competitive from a technological point of view. This includes knowledge-building and developing tools for product development, such as implementation of new knowledge into products.

1.2 The necessity of dynamic HVAC&R models

The components relevant to typical Danish heating systems, are:

- Pumps
- Pipes and fittings
- Radiators
- Valves
- Plate heat exchangers
- Thermal zones and buildings

Following the objectives of IT-Energy there is a need to develop and improve the understanding of each individual component and to develop control strategies for the components interacting in heating systems. Developing new or improving existing products is traditionally a complex and expensive task including numerous tests in the lab followed by final tests of a prototype in situ. Today a large part of these tests can be replaced by computer simulation studies. The use of mathematical models in computer simulation studies has proven to be a perceptive and practical method, and the prospects for further improvement and development are good. Both steady state models and software products of these components exist, and thus the possibility of analyzing the steady state performance of a specific setup. For many applications low-level or steady state models are sufficient. In terms of steady state versus dynamic, the current consensus amongst the modelling community still seems to be that dynamic system operation can be approximated by series of quasi steady state operating conditions, provided that the time step of the simulation is large compared to the dynamic response time of the HVAC equipment (Hensen (1996)). However, dynamic models are critical when close process control is required and where calculations need to be performed almost on a second by second time scale in simulations. When the desired time scale is sufficiently small, e.g. less than one hour, dynamic models are important for energy analysis and closed loop control simulation. More specifically, the use of dynamic models is pertinent to applications such as:

- System design
- Energy analysis
- Design of control strategies
- Fault detection and diagnosis
- Product development

There exist a significant amount of literature on dynamic HVAC&R models. A good overview is given in Bourdouxhe et al. (1998), including a discussion on where the development of new models is required. Simulation aspects using dynamic HVAC&R models are discussed in Hensen (1996).

However, two problems arise concerning the possibility of making dynamic simulations with product specific components where accurate simulation performance is required:

1. The flexibility of the standard software products for dynamic models.
2. The validity of the standard software products for dynamic models when used for product specific setups.

Although a variety of dynamic models exist for most components, there is a strong need to develop models for product specific components within IT-Energy. It is important that the models are based on physical laws, but also that they are validated using empirical data. For this reason a stochastic framework is required. Stochastic modelling of components relevant to water heating systems is treated in (Jonsson (1990); Palsson (1993); Hansen (1997)), but the amount of literature on dynamic HVAC&R models is less than on the traditional deterministic approach. Since models of the product specific components are absent or poorly represented there is a strong need to develop dynamic models that are flexible and can be used for the above mentioned simulation applications within IT-Energy.

1.3 Scope of the thesis

The main objective of this thesis is to apply techniques related to grey box modelling in describing the dynamics of typical HVAC&R components. More specifically the objective is to:

(i) Investigate and demonstrate how statistical methods can be used in combination with physical interpretation in system analysis and identification of grey box models for the HVAC&R components. Emphasis will be put on statistical techniques for experimental design, parameter estimation and model validation.

(ii) By (i) to develop accurate dynamic models of HVAC&R components that can be used in system simulations. The models are intended to be used in a simulation tool within IT-Energy and should be physically interpretable, show good predictive abilities, as well as be suitable for simulation studies.

The objective is considered pertinent and newsworthy because the grey box modelling approach methods are not yet fully recognized as valuable tools for establishing models of HVAC&R components.

Chapter 2

Methodology

This chapter gives an introduction to the methods applied to identification of the dynamic models of the HVAC&R components. First, the concept of grey box modelling will be introduced. Then, grey box models in terms of stochastic differential equations (SDE's) are discussed briefly. Models in terms of SDE's are the foundation of the models presented in this thesis. The three subsequent sections, 2.3-2.5, discuss the methods used for the key techniques in the modelling procedure, namely experimental design, parameter estimation, and model validation. The methods are not necessarily exclusive to grey box modelling, although emphasis will be put on how they can be used in the grey box modelling framework.

2.1 Grey box modelling of energy systems

Basically, there exist two types of knowledge or information that can be applied in describing a system in terms of a mathematical model. The first information is the experience that experts have built, including the literature on the topic, and also the laws of physics. The other type of information is the system itself. Observations from the system and experiments on the system are the foundation for the description of the system and its properties, (Ljung and Glad (1994)). In principle, there are also two different approaches in constructing a mathematical model of a system. The first is to use well-known relationships, e.g. physical laws, and apply them to the system. The second approach is to apply observations from the system, and by doing so, adjust the properties of the model to the system. The first approach is often referred to as a deterministic approach or *white box* approach while the latter approach often is referred to as a stochastic or *black box* approach (Tulleken (1992)). When the two approaches are used in conjunction, the approach is often referred to as a *grey box* approach. According to Jørgensen and Hangos (1995) the following distinction can be made for different types of models:

- White box: The identification is performed without the use of experimental data, e.g. based only on first principles.
- Grey box: Both, a priori process knowledge and experimental data are used for identification, e.g. only a subset of parameters is estimated from experimental data.
- Black box: The identification is performed exclusively from experimental data.

The grey box modelling approach combines the deterministic and stochastic approach such that the model is based on both, prior knowledge about the system and information from data. Grey box modelling is thus a combination of the deductive and the inductive approach. Typically, the initial model structure and model constraints are determined by physical insight, while statistical procedures are applied for evaluating the model structure, estimating the model parameters and for model validation.

An advantage of using grey box models of physical systems is that it is possible to incorporate well known physical facts in the model structure, which is essential for many practical applications. It is also a more adequate approach for modelling of non-linear systems, which is the case for most physical systems, (Brohlin and Graebe (1995)), including HVAC&R systems, because the non-linear description is dictated by the laws of physics. Another key advantage is, that the use of statistical methods can reveal physical phenomena that were not considered initially.

The characterization of a grey box model is somewhat broad depending on the amount of prior knowledge used. For example, in the case no specific physical structural knowledge about the system is available, parameterized grey-box models cannot be used. Identification in black-box models is then the only alternative. However, certain non-structural knowledge about the system is sometimes available, e.g. that the step response is monotonic etc. This knowledge can also be incorporated in 'semi-physical modelling', cf. (Lindskog and Ljung (1995, 2000); Tulleken (1992)) on accounting for the constraints on the model parameters. On the other hand, if a system is well defined wrt. prior knowledge, only minor subsets may have to be considered as stochastic terms. In other words, there are also a variety of grayness in different kind of models, sketched in Fig. 2.1. In Harremoes and Madsen (1999) the balance between simplicity and complexity in model prediction applied to urban drainage structures is discussed. For introduction to and discussion on grey box modelling, see Ljung and Glad (1994); Tulleken (1992); Melgaard (1994).

Traditionally, the deterministic or white box approach has been applied in modelling of HVAC&R components and systems. The explanation may be found in the fact that the physical knowledge, i.e. laws of heat transfer and fluid mechanics, is a well-established science. The deductive physical knowledge and interpretation is the backbone of developing and analyzing models of HVAC&R systems. The physical knowledge is for most systems the foundation of under-

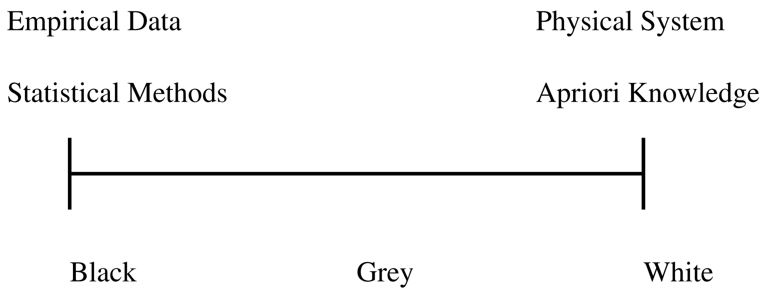


Figure 2.1: Conceptual sketch.

standing the system, and for practitioners of interpreting the model. Whether the practitioner has a background in electrical, mechanical or civil engineering, information technology or physics, the interpretation and understanding of the underlying physical laws is pertinent for almost any application.

Whereas the use of physical laws is recognized as the main approach in building models of HVAC&R components, the advantages of using statistical methods are in the treatment of empirical data from the system or a stochastic description of the system. The need for applying statistical methods is in the treatment of empirical data and establishment of grey box models of HVAC&R systems. This can perhaps best be illustrated by considering the variety of applications where such methods have been proven adequate and useful, e.g. in system identification, system analysis, design of control strategies, and fault detection and diagnosis. For these applications, the combination of empirical methods and physical interpretation is an important tool at every level of the system identification and analysis, from descriptive statistics of empirical data to detailed modelling of complex systems. Thus, the key issue in this thesis is to combine the deductive and inductive approach, and by doing so utilize the best of the two disciplines.

2.2 Grey box models in terms of SDE's

Deterministic models of dynamic systems are often formulated in terms of ordinary differential equations (ODE's) or partial differential equations (PDE's). However, continuous time grey box models are best formulated in terms of SDE's, cf. (Harremoes and Madsen (1999)). Models in terms of SDE's are adequate for several applications. First, the continuous time formulation ensures that a priori knowledge can be described as a set of coupled ODE's. For many physical systems a deterministic description in terms of ODE's may be obtained from physical reasoning. Second, SDE's provides for a description of the stochastic term of the system that is not accounted for elsewhere. Examples of such stochastic terms are noise, model approximations, unrecognized dynamics etc.

Models in terms of SDE's can often be formulated as state space models. The system dynamics are formulated in terms of a system of SDE's and referred to as the system equation (2.1):

$$d\mathbf{X}_t = \mathbf{f}(\mathbf{X}_t, \mathbf{U}_t, \theta, t)dt + \mathbf{G}(\mathbf{U}_t, \theta, t)d\mathbf{w}_t, \quad t \geq 0 \tag{2.1}$$

where $\mathbf{f}$ is a vector function describing the evolution of the system states $\mathbf{X}_t$ in time t. $\mathbf{U}_t$ is the input vector and the θ is a vector of known and unknown model parameters. This is equivalent to a deterministic system of ODE's. However, the vector function $\mathbf{G}$ describes how the noise is entering the system. The noise is modelled as a stochastic process, $d\mathbf{w}_t$ and it is assumed that it can be described by a standard Wiener process, cf. Øksendal (1985). Having discrete time measurements from the system it is possible to link the data to the system using a discrete system equation (2.2):

$$\mathbf{Y}_k = \mathbf{h}(\mathbf{X}_k, \mathbf{U}_k, \theta, t_k) + \mathbf{e}_k, \quad t_k \in \{t_0, t_1, \ldots, t_N\}, \tag{2.2}$$

where $\mathbf{h}$ is a functional relationship between the model output $\mathbf{Y}_k$ (N data points) and the model states, inputs, and parameters. $\mathbf{e}$ is assumed to be a Gaussian white noise sequence independent of $\mathbf{w}$ and describes the disturbance or noise of the measurement equation.

The close relationship between ODE's and SDE's makes the use of SDE's adequate for grey box modelling where prior information about the physics is available, and this is the case for most HVAC&R components. However, the use of models in terms of SDE's have been applied in a variety of fields and not only as grey box models. In some applications, such as modelling the behavior of stock prices, the deterministic component in the dynamics is only of secondary importance, and the volatility or stochastic drift is the phenomena that is sought for to be modelled (Hurn and Lindsay (1997, 1999); Nielsen et al. (2000)). Another example where the diffusion part is of primary interest is in first-passage problems, where SDE's have been applied to model and analyze fatigue crack of machinery components, cf. e.g. Ray and Patankar (1999). Thus, a model in terms of SDE's is not necessarily a grey box model.

In the modelling of HVAC&R components, however, the deterministic part of the SDE is crucial for the physical interpretation of the model. In such grey box models, the stochastic part often serves as a description of noise, model approximations and unrecognized dynamics, cf. Hansen (1997). Furthermore, these applications are often characterized with the fact that the deterministic system description is relatively accurate due to solid physical knowledge. Examples of grey box models in terms of SDE's are extensive. A few examples are stochastic simulation of biotechnical processes (Kinder and Wiechert (1996)), modelling of indoor air quality (Sowa (1998)), stochastic modelling of global atmospheric response to enhanced greenhouse warming with cloud feedback in Szilder et al. (1998), modelling of wastewater treatment processes in Tenno and Uronen (1995), heat exchanger modelling (Jonsson et al. (1992); Jonsson and Palsson (1992, 1994)), modelling of heat dynamics in buildings in Madsen and Holst (1995), modelling of water based central heating systems in Hansen (1997), and heat dynamics in greenhouses in Nielsen and Madsen (1998). More miscellaneous applications are the use of SDE's to estimate the threshold value for systems with hysteresis (Freidlin and Pfeiffer (1998)), modelling of game problems and controls in (Bally (1998)), and modelling of moments in present values in life insurance (Norberg (1995)).

2.3 Experimental design

For grey modelling of a physical system the use of empirical input/output data is involved in the identification of the system. Prior to this identification, the experimental design is an important aspect for most applications, including the modelling of HVAC&R components. A well-planned experiment is the connection between the experiment and the model that the experimenter can develop from the results of the experiment, cf. (Montgomery (1991)). The design of experiments is also the foundation for the subsequent parameter estimation and model validation. As sketched in Fig. 2.2, the modelling procedure is interpreted as an iterative process where the stages interconnect.

In the following, some techniques that have been proven useful for designing experiments with the purpose of a subsequent identification of dynamic models

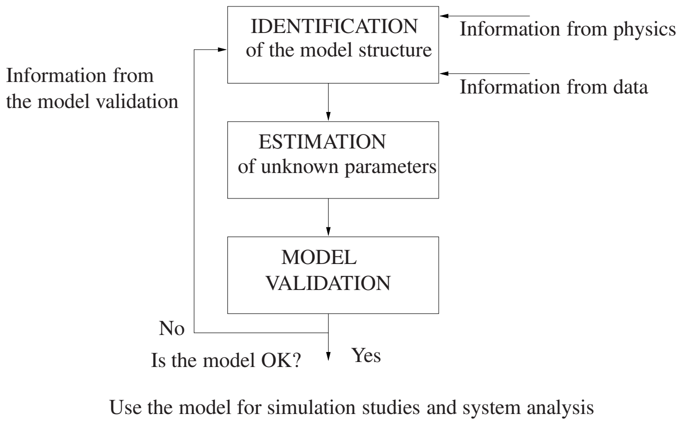


Figure 2.2: The modelling procedure.

will be described. First, some general aspects for practical situations will be discussed, followed by a brief discussion on techniques that may be suitable when the model structure is known a priori. Then, the generation of optimal input sequences will be discussed. Such signals are useful even if the model structure is not predetermined and the object of the experiment is to collect informative data for a subsequent analysis. Finally, some practical aspects and limitations concerning experiments on HVAC&R equipment will be discussed.

When building models to explain observed phenomena, one often uses a priori knowledge, such as physical, chemical or biological 'laws,' to propose possible models. In each case, these laws dictate the model structure, and we may wish to know whether one or more of such structures are adequate for the problem at hand. Given the model structures, the real purpose of experiment design is often to maximize the information content of the data within the limits imposed by the given constraints. Thus, any experimental design must take account of such constraints on the experimental conditions. Following some of the typical constraints that might be met in practice are (Goodwin and Payne (1977)):

1) Amplitude constraints on inputs, outputs, or internal variables.

2) Power constraints on inputs, outputs or internal variables.

3) Total time available for the experiment.

4) Total number of samples that can be taken or analyzed.

5) Maximum sampling rate.

6) Availability of transducers and filters.

7) Availability of hardware and software for analysis.

Given the constraints, the aim of the experimental design is then to ensure that the design variables are chosen in such a way that the experiment is maximally informative relative to the intended application. Other important issues to bear in mind are the class of models to be used, the identification method, and the criteria of best fit, the extent of prior knowledge about the system etc., see Goodwin (1982) for further details.

2.3.1 Optimal design criteria

There exist several possible ways of defining the optimal experiments when constraints and an initial set of models are considered. Examples of such experiments are design of input sequences that consists of maximizing the determinant of the Fisher information matrix, or D-optimality (Goodwin and Payne (1977); Emery et al. (2000)), constructions based on the sensitivity matrix (Point et al. (1996)), designs based on an overall measure of the divergence between the model predictions (Asprey and Macchietto (2000)), or designs based on maximizing the smallest eigenvalue of the Fisher information matrix (Antoulas and Anderson (1999); Sadegh et al. (1998)).

A drawback with the 'optimal solution' is, however, that it typically depends on unknown quantities, like the unknown system that we are trying to identify, e.g. when a prior model of the system might not be at hand, cf. (Forssell and Ljung (2000)). For practical applications it may be difficult or even impossible to construct such optimal designs. In these situations the design may be constructed to be as good as possible with respect to the experimental facilities as well as to the experimental constraints.

2.3.2 Design of input sequences

A possible alternative to the problem when the model structure is not known a priori is to apply perturbation signals designed for system identification (Godfrey (1993)). The signals may provide information that is optimal for a subsequent identification of functional relationships between input and output signals. In this context pseudo-random signals are the most popular choice for the persistently exciting perturbation signals required in system identification. In fact, pseudo-random signals are often close to the signals found by the model specific optimal criteria (Johansson (1993)). The pseudo-random signals can easily be created from any desired implementation and with desirable properties such as having autocorrelation function, $\rho(k)$, and cross-correlation function, $\rho_{uv}(k)$ that are not significant within a period, N, of the pseudo-random sequence. The autocovariance function, $\gamma(k)$, and cross-covariance function, $\gamma_{uv}(k)$, for the input sequences, u and v, are estimated by:

$$\gamma(k) = \frac{1}{N-1}\sum_{i=1}^{N}(u_i - \hat{u})(u_{i+k} - \bar{u}), \tag{2.3}$$

$$\gamma_{uv}(k) = \frac{1}{N-1}\sum_{i=1}^{N}(u_i - \hat{u})(v_{i+k} - \bar{v}), \tag{2.4}$$

where the autocorrelation function and cross-correlation function are found by scaling with the variances $V(u)$ and $V(v)$, respectively:

$$\rho(k) = \frac{\gamma(k)}{V(u)} \quad \text{and} \quad \rho_{uv}(k) = \frac{\gamma_{uv}(k)}{\sqrt{V(u)V(v)}}. \tag{2.5}$$

When input series are uncorrelated within a period of the pseudo-random sequence it holds that:

$$\rho(k) \simeq 0 \quad \text{for } |k| = 1, 2, 3, .., N, \tag{2.6}$$

$$\rho_{uv}(k) \simeq 0 \quad \text{for } |k| = 0, 1, 2, .., N. \tag{2.7}$$

Here, u_i and v_i are input sequences with mean $\bar{u}$ and $\bar{v}$ respectively. The signals may be determined by algorithms for pseudo-random signal generation. Further advantages are the possibility of designing signals, that are particularly suitable for obtaining the characterization of aspects of non-linear behavior, and also suitable for obtaining uncorrelated input sequences for multi-input signals. Methods for generating pseudo-random sequences exist in both, time and frequency domain, cf. (Goodwin and Payne (1977); Godfrey (1993, 1980); Zarrop (1979); Pazman (1986); Yarmolik and Demidenko (1988)).

For simplicity, only the characteristics of and a method for generating pseudo-random binary signals (PRBS) will be described here. As the name indicates, the PRBS sequence is a sequence of numbers having two possible levels. Other methods for generation and signals derived from PRBS signals as well as the extension to multi-level signals are treated in detail in (Godfrey (1993); Yarmolik and Demidenko (1988)).

Following Godfrey (1993) the PRBS has the following properties:

1. The signal has two levels, and it can switch from one level to the other only at certain event points, $t = 0, \Delta t, 2\Delta t...$
2. The signal is predetermined, meaning that the signal is deterministic and thus repeatable.
3. The PRBS is periodic with period $T = N\Delta t$, where N is an odd integer.
4. In any one period, there are $\frac{1}{2}(N+1)$ intervals when the signal is at one level and $\frac{1}{2}(N-1)$ intervals when it is at the other.
5. The autocorrelation function is only significant in lag 0.

Binary m-sequences exist for $N = 2^n - 1$ where n is an integer (>1). They can easily be generated using an n-stage feedback shift register with feedback to the first stage consisting of the modulo-2 sum of the logic value of the last stage and one ore more of the other stages. The binary logic values are taken as 1 and 0 and modulo-2 addition is given by:

$$0 \uplus 0 = 0 \quad 0 \uplus 1 = 1 \quad 1 \uplus 0 = 1 \quad 1 \uplus 1 = 0. \tag{2.8}$$

An example from a four-stage shift register with feedback is shown in Fig. 2.3. For a detailed discussion on the properties see Godfrey (1993).

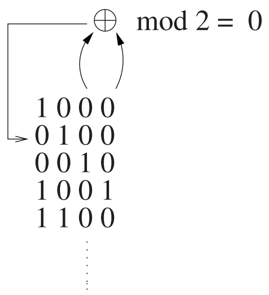


Figure 2.3: Generation of a PRBS sequence, exemplified using a shift register circuit with $n = 4$. In the first step a zero is generated (0+0 mod 2 = 0). The generated number is put back in the shift register. In the second step, a new zero is generated (0+0 mod 2 = 0) followed by a one in the next step (1+0 mod 2 = 1) etc.

In summary, by using simple shift registers, pseudo-random number sequences can be generated and used to determine the level of the input sequences in experiments. The input sequences are optimal in the sense that they are not correlated, which is essential for the subsequent model identification.

2.3.3 Practical Issues

The use of pseudo-random input sequences in this study has been used for experiments with HVAC&R equipment. Having determined the input/output relations for the considered component or system, the input sequences need to be scaled to comply with the system constraints, such as the amplitude constraints, number of controllable levels, sampling rate and time available for the experiment. An example could be a valve opening and closing determined by the PRBS values of 1 and 0, or e.g. a supply temperature switching between 30, 40, 50 and 60 °C using a multilevel signal. Concerning the availability of hardware and software for analysis, an empirical base for experiments of the components in consumer installations has been established within IT-Energy . A photo of a test rig built for IT-Energy is shown in Fig. 2.4. In these setups it is possible to excite the systems with different load characteristics without disturbing any consumers. The use of pseudo-random input sequences is adequate for such practical experiments since they are easily controlled from a computer implementation. Typically, fluid flows may be controlled according to the input sequences, while fluid temperatures cannot be controlled as accurately. The constraints concerning sampling rate and time available for the experiments, are usually not critical.

Figure 2.4: Experimental setup at the Danish Technological Institute.The heating system corresponds in capacity to 20 apartments, consisting of three water tanks, heat exchangers, valves, pumps and pipes as well as data acquisition equipment.

2.4 Parameter estimation

This section presents the applied estimation method for models formulated in terms of SDE's. Parts of the mathematical description follows Melgaard (1994) closely.

There exist several parameter estimation methods for SDE's. An extensive review is given by Nielsen et al. (2000), where different estimation methods, such as the generalized method of moments, the efficient method of moments and indirect inference, are discussed. Furthermore, Monte Carlo methods, transfer function models and non-linear filtering methods are reviewed. In (Hurn and Lindsay (1997, 1999)) a non-parametric method based on the maximum likelihood principle and Monte Carlo techniques is used to estimate parameters of SDE's for simulated data. In Bianchi and Cleur (1996) an indirect estimation method of SDE's is given. The problem of solving SDE's numerically is treated in Kloeden et al. (1997) while more theoretical issues concerning parameter estimation in linear SDE's are discussed in (Khasminskii et al. (1999); Kim (1999)). Finally, parameter estimation in nonlinear SDE's is discussed in Timmer (2000). For a more general discussion on SDE's, see (Øksendal (1985); Gard (1988)).

2.4.1 Maximum Likelihood parameter estimation

In this section, the Maximum Likelihood (ML) estimation method, used in the subsequent analysis, is discussed as applied to models formulated in terms of SDE's.

As discussed in section 2.2, dynamic HVAC&R models can often be formulated in terms of a system equation: (2.9):

$$d\mathbf{X}_t = \mathbf{f}(\mathbf{X}_t, \mathbf{U}_t, \theta, t)dt + \mathbf{G}(\mathbf{U}_t, \theta, t)d\mathbf{w}_t, \quad t \geq 0 \tag{2.9}$$

where discrete time measurements from the system are used to link the data to the system using a system equation (2.10):

$$\mathbf{Y}_k = \mathbf{h}(\mathbf{X}_k, \mathbf{U}_k, \theta, t_k) + \mathbf{e}_k \quad t_k \in \{t_0, t_1, \ldots, t_N\}. \tag{2.10}$$

It should be noted that the state space formulation allows for a description that distinguishes between the uncertainty of the model formulation of the system equation as well as the uncertainty in the measurements and the observation equation.

Assuming that the experimental data $\mathcal{Y}(t) = [\mathbf{Y}_t, \mathbf{Y}_{t-1}, \ldots, \mathbf{Y}_1, \mathbf{Y}_0]$ is a realization of stochastic variables and that the model parameters are normally distributed, it is possible to construct the likelihood function. It is well-known that the ML estimates are found as the estimates which maximize the likelihood function, see e.g. Goodwin and Payne (1977). The unconditional likelihood function $\mathbf{L}(\theta; \mathcal{Y}(N))$ is the joint probability of all the observations assuming that the parameters are known:

$$\mathbf{L}(\theta; \mathcal{Y}(N)) = p(\mathcal{Y}(N)|\theta). \tag{2.11}$$

Applying Bayes rule to the likelihood function can be written as a product of conditional densities:

$$\begin{aligned} \mathbf{L}(\theta; \mathcal{Y}(N)) &= p(\mathcal{Y}(N)|\theta) \\ &= p(\mathbf{Y}_N|\mathcal{Y}(N-1), \theta)p(\mathcal{Y}(N-1)|\theta) \\ &= \left(\prod_{t=1}^{N} p(\mathbf{Y}_t|\mathcal{Y}(t-1), \theta)\right) p(\mathbf{Y}_0|\theta). \end{aligned} \tag{2.12}$$

The conditional likelihood function (conditioned on $\mathbf{Y}_0$) becomes:

$$\mathcal{L}(\theta; \mathcal{Y}(N)) = \prod_{t=1}^{N} p(\mathbf{Y}_t|\mathcal{Y}(t-1), \theta). \tag{2.13}$$

The normal distribution is characterized by the mean and the covariance. In order to parameterize the conditional distributions, the conditional mean and conditional covariance are introduced as:

$$\hat{\mathbf{Y}}_{t|t-1} = \mathrm{E}[\mathbf{Y}_t|\mathcal{Y}(t-1),\theta] \quad \text{and} \quad \mathbf{R}_{t|t-1} = \mathrm{V}[\mathbf{Y}_t|\mathcal{Y}(t-1),\theta], \tag{2.14}$$

respectively. Eq. (2.14) represents the one-step prediction and the associated covariance, which are calculated using a Kalman filter, discussed in Section 2.4.2. The one-step prediction errors are calculated by:

$$\epsilon_t = \mathbf{Y}_t - \hat{\mathbf{Y}}_{t|t-1}. \tag{2.15}$$

Then the conditional likelihood function (2.13) becomes:

$$\mathcal{L}(\theta;\mathcal{Y}(N)) = \prod_{t=1}^{N} \left((2\pi)^{-m/2} \det \mathbf{R}_{t|t-1}^{-1/2} \exp(-\tfrac{1}{2}\epsilon_t' \mathbf{R}_{t|t-1}^{-1} \epsilon_t) \right), \tag{2.16}$$

where m is the dimension of the $\mathbf{Y}$ vector. By taking the logarithm of the conditional likelihood function, we obtain:

$$\log \mathcal{L}(\theta;\mathcal{Y}(N)) = -\tfrac{1}{2} \sum_{t=1}^{N} \left(\log\ \det \mathbf{R}_{t|t-1} + \epsilon_t' \mathbf{R}_{t|t-1}^{-1} \epsilon_t \right) + \text{const}, \tag{2.17}$$

the ML estimate of θ is found as the value that maximizes the conditional likelihood function $\mathcal{L}(\theta;\mathcal{Y}(N))$:

$$\hat{\theta} = \arg\min_{D_{\mathcal{M}}} \sum_{t=1}^{N} \left(\log\ \det \mathbf{R}_{t|t-1} + \epsilon_t' \mathbf{R}_{t|t-1}^{-1} \epsilon_t \right). \tag{2.18}$$

Here $D_{\mathcal{M}}$ is the set of possible values for θ. It should be noted that the ML estimator is asymptotically normally distributed with mean θ and variance:

$$\mathbf{D} = \mathbf{H}^{-1}, \tag{2.19}$$

where $\mathbf{H}$ is the Hessian given by:

$$\{h_{lk}\} = -\mathrm{E}\Big[\frac{\partial^2}{\partial\theta_l \partial\theta_k} \log \mathcal{L}(\theta; \mathcal{Y}(N))\Big]. \tag{2.20}$$

$\{h_{lk}\}$ denotes the element in row l and column k of $\mathbf{H}$, and θ_j denotes element j of θ.

An estimate of $\mathbf{D}$ is obtained by equating the observed value with its expectation and applying:

$$\{h_{lk}\} \approx -\Big(\frac{\partial^2}{\partial\theta_l \partial\theta_k} \log \mathcal{L}(\theta; \mathcal{Y}(N))\Big)\Big|_{\theta=\hat{\theta}}. \tag{2.21}$$

Hereby, Eq. (2.21) is used to estimate the variance of the parameter estimates.

It should also be noted that optimal experimental designs, as discussed in Section 2.3, may be constructed to minimize certain measures of $\mathbf{D}$, i.e. the optimization criteria is to minimize the variance (uncertainty) of the parameter estimates. This requires however, that the model structure is known prior to the experiment.

2.4.2 The Kalman filter

In the ML estimation method described in section 2.4.1, the one step predictions and associated covariances are needed for calculating the likelihood function. In the estimation of the model parameters the Kalman filter is applied to estimate these quantities. It should be noted that the Kalman filter is derived for linear systems. For non-linear systems the extended Kalman filter, based on linearizations of the system equation, is applied. Given that the system equation is described by Eq. (2.9) and the observation equation by Eq. (2.10), the extended Kalman filter for the prediction equations become:

$$\frac{d\hat{\mathbf{X}}_{t|k}}{dt} = \mathbf{f}(\hat{\mathbf{X}}_{t|k}, \mathbf{U}_t, \theta, t), \quad t \in [t_k, t_{k+1}[, \tag{2.22}$$

$$\begin{aligned}\frac{d\mathbf{P}_{t|k}}{dt} &= \mathbf{A}(\hat{\mathbf{X}}_{t|k}, \mathbf{U}_t, \theta, t)\mathbf{P}_{t|k} \\ &+ \mathbf{P}_{t|k}\mathbf{A}'(\hat{\mathbf{X}}_{t|k}, \mathbf{U}_t, \theta, t) \\ &+ \mathbf{G}(\theta, t)\mathbf{G}'(\theta, t), \quad t \in [t_k, t_{k+1}[,\end{aligned} \tag{2.23}$$

where $\mathbf{A}$ is obtained by a linearization of the system equation (2.9):

$$\mathbf{A}(\hat{\mathbf{X}}_{t|k}, \mathbf{U}_t, \theta, t) = \frac{\partial \mathbf{f}}{\partial \mathbf{X}}\Big|_{\mathbf{X}=\hat{\mathbf{X}}_{t|k}}. \tag{2.24}$$

Let $\mathbf{C}$ be the linearization of the observation equation (2.10):

$$\mathbf{C}_k = \mathbf{C}(\hat{\mathbf{X}}_{k|k-1}, \mathbf{U}_k, \theta, t_k) = \frac{\partial \mathbf{h}}{\partial \mathbf{X}}\Big|_{\mathbf{X}=\hat{\mathbf{X}}_{k|k-1}}. \tag{2.25}$$

Applying the Kalman filter the updates at t_k are:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{C}'_k[\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}'_k + \mathbf{S}(\theta, t_k)]^{-1}, \tag{2.26}$$

$$\hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} + \mathbf{K}_k(\mathbf{Y}_k - \mathbf{h}(\hat{\mathbf{X}}_{k|k-1}, \mathbf{U}_k, \theta, t_k)), \tag{2.27}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{C}_k\mathbf{P}_{k|k-1}. \tag{2.28}$$

For a detailed reference on the applied estimation method, see Melgaard (1994).

2.4.3 Practical issues

Given that an initial model is formulated in terms of SDE's, the model parameters θ and the model states may be estimated by the ML method using a software package CTLSM, cf. (Madsen and Melgaard (1993); Melgaard (1994)). The package provides numerical estimates using a quasi-Newton method as well as relevant test statistics. The software package has been used in a variety of modelling applications for physical models, and also in the case studies presented in this thesis.

2.5 Model validation

As the final step in the modelling procedure, the aspects of model validation are considered. Model validation is an important step in identifying models. It concerns the investigation and testing of the established model, and may be used to decide whether the model obtained from the identification procedure can be accepted, and if not, how to improve it. If the identification method has provided the best possible model in the model set, the problem is to decide if the chosen model is a suitable one. This question can be approached in two ways, cf. Ljung (1982). One way is to evaluate the properties of the model and decide whether it meets reasonable requirements. Another approach is to look for other models or model types and make the comparison. The latter procedure can be regarded as an a posteriori choice of model sets.

An appropriate first validation criterion is a preliminary comparison with the model and the system. This comparison, which often is subjectively, may be done by performing simulations with the model and verify that certain criteria, e.g. that the steady state performance, is reasonable. A good insight may be achieved by plotting dependent variables against simulated output. Visualization of the model performance will often give a clear indication if the model captures the important features of the system, such as time constants, power constraints, steady state conditions etc. Furthermore, visualization of data can detect obvious model limitations and flaws, or reveal structures in the data that cannot be absorbed in any other way, cf. Cleveland (1993).

2.5.1 Residual analysis

Many objective methods for validation of grey box models are closely related to validation techniques for black box models. For both types of model classes the residual analysis, where residuals are defined as the deviation between the model and observations from the system, is an important part of the model validation. The purpose of residual analysis is primarily to check whether any obtained information contradicts the assumptions upon which the models and methods are built, cf. (Holst et al. (1992)). Furthermore, the residual analysis may be used to indicate how the model should be extended.

If the measurement uncertainty of the data is known, the amplitude of the model residuals can be compared with the experimental uncertainty. If the uncertainty of the model output is not known directly, but is some function of measured data, the model uncertainty can be calculated using the method of propagation of errors, (Coleman and Steele (1999)). Consider a model such that $\Phi = f(\delta, \gamma, \lambda)$ and assume that δ, γ, and λ are measured with known uncertainties σ_δ^2, σ_γ^2 and σ_λ^2, respectively. Then the uncertainty σ_Φ^2 in Φ can be approximated by the error propagation formula assuming the measurement errors of different variables to be uncorrelated, i.e., assuming the covariance terms to be zero:

$$\sigma_\Phi^2 = \sigma_\delta^2 (\frac{\partial f}{\partial \delta})^2 + \sigma_\gamma^2 (\frac{\partial f}{\partial \gamma})^2 + \sigma_\lambda^2 (\frac{\partial f}{\partial \lambda})^2. \tag{2.29}$$

Usually the sequence of model residuals, $\epsilon_i \;\; i = 1..N$, is assumed to be normally distributed with mean 0 and variance σ_ϵ^2, i.e. $\epsilon_i \in N(0, \sigma_\epsilon^2)$. Thus, the value of the estimated variance σ_ϵ^2 can then be compared to Eq. (2.29).

The assumption that the model residuals are normally distributed is realistic from a theoretical point of view, cf. (Coleman and Steele (1999)). Important identification methods, such as the ML method in section 2.4, are built on this assumption. Therefore, it is important to analyze whether or not the model residuals can be regarded as being normally distributed. In general, an adequate model should have residuals that are free of systematic patterns that the model otherwise is failing to explain. Systematic deviations can be analyzed in multiple ways. A natural first step in the residual analysis is to graph the residuals and check for outliers and non-constant variance. It is also helpful to graph the residuals against input and output variables in order to investigate systematic patterns in the residuals.

Time domain tests

There exist several statistical tests, that can be performed as a supplement to the graphical test. These tests may be performed in both the time domain and in the frequency domain. In the time domain, the residuals may be tested using non-parametric methods (Haerdle (1990)), or by applying standard parametric techniques (Box and Jenkins (1970)). The non-parametric techniques can be used for analyzing the distribution of the residuals. The distribution estimate can be used in the grey box validation for the residual sequence and for the model as a whole (Holst et al. (1992)).

Especially for dynamic models, the autocovariance function can be used to test the model residuals in the time domain. The autocovariance function is estimated by:

$$\hat{\gamma}_\epsilon(k) = \frac{1}{N}\sum_{i=1}^{N}(\epsilon_i - \bar{\epsilon})(\epsilon_{i+k} - \bar{\epsilon}), \tag{2.30}$$

where $\bar{\epsilon}$ is the estimated mean of the residual sequence ϵ_i. The estimated correlation coefficient at lag k is:

$$\hat{\rho}_\epsilon(k) = \frac{\hat{\gamma}_\epsilon(k)}{\hat{\gamma}_\epsilon(0)}. \tag{2.31}$$

If the residuals ϵ_t of a scalar process are normally distributed, the estimated autocorrelation function is:

$$\hat{\rho}_\epsilon(k) = \begin{cases} 1 & k = 0 \\ 0 & |k| = 1, 2, .., \end{cases} \tag{2.32}$$

and

$$\hat{\rho}_\epsilon(k) \in_{approx} N(0, \frac{1}{N}). \tag{2.33}$$

Approximative 95% limits for this distribution are $\pm 2\sigma = \pm 2/\sqrt{N}$. Plots of both, the estimated autocorrelations for the residuals and the confidence limits, provide a graphical way to test the hypotheses about the noise assumption, i.e. if the residuals are normally distributed.

Frequency domain test

In a similar manner, the model residuals may also be analyzed in the frequency domain. The periodogram for the residuals for the frequencies $\nu_i = \frac{i}{N}, \quad i = 0, 1, \ldots, N/2$, is:

$$\hat{I}(\nu_i) = \frac{1}{N}\left[\left(\sum_{t=1}^{N} \epsilon_t \cos 2\pi\nu_i t\right)^2 + \left(\sum_{t=1}^{N} \epsilon_t \sin 2\pi\nu_i t\right)^2\right]. \tag{2.34}$$

The periodogram is a frequency domain description of the variation of the residuals, as $I(\nu_i)$ indicates how much of the variation of the residuals is present at the frequency ν_i. The normalized cumulative periodogram becomes:

$$\hat{C}(\nu_j) = \frac{\sum_{i=1}^{j} \hat{I}(\nu_i)}{\sum_{i=1}^{N/2} \hat{I}(\nu_i)}, \tag{2.35}$$

which is a non-decreasing function, defined for the frequencies ν_i. For normally distributed residuals the variation is uniformly distributed over the frequencies, and is often referred to as white noise, due to the same spectral properties for white light. The total variation for N observations is $N\sigma_\epsilon^2$, and hence the theoretical periodogram for white noise is:

$$I(\nu_i) = 2\sigma_\epsilon^2. \tag{2.36}$$

The theoretical cumulative periodogram is thus a straight line from (0,0) to (0.5, 1). If the residuals are white noise, it is expected that $\hat{C}(\nu_i)$ is close to this line. Confidence intervals around the straight line can be calculated using a Kolmogorov-Smirnov test, see (Box and Jenkins (1970)). The advantage of applying frequency domain test may be that it is possible to determine the frequencies of dynamic patterns.

Further issues concerning residual analysis

The above-mentioned techniques for residual analysis are not always adequate. Both, the test in the autocorrelation function and frequency domain test, are linear tools and may not be able to detect non-linearities. Non-linear treatment could be investigated non-parametricly functionals (Haerdle (1990)), that accounts for non-linearities, such as the mutual information coefficient Granger and Lin (1994), or the lag dependent functions, proposed by Nielsen and Madsen (2001). For alternative methods for residual analysis, cf. (Wong (1997); Li and Hui (1994); Li (1998)).

2.5.2 Tests and physical interpretation of grey box models

If the residual analysis supports the model assumptions, the natural next step is to investigate the physical characteristics of the model in more detail. Typically, this includes the investigation of estimated parameter values, that the time constants and model states agree with the physical characteristics etc. For an experienced practitioner these tests might be done intuitively based on a solid insight into the system. However, if the model at a first sight seem reasonable, more sophisticated mathematical methods can be applied to investigate the system in more detail. These methods can indicate flaws that are not easily found otherwise.

The estimated model parameters may be tested for significance using the fact that the ML method provides estimates that are approximately normally distributed, cf. section 2.4. This enables test for the significance of the parameters, i.e. to test whether the parameters are significantly different from zero:

$$H_0 : \theta_j = 0 \quad \text{against} \quad H_1 : \theta_j \neq 0. \tag{2.37}$$

The test quantity is $\frac{\hat{\theta}_j}{\hat{\sigma}_{jj}}$, where $\hat{\theta}_j$ denotes the j-th parameter estimate and $\hat{\sigma}^2_{jj}$ the associated variance estimate. As the parameter estimates are asymptotically normally distributed, the test value is t distributed, and then a t-test of the hypothesis can be performed. Using this test it is possible to neglect model parameters in the model which are not significant with the data at hand.

Often initial guesses or a priori information about parameter values are available which allow for a comparison if the parameter values are reasonable. Thus, the above test Eq. (2.37) may be used to test for other parameter values than zero, e.g. test if an estimated physical parameter can be assumed to equal some value that corresponds to physical interpretation.

2.5.3 Cross validation and model selection

Another important test is to perform a cross validation study, i.e. to investigate the model performance using independent data. In such a study the residual analysis (quantitative) is not as important as a more qualitative evaluation. This includes the extrapolation abilities of the model, which is usually assumed stronger for a grey box model compared to a black box model.

There exist several methods to compare sets of models. One criteria is to evaluate if different models fulfill the criteria on residual analysis, the degree of physical interpretation, cross validation (Shao (1993)), and extrapolation abilities. Other methods are based on model selection criteria, cf. e.g. McQuarrie et al. (1997), or likelihood ratio tests Holst et al. (1992).

2.5.4 Practical issues

In general, it is advisable to combine several validation tools, since some methods may be more adequate in certain situations than others. Also, the identification proceduce is an iterative process, and it is usually necessary to validate the adequacy of model modifications. In the presented case studies in this thesis, however, only the final validation results are presented. These validation methods may easily be done using standard software, and in some software packages, validation results are presented along with the parameter estimation results.

Chapter 3

Overview of included papers

This section gives an overview of the included papers in the thesis. The focal point of each paper is the application aspect, i.e. each paper either discusses grey box modelling techniques or models for HVAC&R components related to practical applications. Although different models and methods are presented, the main issue of each paper is the same, namely addressing the use of grey box modelling methods for HVAC&R components and applications. The grey box methods for experimental design, parameter estimation, and model validation, have been applied throughout the papers.

Paper A is an introductory paper, which presents the modelling perspectives of the joint project IT-Energy. A short introduction to the scope of application, with emphasis on the design of simulation models for energy analysis of HVAC&R components and systems, is given. Issues concerning the development of the simulation models are discussed, e.g. on determination of the model interface, relevant time scale and aspect of simplicity versus complexity. This discussion is of more or less general nature and an indirect result of the research work developed through IT-Energy and inspired from topics in the literature. The paper also presents more project related issues, such as experimental setups and model validation issues as well as a discussion on the applied modelling method. The paper is modified from its original version by excluding the modelling examples. These are instead referred to as Paper C and F in the thesis. The paper is intended as a presentation of IT-Energy to other researchers within the field of HVAC&R and is therefore somewhat broad. In this thesis it serves as an introduction to and background for the subsequent papers, which deal more specifically with detailed modelling of the HVAC&R components.

As the first paper on modelling of HVAC&R components, Paper B discusses the aspects of experimental design related to heat exchanger modelling. The aim of the experiment is to collect data for a succeeding modelling of the heat dynamics of the heat exchanger. The paper addresses issues relevant to the design of input sequences and discusses a method for generating input signals. The experimental setup is presented as well as an analysis of the experimental data. Finally, the data is used to validate a dynamic model of the heat dynamics of the heat exchanger. The main objectives of the paper are 1) to stress the importance of and 2) to illustrate techniques for designing appropriate experiments for HVAC&R components. The design of input sequences is related to design of Pseudo Random Number Signals. These signals have proven useful in system identification and can easily be applied to other components. Also, practical aspects are discussed. These include the choice of experimental setup, sampling time, and investigation of time constants.

The modelling of the fluid dynamics in pipes is presented in Paper C. Emphasis is put on how statistical methods can be used to build physical threshold models. Threshold models may be useful in the modelling of HVAC&R components since it is a method to account for model discontinuities, e.g. change in flow characteristics or a valve closing. The method of using a smooth threshold function may be useful for parameter identification and for simulation purposes. The paper describes the experimental setup and design of input variables as well as the applied threshold function. The estimation procedure is discussed and the results are presented. Although applied to the modelling of the fluid dynamics in pipes, the method may be used for other applications as well.

A dynamic model of a radiator is presented in Paper D. The model is obtained from physical reasoning and experimental data. A main issue about the model is that the lumping of the radiator is done with varying section sizes depending on the flow. The lumping technique of equal section sizes is well known from many systems and the resulting model is obtained in terms of ODE's. However, preliminary analysis and physical interpretation suggest that there is an advantage in using varying section sizes. Especially it is found that the varying section sizes are helpful for reducing the model order. The model performance is illustrated and applications and limitations are discussed.

Paper E presents a model of a thermostatic valve with hysteresis effects. Hysteresis is a phenomenon that appears in many physical systems and which is not easily modelled. In this paper an adaptive model for friction compensation accounts for the hysteresis. Using experimental data the effects from hysteresis are clearly seen and has been analyzed. The valve position is determined by a steady state expression based on physical reasoning. By formulating the model in terms of SDE's the unknown quantities, especially the hysteresis force, are estimated. The model performance is illustrated and further extensions of the model are discussed.

While Papers B-E concern modelling and modelling aspects of single components, a model of a system, namely a model for the heat dynamics of a building, is presented in Paper F. The model is used for a building consisting of three rooms that are affected by inputs from the sun, ambient air, and radiators. Simple models are used to describe these external sources. The model approximations, and especially the simplified grey box approach, is discussed as compared to more traditional methods. The model is estimated and validated using experimental data from designed experiments and both physical interpretation and statistical techniques are used to validate the model. In summary, this model shows how simple models can be connected in order to model more complex system and also stresses the importance of empirical methods in deriving such models.

Finally, Paper G presents a regression approach (Error in Variable) as a means of identifying unbiased physical parameter estimates. The paper is different from Paper A-F in the fact that the model, which is a grey box model of the Coefficient of Performance for a commercial chiller, is only used for steady state calculations, and not in terms of SDE's. The paper investigates the use of the model for fault detection and diagnosis. It is illustrated that sources of noise, such as measurement error, may influence the physical parameter estimates in terms of bias and thus influence on the quality of the fault detection. The error in variable approach is applied to illustrate how the bias arising from measurement errors can be removed. Furthermore, the paper suggest different methods for correction for the measurement error bias for steady state grey box models.

Chapter 4

Conclusion

The main issue in this thesis has been the application of the grey box modelling method to dynamic models of HVAC&R components. The application aspects and the applied methodology have been presented in the summary report, while the research work consists of 7 case studies. The case studies demonstrate the techniques and benefits in combining the deductive and inductive approach in the modelling of typical HVAC&R components. As in the summary report, the papers focus on the grey box modelling process and distinguish between the disciplines of experimental design, parameter estimation, and model estimation, as interrelated disciplines in building models. These methods are of primary interest, whereas the models are found as the best choice given the data. In that respect, the thesis does not give a set of 'optimal' models, but points out how statistical methods can be applied to handle modelling problems that cannot be solved easily otherwise.

As the first step in the modelling procedure, Paper B discusses issues relevant to experimental design for subsequent dynamic modelling of HVAC&R equipment. The paper discusses the design of input sequences for heat exchanger modelling and the quality of the acquired data. Experimental designs have also been applied in the modelling of the flow in pipes in Paper C, and in Paper F concerning the modelling of the heat dynamics of buildings. From the research work it is concluded, that there are, in general, obvious benefits in performing experiments with a system whenever this is possible. By proper experimental design it is possible to acquire data that is informative and representative for the system. This is helpful in the subsequent model identification compared to non-intrusive experiments, where the variables are measured but cannot be controlled. Even if the input variables can be controlled, there is most often a set of constraints to a given experiment that has to be accounted for. However, a well-planned experiment can ensure that data spans the modelling space and that it is possible to identify the dynamics of interest given these experimental constraints. For the experimental designs applied in this study it has been found that the use of Pseudo Random Number Signals (RPNS) are adequate and the signals can be constructed to meet typical experimental constraints. In that respect the PRNS provide input sequences that are optimal for the subsequent parameter estimation and model validation. It can be argued that more sophisticated experimental designs could have been applied in case an a priori model structure was assumed. The use of PRNS included a trade off between very detailed experimental plans and simple or non-intrusive measurements. It is concluded that the use of PRNS along with the well-defined experimental conditions available has provided high quality data for subsequent analysis. The PRNS have desirable statistical properties and are easy to generate and control in experiments. However, an interesting aspect would be to investigate the benefits and difficulties with more sophisticated experimental designs.

The ML method for parameter estimation has been applied in the Papers B-F. The advantage of determining model parameters by statistical methods is that these parameter values may not easily be known or calculated otherwise. It is concluded that the ML method is adequate for modelling of HVAC&R components because it allows for estimating both, model parameters and model states, in models formulated in terms of SDE's. A key advantage of the ML method is that the continuous time formulation ensures that the models are physically meaningful and that the model parameters can be interpreted directly. Furthermore, the ML parameter estimation technique may provide estimates of state variables that are non-observable and of interest in simulation studies, e.g. the hysteresis force of the thermostatic valve in Paper E or the temperature of the floor in the heat dynamics of buildings in Paper F. From a statistical point of view, the parameter estimates provided by the ML method are normally distributed and this can be exploited by statistical tests for parameter significance and model reduction. Hereby, it can be statistical determined which terms that are significant in a physical model. A drawback is, however, that the ML method requires knowledge in statistical modelling. Furthermore, the ML method provides estimates that are optimal to fit the data and does not question whether the model is adequate or not, with respect to physical meaning. In such cases physical insight can be used to determine whether the model gives sound results or not. If the model assumptions for the ML method are not valid there may also be doubts for the subsequent inference made by the model. In such cases the method could be regarded as a prediction error method and the parameter estimates as values that gives the best fit in terms of predictive ability. Another aspect is that the ML method provides estimates optimal for one-step prediction while for many applications the real purpose is simulation. It would be possible to modify the optimization criteria to account for the simulation performance of the model instead. The drawback with this issue is, however, that the properties and assumptions of the model may fall apart and influence on the tests for model validation. In such case a more subjective validation criteria must be applied.

The techniques for model validation give some criteria for model verification. These techniques are well known from system identification and serve as a complement to the physical interpretation of the system. The techniques demonstrated in paper B-F are to some extent objective measures and related to residual analysis. This makes sense because it can be used to validate the prediction performance as well as the model assumptions. These techniques do not, however, reveal whether the model is adequate or detailed enough for other applications, for e.g. extrapolation studies, although a grey box model should be more reliable in extrapolation compared to a black box model. To determine the extrapolation abilities, the physical interpretation of the model can be used for further analysis, such as sensitivity analysis. A conclusion is that the techniques related to residual analysis are adequate to determine whether the model explains the data. The physical interpretation and tests for parameter significance are also helpful in deciding whether the model is adequate. However, when the physical interpretation of the model is crucial, the presented validation methods should only be used as a supplement to the model selection process. The presented validation methods primarily reveal features in the data, and cannot be used as a substitute for physical insight.

From a practical viewpoint the presented models have been used in the development of a simulation tool for heating systems. Although the thesis does not directly concern the development of the simulation tool, it is argued that the grey box modelling method can be used to establish adequate models for simulation purposes compared to other modelling approaches. The term 'adequate' is relative to the modelling objective. In simulation studies a low model order can reduce the computational effort and hereby increase the speed of the simulations. In some cases a low model order is necessary in order to make simulations computationally feasible. Examples of models that are computationally 'fast' are presented in Paper B, C and D. Paper D presents a dynamic model of a radiator where the model order is reduced due to the introduction of variable section sizes in the lumping of the radiator. This makes simulations computationally faster compared to a model with a large number of sections. Other modelling aspects are influencing the computational speed as well. Paper C presents a dynamic model of a pipe where the regime between the turbulent and laminar flow is modelled using a smooth threshold. This is an advantage in simulation studies because algebraic loops may be avoided (solvability), i.e. an algebraic routine that determines whether the actual flow is laminar or turbulent. As a third example, the heat exchanger model is evaluated in paper B. The statistical expression used for the heat transfer coefficient is both, accurate and computationally fast, as compared to deterministic expressions for the heat transfer coefficient. The other main characteristic is that the models are directly physically interpretable. This means that practitioners not familiar with statistical issues or modelling techniques can interpret the model parameters and model states. For simulation purposes the practitioner may experiment with different parameter values to investigate the behavior of the component because the parameters have physical meaning. For example, the behavior of the thermostatic valve presented in Paper E can be investigated when certain characteristic parameters are modified. This is beneficial e.g. in developing new components or improving the understanding of the existing ones. Another example is the chiller model evaluated in Paper G. Here the COP of the chiller may be investigated as the thermal resistance is varied.

In summary, stochastic models of HVAC&R components have been presented with emphasis on dynamic grey box modelling and it's benefits for practical applications. As with most modelling problems there is always room for improvement. Possible improvements include performing of more sophisticated experiments, and investigating different model types and modelling techniques, e.g. residual analysis for non-linear effects. However, the overall conclusion of the research work is, that grey box models, in terms of SDE's, are suitable for describing the dynamics of the considered HVAC&R components. This is due to the fact that both, the physical knowledge about the components is well established and that data is possible to control and acquire.

II

Part II

INCLUDED PAPERS

List of papers

The case studies presented in the following have been published, presented, or submitted to the following journals or conferences:

Paper A: Andersen, K.K. and Poulsen, H. (1999). Building integrated heating systems. *Building Simulation '99, Kyoto, Japan*, **1**, 105–112.

Paper B: Andersen, K.K., Palsson, O.P., Madsen, H., and Knudsen, L.H. (2000). Experimental design and setup for heat exchanger modelling. *International Journal of Heat Exchangers*, Accepted for publication.

Paper C: Andersen, K.K., Lundby, M., Madsen, H., and Paulsen, O. (1999). Identification of continuous time smooth threshold models of physical systems. *Presented at the Joint Statistical Meetings*, Baltimore, 1999.

Paper D: Andersen, K.K., Hansen, L.H., and Madsen, H. (2000). A model for the heat dynamics of a radiator. Submitted.

Paper E: Knop, O., Andersen, K.K., Madsen, H., Gregersen, N.H., and Paulsen, O. (2000). Modeling of a thermostatic valve with hysteresis effects. Submitted.

Paper F: Andersen, K.K., Madsen, H., and Hansen, L.H. (2000). Modelling the heat dynamics of a building using stochastic differential equations. *Energy and Buildings* **31** 1, 13-24.

Paper G: Andersen, K.K., and Reddy, A. (2000). The error in variable (EIV) regression approach as a means of identifying unbiased physical parameter estimates. Submitted.

A

Paper A

BUILDING INTEGRATED HEATING SYSTEMS

ABSTRACT

This paper presents the preliminary modelling perspectives of an ongoing project where a flexible simulation tool for component and system analysis of district heating consumer installations is developed. The simulation tool makes it possible to simulate district heating consumer installations containing water based central heating systems, domestic hot tap water systems, buildings as well as load predictions of the systems. The main purpose of the project is to improve the inter action of the system components, decrease the energy consumption and analyze the performance of the heating system and the relation between the building and the heating system, Building Integrated Heating System. This is done with emphasis on increasing the system performance of the district heating system. The paper presents the modelling approach with main emphasis on lumped parameter modelling using statistical methods. Some prior results are discussed and illustrated through examples, including analysis of the heat dynamics of a building and a new approach on handling discontinuities exemplified in the modelling of the transient flows pipes.

Keywords: Building Integrated Heating Systems, District Heating Systems, Lumped Parameter Modelling, Simulation studies.

1 INTRODUCTION

Since the first energy price crisis in the early seventies the energy policy in Denmark has been focused on how to reduce the consumption and the dependency of a single fuel. In recent years the energy demand in new buildings in Denmark has been reduced considerably due to new building regulations. The reduced heat demand and the growing focus on the indoor climate increases the requirements of the performance of the heating system and the interaction between building and the heating system, Building Integrated Heating System. More than 60 % of the heat consumption is based on district heating (DH) due to the national energy plans since the second energy crisis in the late seventies. Today the systems are characterized by pooled operation with combined heat and power plants mainly based on incineration, coal and natural gas. The suppliers have made a great effort on optimizing the system performance in order to minimize the costs; during the last fifteen years the optimization process has focused on the production plants and the distribution systems. Integration of the consumer installations in the process started around five years ago, as the natural last step in optimizing the overall efficiency. The design and operation of the consumer installations are improved with respect to the system performance, i.e. the installations must operate at a low temperature level with a high cooling and few peaks. The manufacturers of the components are very important in this process and the product development has changed from focusing on the component (sub optimization) to focusing on the inter action of the components (system optimization). Developing new products or improving existing is traditionally a complex and expensive task including numerous tests in the lab followed by final tests of a prototype in situ. Today a large part of these tests can be replaced by computer simulation studies. The use of dynamic models in computer simulation studies has proven to be a perceptive and practical method to analyze the performance and the control strategies in heating systems. Thence, the dynamics of the heating system can be analyzed second by second with varying heat supply from persons, machines and the sun in order to minimize the heat demand in the buildings and optimize the thermal comfort. However, the prospects from applying simulation studies in development of system components depend strongly on the accuracy and flexibility of the models used for the simulation. The paper presents the modelling approach and the idea behind a project concerning the development of a model library for simulation purposes.

2 THE MODELLING OBJECTIVE

The project is a joint project with collaborators from some of the largest manufacturers of DH components in Denmark and public research institutes. The modelling objective of the project is to establish a model library of the components in typical DH systems, sketched in Fig. 1, as a foundation for system analysis, simulation studies and product development. An essential aspect of the project, is, that the modelling of system components is continuously validated on experimental data to ensure accuracy and reliability in analysis and simulation studies. For this reasons an experimental setup are used as an empirical base for model validation.

In order to analyze the influence from the heating system on the indoor climate the models are divided into two categories, namely:

(i) Models of system components.

(ii) Models of the total system.

The models of system components are gathered in the model library and may be used for analyzing purposes and simulation studies. The user of the model library can select components from a graphical user interface (GUI) environment. Specific arrangements or setups are created using a drag-and-drop technique of the components into a worksheet, where the components are connected into a desired setup. The model library includes models of:

- Plate heat exchangers.
- Radiators.
- Thermostatic and pressure relief valves.
- Pipes and fittings.
- Pumps.
- Thermal zones and buildings.

By descending order, each class of the components cf. above, contain different models of each component. The models may vary in complexity, control strategies and user interface. Furthermore, each component may be adjusted to a specific product, i.e. physical properties such as heat capacities, resistances, sizes etc. may be controlled by the user.

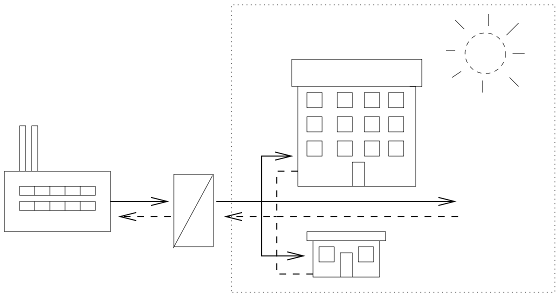


Fig. 1: Sketch of a typical DH system. The combined heat and power plant supplies the consumer via heat exchangers with hot water (solid line) for heating and domestic water. After consumption, the cooled water (dotted line) returns to the power plant.

The model interface

As sketched in Fig. 1 the total DH system besides the dynamics of the power plant is considered. The system is affected by climate variables as well as the load from the power plant, interaction between components, human activity etc. In order to achieve a well-defined user interface, only the models of the total system and entire rooms and buildings may be affected directly by the climate variables, such as the outdoor temperature, solar radiation etc. System components, such as radiators and valves, are only indirectly affected by the climate variables, i.e. via the building in which they are placed. The interface of the system components are indoor temperatures, supply- and return water temperatures as well as flow and pressure drop in pipes.

The models of system components used in the simulation tool are mainly lumped parameter models. There are several arguments for preferring and using lumped parameter models for system components compared to black box models and models in terms of partial differential equations. Firstly, lumped parameter models based on the laws of nature are directly physically interpretable, a key issue where the black box models fail. Secondly, lumped parameter models in terms of ordinary differential equations are far simpler to establish and for simulation purposes compared to distributed parameter models, i.e. models in terms of partial differential equations. The lumped parameter models may be used for 'stand alone' simulation studies or together with other models of components in simulation studies of component interactions. The models of a total system may in principle be a system of lumped parameter models, corresponding to a setup of lumped parameters models of the components in the system. However, for large scale simulations, this may not be necessary or even to complex a task. When only the relation between the input and the output signals of the system is of interest, the transfer function representation is regarded as a sufficient description of the system. The user interface is still physically interpretable, e.g. flow rates, supply temperatures and pressure drops, whereas the transfer function from input to output is not physically interpretable.

3 THE EXPERIMENTAL SETUP

In order to get an empirical base for the dynamical models of the components in the consumer installation the input-output relationship and dynamical characteristics of each component as well as the interaction between the system components are investigated in an experimental setup. This section presents a designed test rig for detailed testing of DH system components.

The experimental setup, a 100 kW test rig, corresponding to 20-30 apartments, represents a typically consumer installation in Western European countries. The heating system is separated from the DH system by a heat exchanger. The radiator system is divided into three individually zones, where hot water tanks emulate the dynamics of each zone. The load in the domestic hot water system is emulated by predictions, i.e. tap programs from a computer. The domestic hot water is produced instantaneous in a heat exchanger. In the main setup of the heating system only the load from the radiator system is of importance, not the performance of the individual radiator. Therefore, the empirical data of the radiator is assembled in a separate set up, where thermo-vision is used to verify the heat transfer from the radiator. The heat exchangers, radiators and hot water tanks are connected through pipes and fittings where pumps and valves controlled by a computer may generate pre-specified operating conditions. Furthermore the setup has the possibility for hardware-in-the-loop simulations, where new products, such as controllers and pumps, can be analyzed. The test rig are used for a variety of experiments and applications, such as model analysis and hardware-in-the-loop experiments, diagnosis and failure detection as well as data acquisition for modelling and simulation purposes.

4 THE MODELLING METHOD

The modelling of system components are based on the physical characteristics of the components and empirical data from the experimental setup. This section presents briefly the modelling method for establishing physically interpretable lumped parameter models of system components using state-of-the-art modelling techniques. The emphasis is made on the total modelling process and the modelling method is illustrated through examples.

4.1 The grey box modelling method

The lumped parameter models of the system components included in the model library are designed wrt. two main criteria:

- Accurate static and dynamic simulation performance.
- The model interface.

To comply with the design criteria, (i) and (ii), a novel modelling method, the *grey box modelling method*, see Melgaard (1994), is applied. The method has proven adequate for modelling components in DH systems, see e.g. Palsson (1993); Sejling (1993); Hansen (1997). The total modelling process may be described by a flowchart, consisting of the three stages, sketched in Fig. 2. The modelling process is characterized by that both physically insight, statistical methods and experimental data are used to identify the model structure, estimate model parameters, estimate the model uncertainty and validate the model.

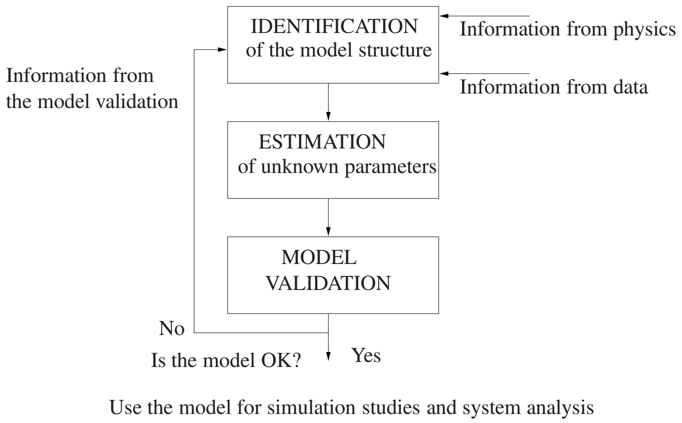


Fig. 2: *The modelling procedure.*

4.2 Identification

Using experimental data from the test rig and physically insight, such as well known hydraulic and thermodynamic relationships, a model structure may be identified, i.e. the first step in Fig. 2, in terms of stochastic differential equations:

$$d\mathbf{X}_t = \mathbf{f}\left(\mathbf{X}_t, \mathbf{U}_t, \theta, t\right) dt + \mathbf{G}(\theta, t) d\mathbf{W}_t, \qquad (1)$$

where $\mathbf{X}_t$ is a vector of system states and the vector $\mathbf{U}_t$ contains the known inputs. $\mathbf{f}$ is a known function describing the evolution of the system states. Finally, θ is a vector of parameters, $\mathbf{W}_t$ is a Wiener process and $\mathbf{G}(\theta, t)$ is a function describing how the disturbance is entering the system. There are several reasons for applying a stochastic term, i.e. the system dynamics is described in terms of stochastic differential equations, referring to Madsen and Holst (1996):

- Modelling approximations. The evolution of the system states, described by the function $\mathbf{f}$, might be an approximation to the true system.

- Unrecognized and unmodelled inputs may affect the evolution of the states.
- Measurements of the input are noise-corrupted.

Thus, a model formulation in terms of stochastic differential equations accounts for that the function $\mathbf{f}$ only is an approximation to the true evolution of the system states. The continuous time model formulation based on laws of nature ensures that the the model structure is directly physical interpretable.

4.3 Parameter estimation

Having measured some of the state variables, a state space representation can be formulated

$$d\mathbf{X}_t = \mathbf{f}\left(\mathbf{X}_t, \mathbf{U}_t, \theta, t\right) dt + \mathbf{G}(\theta, t) d\mathbf{W}_t, \tag{2}$$

$$\mathbf{Y}_t = \mathbf{h}\left(\mathbf{X}_t, \mathbf{U}_t, \theta, t\right) + \mathbf{e}_t, \tag{3}$$

where $\mathbf{Y}_t$ is a vector of the actually observed state variables. Eq. (2) is a continuous-time system equation and Eq. (3) is the discrete time observation equation. The function $\mathbf{h}$ describes the relationship between the state variables $\mathbf{X}_t$ and the measurements $\mathbf{Y}_t$. $\mathbf{e}_t$ is a vector describing the measurement noise which is assumed to be Gaussian distributed. The parameter vector θ contains the equivalent thermal and hydraulic components, i.e. capacitances, resistances etc., and is estimated by a Maximum Likelihood (ML) method. Descriptions of the ML method can be found in Ref. Melgaard (1994).

4.4 Model validation

The last stage in the modelling process concerns the model validation. The models are validated using both experimental data, statistical validation techniques and physically interpretation. The latter by comparing the model parameters, such as time constants and the estimated model parameters, $\hat{\theta}$, with the equivalent physical characteristics of the system. As a supplement to the physically interpretation of the model, statistical methods are applied. By using stochastic independent data from the test rig and statistical methods, the models are cross validated. Hereby, it is possible to determine the goodness-of-fit using statistical criteria such as residual analysis, test for model order, test for parameter significance etc. If the model validation indicates an inadequate model performance, information from the model validation can indicate how to improve the model structure, as indicated in Fig. 2.

5 MODELLING EXAMPLES

This section gives two examples of models included in the model library. The first example illustrates the modelling of a total system, namely an approach on the modelling of the heat dynamics of a building. It is argued how to extend the analysis by applying dynamic models of system components such as radiators and thus achieve a more detailed model. The second example concerns the analysis and modelling of a single component, namely the transient flow rates in pipes. Emphasis is made on a method for handling model discontinuities.

$*$ **The modeling examples are left out and given in:**
Paper F: *Modelling the heat dynamics of a building using stochastic differential equations*
and
Paper C: *Identification of continuous time smooth threshold models of physical systems*.

6 CONCLUSION

The paper described the aim and preliminary analysis of an ongoing project. The aim of the project is to develop a simulation tool for simulation purposes of DH systems in order to improve on the interaction between system components and between the building and the heating system. The modelling method has been described with emphasis on the identification process of lumped parameter models where statistical methods are applied in estimation and model validation. The simulation tool may be used for simulation of system components or for simulation of total systems. The latter may be done by creating a specific setup using models of system components. The modelling method was illustrated through two examples. The first example described the modelling of the heat dynamics in buildings. The model parameterization was found adequate using statistical methods. It was discussed how a simulation study could be more detailed by applying dynamic models of system components. The second example illustrated a new method for handling system discontinuities. The method was illustrated in the modelling of single component, namely a model of the transient flow in pipes.

Paper B

B

EXPERIMENTAL DESIGN AND SETUP FOR HEAT EXCHANGER MODELLING

ABSTRACT

This paper deals with the design of input variables for an experimental setup of a counter-flow heat exchanger arrangement. The aim of the experiment is to collect data for a succeeding modeling of the heat dynamics of the heat exchanger. The paper discuss issues relevant to the design of input sequences and a method for generating multi-level random signals is introduced. The signals are designed to excite the non-linear dynamics of the system properly and to avoid correlation between input variables. The quality and statistical properties of the data from the experiment are discussed and the simulation performance of a dynamic model is illustrated.

Keywords: Counter-flow heat exchanger; Experimental Design, Pseudo Random Number Signals, Non-linear system.

Nomenclature

b	=	slope
C	=	constant
C'	=	constant
$\mathbf{f}$	=	function describing the system dynamics
$\mathbf{h}$	=	function
h	=	heat transfer coefficient
k	=	thermal conductivity
k'	=	time lag
K'	=	constant
$\dot{m}$	=	constant
n	=	order of the PRBS sequence
n^*	=	number of compartments
Nu	=	Nusselt number
Pr	=	Prandtl number
q^{-1}	=	back shift operator
Re	=	Reynolds number

T	=	temperature
$\mathbf{U}$	=	input vector
$\mathbf{X}$	=	state vector
$\mathbf{Y}$	=	output vector

Greek symbols

θ	=	parameter vector
λ	=	base time interval in the PRBS sequence
μ	=	dynamic viscosity
τ	=	time constant
ϕ	=	filter parameter
Δt	=	sampling interval

Subscripts and superscripts

c	=	cold side
h	=	hot side
in	=	input
k	=	discrete time index
out	=	output
t	=	continuous time index
x	=	exponent
y	=	exponent

1 INTRODUCTION

The use of statistical methods is well recognized in identification of models for HVAC components. Less attention has been given to the aspect of experimental design, an important step in establishing reliable models. Basically, experimental design deals with the methods and techniques used in the design and analysis of experiments. It is the connection between the experiment and the model that the experimenter can develop from the results of the experiment (Montgomery (1991)).

In order to identify empirical relationships for models of HVAC components the quality and statistical properties of data used for model identification and for model validation should be adequate. It is important that significant variables are measured at a suitable sampling rate, that strong correlation between variables is avoided, that the data is accurate and span the modeling space well, that data are persistently exciting etc. A well-planned experimental design can help achieving this goal. For identification of dynamics systems optimal input sequences may be obtained by using Pseudo Random Number Signal (PRNS) sequences. The generation of PRNS sequences is discussed intensively in the literature for applications in engineering, see e.g. Godfrey (1980); Yarmolik and Demidenko (1988); Nowak and Veen (1993); Godfrey (1993). However, only a few studies report the use of such signals for experiments concerning HVAC equipment. Binary multi-frequency signals have been applied in identification of a large water-heated cross-flow heat exchanger in Franck and Rake (1985). Binary signals have also been applied for linear models of the heat dynamics of buildings, see Madsen and Holst (1995); Hansen (1997); Andersen et al. (2000).

This paper deals with the analysis of an experimental setup and input design of a heat exchanger arrangement. Important issues concerning the quality and statistical properties of the experimental data are discussed. In Section 2 the experimental setup is described. The experimental design is discussed in Section 3. Emphasis is put on the generation of PRNS sequences that are informative and not significantly cross-correlated. Since the heat transfer coefficient is a non-linear function of both mass flows and temperatures a method for generation of input sequences for non-linear systems is proposed. The experimental data is presented in Section 4. The statistical properties of the data are discussed and the simulation performance of a heat exchanger model is illustrated using the experimental data. Finally a conclusion with discussion is given.

2 THE EXPERIMENTAL SETUP

The considered counter-flow plate heat exchanger has the nominal power capacity of 105 kW and the convector fluid is water on both sides. The dimensions are 1174(h) $*$ 368(w) $*$ 458(l) mm. In the standard setup it consists of 31 plates, 0.6 mm stainless steel. The total active area is 4.93 m^2, the volume is 8.6 l and the total mass of the heat exchanger material is 180 kg. The heat exchanger is characterized as a counter flow type even though a few plates passages actually have parallel flow on each side. Furthermore, adding or removing plates can vary the capacity of the heat exchanger. The experimental setup is sketched in Fig. 1. The experimental setup is divided into a hot and a cold side, and the inlet and outlet temperatures as well as the mass flow are measured on each side. For subsequent modeling purposes the variables are characterized as follows:

- Controllable Inputs: The two inlet temperatures and the two mass flows.
- Measured Outputs: The two outlet temperatures.

During experiments it is possible to switch the flow within 4 levels and the temperature within 2 levels at each side, respectively. On both the hot and the cold side of the setup two tanks of water with the desired supply temperatures are acting as buffers in order to keep the supply temperatures constant. By using a PID-controller and several valves and a by-pass loop (not sketched) the flow can be controlled at 4 different levels at each side. The temperature sensors, 1.0 mm, type K (Ametek), are placed close to the inlet and outlet of the heat exchanger. The accuracy of the sensors is estimated to be within the $\pm 0.5°$C range of the actual temperature. The mass flow on both sides of the experimental setup is calculated indirectly by measuring the pressure drop

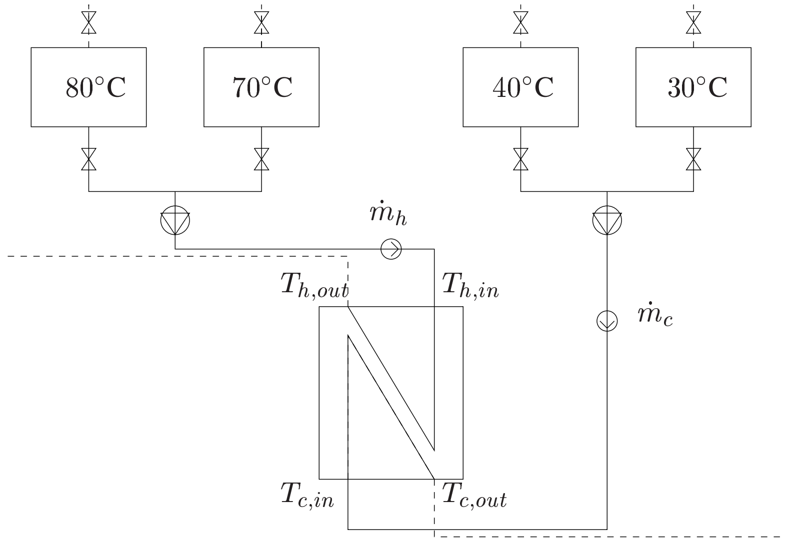


Fig. 1: Sketch of the heat experimental setup. T denotes temperature and $\dot{m}$ denotes mass flow. The index h is used for the hot side and c is used for the cold side.

while the water passes through a blender. The sensors, type Endress Hauser, Deltabar PMD 130, are estimated to measure the actual flow within an accuracy of $\pm 0.1\%$ in the considered range. The motivation for using a blender instead of a traditional flow sensor is, that time constant of the blender is much smaller than a typical flow sensor. However, a traditional flow sensor, Danfoss, is used in the experimental setup, to verify the measurements by the blenders. A converter, 6BP16-2 Backplane (National Instruments) and data acquisition system, Labview, are used to control the input variables and to collect the desired measurements during the experiment.

3 DESIGN OF INPUT VARIABLES

This section presents the design of input variables for the heat exchanger setup. Emphasis is put on how to combine physical knowledge and statistical methods in order to design the input variables.

The physical structure and interpretation of HVAC models and systems is important for analysis and applications, e.g. for system design, simulation studies, design of control strategies, fault detection and diagnosis etc., see e.g. (Jonsson et al. (1992); Jonsson and Palsson (1992, 1994); Bourdouxhe et al. (1998); Weyer et al. (2000)) for heat exchanger applications. Furthermore, it is attractive to combine empirical data and physical knowledge in order to identify and validate these models (Ljung and Glad (1994)). In the design of experiments the physical knowledge of the system can guide the selection of appropriate input levels, sampling rate and in determining what variables that should be measured etc. In the experimental design, statistical methods are useful as a supplement to the physical interpretation of the system. Statistical methods are useful in constructing input sequences that maximize the information in empirical data and in ensuring that the experiment is randomized properly and that the measurements are not significantly correlated.

In this study the input variables may be controlled. Hence, the aim of the experimental design is to generate input sequences that both excite the dynamics of the heat exchanger properly and span the modeling space well. The range and level of the input variables are determined from physical interpretation, i.e. corresponding to the maximum range of the typical operating conditions as well as levels in between. The switching of the signals is determined by Pseudo Random Number Signal (PRNS) sequences. Using PRNS sequences, as input signals in experiments with dynamic systems, is attractive because of the nice statistical properties for system identification. The PRNS have properties similar to white noise, i.e. a PRNS signal may be constructed and controlled so that it is not significantly correlated to other measured variables. This important property makes it easier to separate the impact from different variables on the system and thus establish reliable empirical relationships (Chatterjee (1991)).

The generation and statistical properties of PRNS sequences is investigated intensively in the literature, see e.g. Godfrey (1980, 1993) or Yarmolik and Demidenko (1988). The PRNS sequences are absolute stable and determined by the algorithm for pseudo random number generation. The sequences are deterministic and may be repeated from any desired implementation segment so there is no need for storage capacity. In this paper we distinguish between two types of PRNS signals, the Pseudo Random Binary Signal (PRBS) sequences and the Pseudo Random Multilevel Signal (PRMS) sequences. The PRBS has two levels and the PRMS has three or more levels. A possible method to generate a PRBS sequence is to use a shift register, sketched in Fig. 2. The PRBS is constructed by selecting the shortest (λ) and the longest time interval ($n\lambda$) where the signal is constant. The signal may switch from one level to the other at certain intervals of time, $t = 0, \lambda, 2\lambda, .., n\lambda$. The shift register determines the switching of the signal. The signal is periodic with the period $\lambda(2^n - 1)$. The shift register has feedback to the first stage, and consists of the modulo-2 sum of the logic level of the last two of the other stages. See Godfrey (1993) for other n and a detailed discussion on generation of PRBS sequences. However, concerning heat exchanger modeling, the heat transfer coefficient is known to be a non-linear function dependent upon both mass flows and temperatures. Due to the limitation of the PRBS in only consisting of two levels, PRBS may not sufficiently excite nonlinear systems, see Nowak and Veen (1993). For non-linear system identification, PRMS sequences may be more adequate, since a higher number of levels may be selected. We propose a method to construct the PRMS sequence. The method is simple and straightforward and quite similar to the above method for generation of PRBS sequences. Using two different shift registers, two independent PRBS sequences are generated. The four possible combinations of the two PRBS signals determine the four-stage PRMS sequence, i.e. $0 \uplus 0 = 1,\ 0 \uplus 1 = 2,\ 1 \uplus 0 = 3$ and $1 \uplus 1 = 4$.

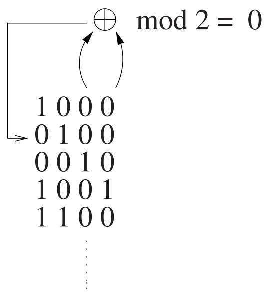


Fig. 2: Generation of a PRBS sequence, exemplified using a shift register circuit with $n = 4$. In the first step a zero is generated (0+0 mod 2 = 0). The generated number is put back in the shift register. In the second step a new zero is generated (0+0 mod 2 = 0) followed by a one in the next step (1+0 mod 2 = 1) etc.

Since the mass flows could be controlled within 4 levels and the temperatures each within 2 levels the temperature signal has been determined by PRBS sequences and the mass flow by a PRMS sequences. The selected input levels are listed in Tab. 1. Based on steady state calculations provided by the manufacturer, the time constants, τ, were assumed to be in the interval 4-25 seconds. These time constants are listed in Tab. 2. The calculated time constants indicate that the sampling time should be selected not larger than 2 seconds since it would not possible to estimate the fast dynamics using a lower sampling rate (Ljung and Glad (1994)). Due to the uncertainty of the estimate of the time constant, a sampling time of 1 second is found appropriate. The time constant of the sensors is found to be much smaller $[0.15 - 0.25]$ s and assumed not to influence on the accuracy of the measurements. The shift registers and time intervals, used for generation of the PRBS sequences controlling the inlet temperatures, are listed in Tab. 3. The PRMS sequences controlling the mass flows are listed in Tab. 4. The time intervals, λ_i, are selected with respect to the calculated time constants, i.e. the smallest time interval is larger than expected time constants.

4 THE EXPRIMENTAL DATA

Several different PRNS sequences were applied as input variables to the experimental to make sure that both fast and slow excitations could be identified and simulated by a dynamic model. Each time series had a sampling frequency of 1 Hz and the length of an experimental run was 3500 s.

Table 1: The input level of the mass flows and inlet temperatures.

Level	$T_{h,\text{in}}$ [°C]	$\dot{m}_h$ [kg/s]	$T_{c,\text{in}}$ [°C]	$\dot{m}_c$ [kg/s]
1	70	0.31	30	0.32
2	80	0.37	40	0.39
3	-	0.50	-	0.52
4	-	0.71	-	0.77

Table 2: Estimated time constants.

	Step (from $\rightarrow$ to)	τ ($T_{h,out}$)	τ ($T_{c,out}$)	Units
$\dot{m}_h$	0.72 $\rightarrow$ 0.36 (kg/s)	9.5	13.8	s
$\dot{m}_c$	0.84 $\rightarrow$ 0.42 (kg/s)	25.0	9.9	s
$T_{h,in}$	80 $\rightarrow$ 70 (°C)	-	4.0	s
$T_{c,in}$	40 $\rightarrow$ 30 (°C)	5.0	-	s

Table 3: PRBS sequence controlling the inlet temperatures.

Input	Time interval $[s]$	Shift Register
T_h	$\lambda = 27$	$[0101001010110]$
T_c	$\lambda = 41$	$[100110]$

A time series from an experimental run is shown in Fig. 3. It should be stressed that the results presented in the following are representative for the other experiments as well. In general, the input variables are fairly close to the deterministic PRNS sequences. More precisely, the mass flows are very accurate while the inlet temperatures deviate slightly wrt. the input design. For the time series in Fig. 3 the estimated correlation coefficients between the flow and temperature signals are listed in Tab. 5. It is seen that none of the signals are strongly correlated with each other. The data are not considered noisy and the time series contain non or only very few outliers. In Fig. 4 a scatter plot of the outlet temperature against the inlet temperature and the mass flow on the

Table 4: The two PRMS sequence controlling the mass flows. The four stage PRMS sequence is determined by the four possible combinations of PRBS1 and PRBS2 and by PRBS3 and PRBS4, respectively.

Input	Time interval $[s]$	Shift Register
$\dot{m}_h$	$\lambda_1 = 9$	PRBS1: $[0101001]$
	$\lambda_2 = 17$	PRBS2: $[11000101011]$
$\dot{m}_c$	$\lambda_1 = 5$	PRBS3: $[0101001]$
	$\lambda_2 = 12$	PRBS4: $[011001010]$

Table 5: Estimated correlation coefficients between input variables.

$\rho(\dot{m}_h, \dot{m}_c)$	$\rho(\dot{m}_h, T_h)$	$\rho(\dot{m}_h, T_c)$	$\rho(\dot{m}_c, T_h)$	$\rho(\dot{m}_c, T_c)$	$\rho(T_h, T_c)$
-0.05	0.08	0.32	-0.23	0.02	-0.06

hot side is shown. It is seen that the data spans well and that the data are not strongly correlated. The same holds for the cold side (not plotted) and for other combinations of input and output variables.

The range and variations of the data ensures that the dynamics of the considered heat exchanger may be validated for a wide range of operating conditions. A dynamic model, reported in Jonsson et al. (1992); Jonsson and Palsson (1992, 1994) to perform well for a small counter-flow heat exchanger has been applied. The model is a lumped parameter model where both physical

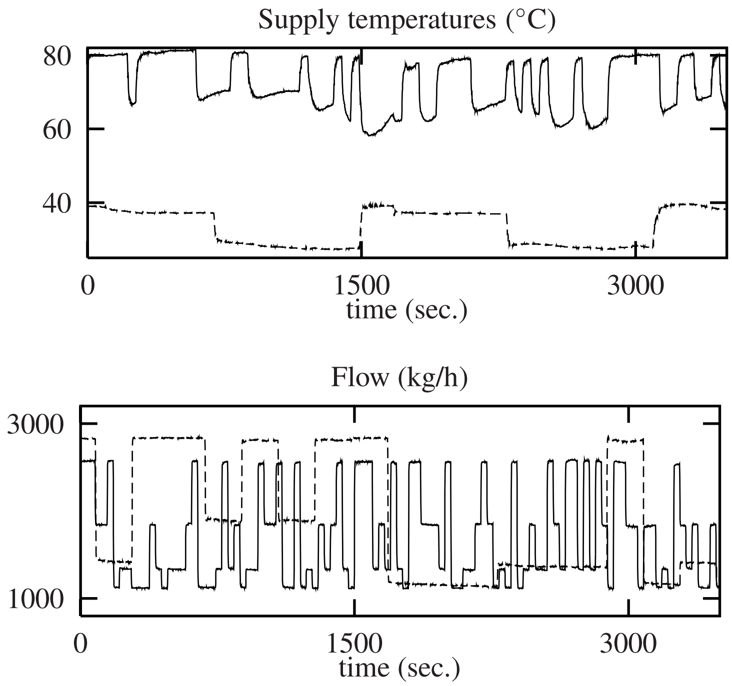


Fig. 3: Measured data: cold side (-.-) dotted line, hot side (–) solid line.

knowledge and statistical methods have been used in establishing an adequate model structure. Referring to Palsson (1993) it is assumed that the heat transfer to the surroundings is negligible and that there is no heat conduction in the fluids. The distributed character of the heat transfer can be approximated by introducing compartments in the system (lumping). As sketched in Fig. 5, the heat exchanger is divided into compartments. $\dot{m}_h(t)$ and $\dot{m}_c(t)$ are the mass flows on the hot and cold side, respectively. Similarly, $T_{h,in}(t)$, $T_{h,out}(t)$ and $T_{c,in}(t)$, $T_{c,out}(t)$ are the inlet and outlet temperatures on each side. $T_{h,i}(t)$, $T_{c,i}(t)$ are temperatures in each compartment, $T_{m,i}(t)$ are the temperature of the intermediate metal layer compartments.

The differential equations which describes the heat dynamics of the heat exchanger are formed by considering the energy balance equations Eq. (1-2):

$$\frac{d(\text{Heat stored})}{dt} = \sum \text{Power in} - \sum \text{Power out}, \tag{1}$$

$$\Rightarrow \quad C_i \frac{dT_i}{dt} = \sum \Phi_{in} - \sum \Phi_{out}. \tag{2}$$

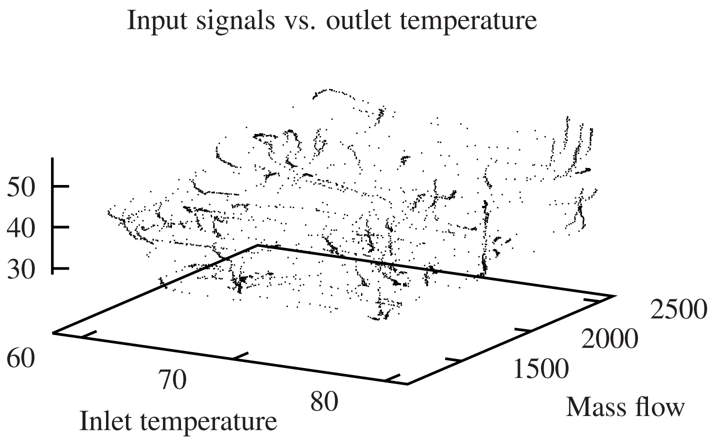


Fig. 4: Distribution of the outlet temperature as a function of the inlet temperature and the mass flow (hot side).

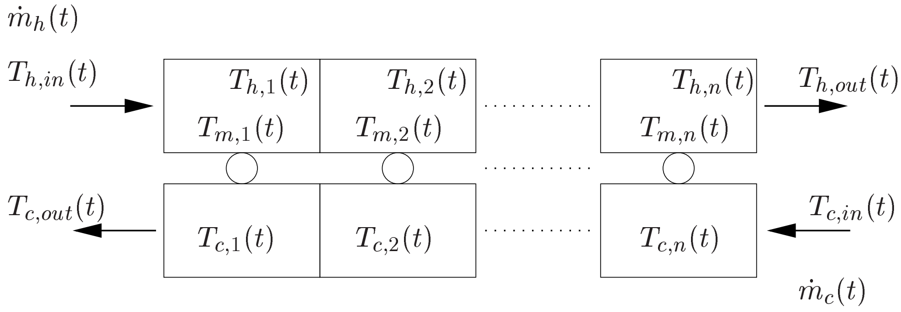


Fig. 5: The compartmental counter flow heat exchanger model.

Dividing the heat exchanger into n equal compartments, implies for the compartment of number i, the following differential equations:

$$\frac{M_h}{n} c_h \frac{dT_{h,i}(t)}{dt} = \dot{m}_h(t) c_h [T_{h,i-1}(t) - T_{h,i}(t)] - \frac{A_h}{n} U \Delta T_i(t), \quad (3)$$

$$\frac{M_c}{n} c_c \frac{dT_{c,i}(t)}{dt} = \dot{m}_c(t) c_c [T_{c,i-1}(t) - T_{c,i}(t)] + \frac{A_c}{n} U \Delta T_i(t). \quad (4)$$

U is the overall coefficient for the heat transfer from one flow to the other. Here it is written as an independent parameter though it is actually a function of mass flows and temperatures. $\Delta T_i(t)$ is the driving force for the heat transfer representing the temperature difference between each pair of hot and cold compartments. The indexes h and c denotes the hot and cold fluid, Mc denotes the heat capacity on each side and A the surface area of each compartment. Dividing the heat exchanger into n compartments yields a $2n$ order model. In Eq. (3-4) the heat capacity of the metal between the liquids are neglected. However, the extension of including the heat capacity of the metal is straight forward, see Sejling (1993). The driving force for the heat transfer represented by ΔT_i are modeled as a weighted average of the temperatures on each side:

$$\Delta T_i(t) = \frac{1}{2}[T_{h,i-1}(t) - T_{h,i}(t)] - \frac{1}{2}[T_{c,i-1}(t) - T_{c,i}(t)]. \tag{5}$$

$$\tag{6}$$

It is well known, that the heat transfer between the fluid and the intermediate metal is a non-linear function of both temperature and mass flow. In Jonsson and Palsson (1992, 1994) the heat transfer coefficient, $h(T, \dot{m})$, are modeled by physical interpretation of empirical relations and statistical methods, by simplifying (applying) the empirical relation:

$$\text{Nu} = C\text{Pr}^x\text{Re}^y, \tag{7}$$

and inserting expressions for Nu, Pr and Re, it is shown in e.g. Jonsson and Palsson (1994) that the heat transfer can be give by:

$$h(T, \dot{m}) = K \underbrace{\mu(T)^{(x-y)}k(T)^{(1-x)}}_{i} \underbrace{\dot{m}^y}_{ii}, \tag{8}$$

where $\mu(T)$ is the dynamic viscosity and $k(T)$ is the thermal conductivity both with dependency of the temperature, where $\dot{m}$ is the mass flow, K is a constant and y and x are exponents depending on the type of flow. In Eq. (8) (i) is temperature dependent and (ii) mass dependent. In Jonsson and Palsson (1994) the approximations Eq. (9) and (10) are found appropriate:

$$h(T, \dot{m}) = K^{'}(1 + bT)\dot{m}^y, \tag{9}$$

$$h(\dot{m}) = C^{'}\dot{m}^y, \tag{10}$$

where $C^{'}$ and $K^{'}$ are constants and b is a slope. In Eq. (9) the heat transfer is both mass flow and temperature dependent, but in Eq. (10) is only mass flow dependent. Finally, neglecting the resistance of the metal between the mass flows, the heat transfer coefficient U can be formulated as:

$$\frac{1}{U} = \frac{1}{h_h} + \frac{1}{h_c}. \tag{11}$$

The heat transfer coefficient (11) may now be included in the models of the heat exchanger. In summary, the lumped model describes the outlet temperatures as a function of the inlet temperatures and the mass flows.

By lumping the model into 3 sections at each side the model consists of 6 differential equations. The model has been applied to the experimental data. Different time series have been used for parameter estimation and others for model validation in order to validate the simulation performance of the model. A cross validation of the model using an is plotted in Fig. 6. It is seen that the model is able to simulate the outlet temperatures with a very high accuracy.

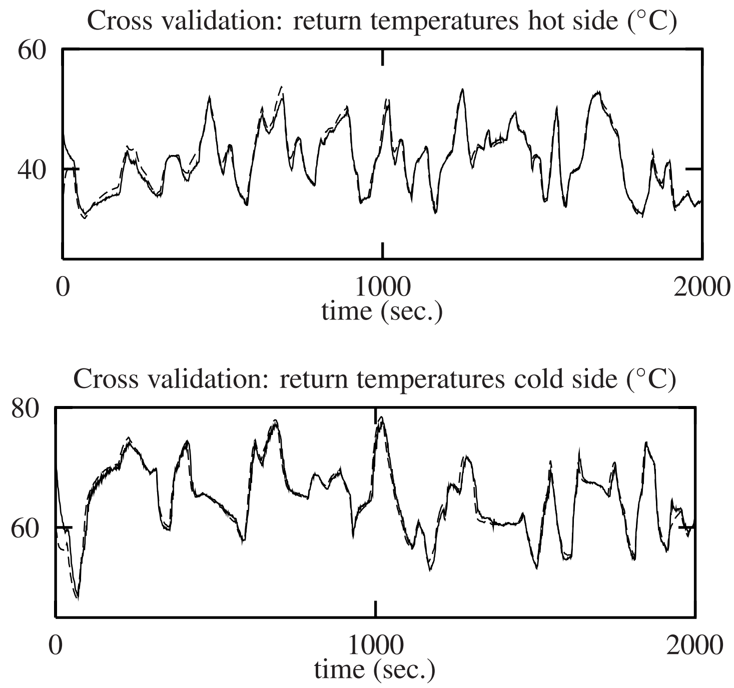


Fig. 6: Measured data (-.-) dotted line, Simulated data (–) solid line.

5 CONCLUSION

An experimental setup and the design of input variables for a counter-flow heat exchanger arrangement has been presented. The aim of the experiment was to collect data for the succeeding modeling of the dynamics of the heat exchanger. The heat exchanger was described and the construction of the experimental setup was discussed. Physical interpretation and statistical methods were used to plan the experiment and to design the input sequences for the heat exchanger setup.

It was argued that in order to identify empirical relationships for models of HVAC components the quality and statistical properties of data used for model identification and for model validation should be adequate. To achieve this goal PRNS sequences were applied to generate input sequences and a new method for generating such sequences was described. These signals were applied and the experimental data were presented. The quality of the data was analyzed and it was argued that by applying the PRNS sequences

(i) The range and distribution of the data ensures that the model is valid for a wide range of operating conditions.

(ii) The input signals are not strongly correlated it is possible to identify and estimate the impact from both mass flows and temperatures on the dynamics of the heat exchanger.

The generation of PRNS sequences is straight forward and has proven useful in experiments of dynamic systems. Furthermore, the generation of PRMS sequences may be extended to have more levels than applied in this study. For successful experiments it is required that the signals can be controlled at the desired levels. As the number of levels increases so do the need for longer and more complex experiments in order to obtain data with adequate statistical properties. Thus, the price for good quality data may be a more complex experimental setup. However, even simple experimental design may as discussed is this study can greatly improve the quality of data and make a subsequent model identification easier.

Paper C

IDENTIFICATION OF CONTINUOUS TIME SMOOTH THRESHOLD MODELS OF PHYSICAL SYSTEMS

C

ABSTRACT

This paper presents a technique for identifying accurate and physically interpretable threshold models of physical systems by the means of physical knowledge and statistical methods. The modeling approach is illustrated by considering typical components in a water based central heating system. The model formulation is based on the laws of nature and statistical analysis of experimental data. It is shown that in many cases the dynamic characteristics of such a physical system depend strongly upon the state of the system, e.g. laminar or turbulent flow in a pipe, and it may be difficult to obtain models that are valid for different excitation of the input variables. A solution to this problem is suggested by introducing smooth threshold models. The models are formulated in continuous time by lumping the system and using stochastic differential equations. The continuous time formulation ensures that the model is directly physical interpretable. A Maximum Likelihood method is used to estimate the model parameters and the model states. Statistical methods are applied to verify that the model provides a reasonable description of the system.

Keywords: Smooth threshold models, Lumped parameter models, Stochastic differential equations, Maximum Likelihood method.

1 INTRODUCTION

Many physical systems are distributed systems for which the dynamics in principle may be described by physical laws and thus by partial differential equations (PDE). In many applications a model formulation in terms of PDE's is difficult to obtain since the system is to complex and may not be known completely. Under such conditions a black box model (Box and Jenkins (1970)) may be appropriate, even if no prior information about the process is at hand. However, in cases where some prior knowledge about the process is present, it may be desirable to incorporate this physical knowledge into the model structure. This modeling approach, where physical knowledge is accounted for and used together with the information from data is referred to as the grey box modeling approach, see e.g. Tulleken (1992); Melgaard (1994); Ljung and Glad (1994).

Different modeling approaches may be applied to the same problem or physical system. One example is in the modeling of Heating, Ventilation and Air Conditioning (HVAC) components, such as heat exchangers and air conditioners, where different modeling methods can be applied, see Bourdouxhe et al. (1998). However, when analyzing HVAC components, e.g. in the design of control strategies or in the developing of new products it is often desirable to incorporate physical knowledge at hand such as the laws of heat and mass transfer, or relevant physical parameters such as heat capacities and resistances. In such situations the grey box approach can be a powerful tool. The dynamical characteristics of a given system may be described approximately by lumping the partial differential equations, i.e. obtain a model formulation in terms of ordinary differential equations (ODE). For more complex systems the ODE's may be obtained by considering mass and heat balances. By using experimental data and including stochastic terms, the model can be formulated as a system of stochastic differential equations. An advantage using this approach is that it is directly physical interpretable. Furthermore, the use of statistical methods permits a description of model uncertainties and provides a very useful tool in the model validation, i.e. to determine whether or not the model structure is reasonable.

However, for some HVAC components the dynamical characteristics of the system depend strongly upon the state of the system. Examples of such model discontinuities are valves shutting down, phase transition in heat exchangers and the transition from laminar to turbulent flow in a pipe. Therefore it may be difficult to obtain a total model which is valid for very different excitation of input variables by lumped parameter models. A commonly used approach is to restrict the model wrt. the input signals, i.e. the model is only valid for certain ranges of the input variables, e.g. a pipe model which is only valid for turbulent flow.

In this paper an approach towards a total model for the flow in a pipe is presented. The basic idea is to identify two sub-models, one for turbulent and one for laminar flow, respectively, and then combining these sub-models using a smooth threshold. The threshold, which is a function of the flow, is then used to determine which sub-model that should be used at some time instant. A statistical framework is used to estimate the model parameters and to validate the model.

2 THE SMOOTH THRESHOLD FUNCTION

This section presents the idea of using a smooth threshold in a continuous time physical model formulated in terms of ODE's. The use of discrete time threshold models are well known from non-linear time series analysis, see Tong (1990). In Bottin and Chaté (1998) a two dimensional discrete time threshold model is used to model the transition to turbulence flow (in plane Couette flow). Although the transition to turbulence flow is deterministic it can be argued that the transition is only defined statistically. The results in Bottin and Chaté (1998) is consistent with a threshold defined in terms of an equivalent probabilistic process. In this study the hyperbolic function, $\tanh(x),\ x \in \mathbb{R}$, is used as the foundation of a smooth threshold function in order to model the transition between laminar and turbulent flow. The smooth threshold function can be interpreted as a weight function, $w(x) = (\tanh(x) + 1)/2$, shown in Fig. 2. The weight function, $w(x) \in [0\ 1]$, can be interpreted as a probability of the flow being laminar or turbulent. This makes it possible to obtain a total

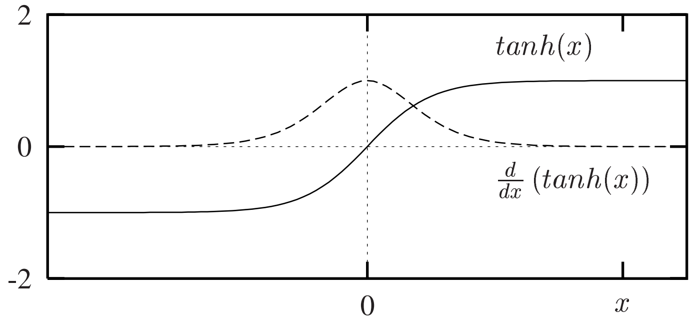


Fig. 1: The threshold function, $\tanh(x)$ (solid line) and the derivative, $\frac{d}{dt}(\tanh(x)) = 1 - \tanh(x)^2$ (dotted line).

model, $\mathbf{f}$, in terms of ODE's on the form:

$$\mathbf{f} = (1 - w(\mathbf{x}))\, \mathbf{f}_1(\mathbf{x}, \mathbf{u}, \mathbf{\Theta}, t) + w(\mathbf{x})\mathbf{f}_2(\mathbf{x}, \mathbf{u}, \mathbf{\Theta}, t),$$

where $\mathbf{f}_1$ is a sub-model valid for laminar flow and $\mathbf{f}_2$ is a sub-model valid for turbulent flow. The vector $\mathbf{u}$ contains the known inputs and $\mathbf{\Theta}$ is a vector of parameters. Finally, the weight function, $w(\mathbf{x})$ is a function of the system states, $\mathbf{x}$. The derivative of the weight function is continuous since $\frac{d}{dt}(\tanh(x)) = 1 - \tanh(x)^2$, is continuous $\forall x \in \mathbb{R}$. Hereby, it should be possible to formulate a total model which is continuous $\forall x \in \mathbb{R}$. The threshold function and its first order derivative is shown in Fig. 1. Now $w(x)$ is modified slightly:

$$w(x) = (\tanh(a(x - b)) + 1)/2, \tag{2}$$

where a and b are scale and location parameters to be estimated. It is seen, that the threshold is now centered on the x-axis to a point b and the shape is formed by the parameter a. In the following section, a total model for the transient flow in a pipe based on the threshold model in Eq. (1), will be proposed.

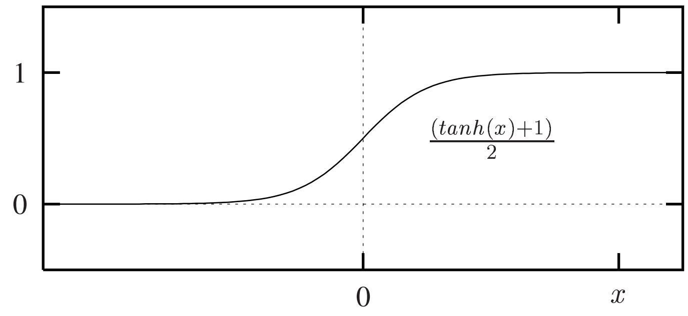


Fig. 2: The weight function, $w(x) = \frac{(\tanh(x)+1)}{2}$.

3 DYNAMIC MODELLING OF PIPES

This section presents a lumped parameter model of the flow in a pipe, where Eq. (2) discussed in Section 2 is applied as a smooth threshold between two sub-models, one for laminar flow and one for turbulent flow. Using the second law of Newton it easily follows that the flow, q, of an incompressible fluid out a pipe with the pressure drop, Δp_p, can be determined by the first order non-linear differential equation:

$$\frac{dq}{dt} = \frac{A}{l}\left(\frac{\Delta p_p}{\rho} - (\psi + \frac{\lambda l}{d})\frac{q^2}{2A^2}\right), \tag{3}$$

where A is the cross-section area of the pipe with diameter d and length l, ρ is the density of the water, ψ is the minor loss coefficient and λ is the friction factor. However, Eq. (3) is not valid for both laminar and turbulent flow, since the friction factor depends on the actual flow. For laminar flow the friction factor can be calculated using:

$$\lambda = \frac{64}{\text{Re}}, \quad \text{Re} = \frac{ud}{v}, \tag{4}$$

where Re is the Reynolds number, v is the kinematic viscosity of the fluid and u is the average speed of flow. For fully developed turbulent flow, the Colebrook formula (Fox and McDonald (1985)) holds:

$$\frac{1}{\sqrt{\lambda}} = -2\log_{10}\left(\frac{\gamma}{3.7d} + \frac{2.51}{\text{Re}\sqrt{\lambda}}\right), \tag{5}$$

where γ is the absolute pipe roughness. Eq. (5) is transcendental and has to be solved by iteration. The turbulent region can be divided into smaller regions. For example, when the Reynolds number is less than $1 \cdot 10^5$ and the pipe is smooth, the friction factor can be described by the Blasius equation (Fox and McDonald (1985)):

$$\lambda = \frac{0.316}{\text{Re}^{0.25}}. \tag{6}$$

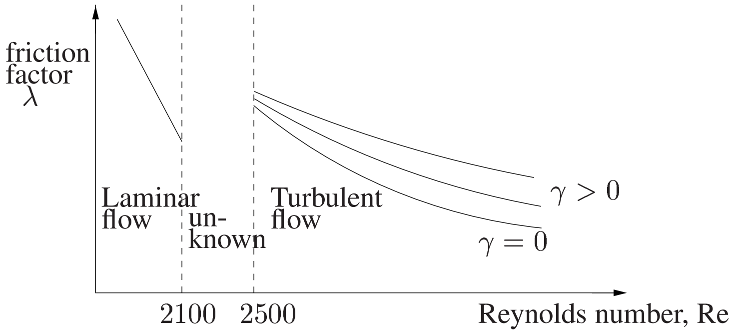


Fig. 3: Sketch of the friction factor as a function of the Reynolds number.

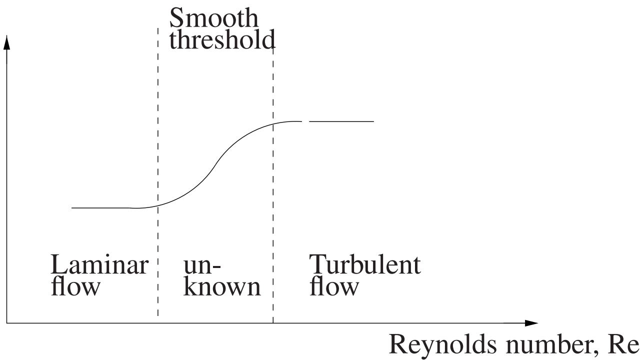


Fig. 4: Sketch of the weight function as a function of the Reynolds number.

Depending on the physical properties of the actual pipe, such as the roughness of the pipe, the flow in a pipe is known to be laminar for approximately Re ≤ 2100 and turbulent for Re ≥ 2500. In the region between, $2100 \leq$ Re ≤ 2500, the flow is neither purely laminar nor purely turbulent, but a combination. In this interval, no expression for the friction factor, λ, exist. The friction factor as a function of the Reynolds number is sketched in Fig. 3. In the following the friction factor at laminar and turbulent flow are denoted λ_l and λ_t, respectively. The sub-models describing the flow are written on the form Eq. (3) and denoted $\mathbf{f}_1$ for laminar flow and $\mathbf{f}_2$ for turbulent flow. Furthermore, the weight function, $w(\mathbf{x})$ is a function of the Reynolds number, i.e. $w(\mathbf{x}) = w(Re)$, where:

$$w(Re) = (\tanh(a(Re - b)) + 1)/2, \tag{7}$$

where a and b are parameters to be estimated. It should be noted, that the location parameter b is expected to be in the interval $[2100; 2500]$. To summarize, the model can be written:

$$\frac{dq}{dt} = (1 - w(Re))\,\mathbf{f}_1(q, \Delta P, \Theta, t) + (w(Re))\,\mathbf{f}_2(q, \Delta P, \Theta, t),$$

where the vector Θ contains the model parameters:

$$\Theta = [A,\ l,\ \rho,\ ...\ a,\ b]. \tag{9}$$

4 PARAMETER ESTIMATION

In this section the parameter estimation method is discussed briefly. The section follows Melgaard (1994) closely. The total model is written:

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\Theta}, t)\,dt, \tag{10}$$

where $\mathbf{x}_t$ is a vector of the system states and the vector $\mathbf{u}_t$ contains the known inputs. Finally, $\boldsymbol{\Theta}$ is the vector of the parameters to be estimated. The model, Eq. (10), provides a deterministic description of the evolution in time of the states of the system. It is obvious that any description of the form (10) gives only an approximation of the evolution of the true system. In order to use Eq. (10) as a description of the true variation of the states a stochastic term is included in Eq. (10) leading to the following stochastic differential equation:

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\Theta}, t)\,dt + \mathbf{g}(\Theta, t)d\mathbf{w}_t, \tag{11}$$

where $\mathbf{w}_t$ is a Wiener process and $\mathbf{g}(\Theta, t)$ is a function describing how the disturbance is entering the system. Having measured some function of the state variables, a state space representation can be formulated:

$$d\mathbf{x}_t = \mathbf{f}\left(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\Theta}, t\right) dt + \mathbf{g}(\Theta, t) d\mathbf{w}_t, \tag{12}$$

$$\mathbf{y}_t = \mathbf{h}\left(\mathbf{x}_t, \mathbf{u}_t, \boldsymbol{\Theta}, t\right) + \mathbf{e}_t, \tag{13}$$

where $\mathbf{y}_t$ is a vector of the actually observed variables. Eq. (12) is a continuous-time system equation and Eq. (13) is the discrete time observation equation. The function $\mathbf{h}$ describes the relationship between the state variables $\mathbf{x}_t$ and the measurements $\mathbf{y}_t$. Finally $\mathbf{e}_t$ is a vector describing the measurement noise which is assumed to be Gaussian distributed. The model parameters are found using discrete time observations of the pressure drop and flow, respectively. Let $\mathbf{y} = y^k$ and $\mathbf{u} = u^k$ denote the output/input sets of observations:

$$y^k = [y_k, y_{k-1}, y_{k-1}, ..., y_1, y_0], \tag{14}$$

$$u^k = [u_k, u_{k-1}, u_{k-1}, ..., u_1, u_0], \tag{15}$$

and let Θ be a vector of the N parameters to be estimated. The problem of estimation is a matter of how to use the information contained in the data Eqs. (14-15) to select a proper value $\hat{\Theta}$. In this study a Maximum Likelihood (ML) method is used to estimate the parameter vector $\hat{\Theta}$. The conditioned likelihood function is the joint probability density of all the observations assuming that the parameters are known (u^k omitted for convenience):

$$\begin{aligned} \mathcal{L}(\Theta, y^k) = p(y^k|\Theta) &= p(y_k|y^{k-1}, \Theta)p(y^{k-1}, \Theta) \\ &= \left(\prod_{i=1}^{k} p(y_i|y^{i-1}, \Theta)\right), \end{aligned} \tag{16}$$

where Bayes rule $p(A, B) = p(A|B)p(B)$ is applied. It is now assumed that the sequence of innovations, ϵ_k, are zero-mean, independent stochastic variables with the probability density function $p(\epsilon(\Theta))$. Then, Eq. (16) can be formulated:

$$\mathcal{L}(\Theta, y^k) = \left(\prod_{i=1}^{k} p(\epsilon_i(\Theta))\right). \tag{17}$$

The ML estimates, $\hat{\Theta}_{ML}$, is found by minimizing the function $-\log L(\Theta, y_k)$:

$$\hat{\Theta}_{ML} = \arg\min_{\Theta} \left(-\sum_{i=1}^{k} \log p(\epsilon_i(\Theta))\right). \tag{18}$$

When the prediction errors are assumed to be Gaussian with zero mean, and covariance matrix $R_k(\Theta)$, then

$$-\log\mathcal{L}(\Theta, y^k) = \frac{1}{2}\sum_{i=1}^{k}(\epsilon_i^{\top} R_k^{-1}\epsilon_i + \log \det R_i + s \log 2\pi). \tag{19}$$

where s is the dimension of y. It should be noted that if the assumption that the prediction errors are Gaussian distributed, the method can be considered as a prediction error method. The estimated model is validated using physical interpretation and statistical methods, such as prediction errors analysis, simulation studies and cross validation.

5 THE EXPERIMENTAL SETUP

In order to collect data for the modeling and validation process of the transient flow an experimental setup was built at the Danish Technological Institute. The experimental setup is sketched in Fig. 5. During the experiment the pressure drop ΔP $[Pa]$, the flow q $[m^3/s]$ as well as the inlet and the outlet temperatures $[°C]$ were measured. Using a Moody's diagram the dimensions of the pipe were selected to be 12.0/0.01 (length/diameter$[m]$) in order to be able to generate both laminar and turbulent flow. The pipe material was made of soft copper and winded up to avoid sharp edges. Finally, a computer was used to control the flow as well as the inlet temperature.

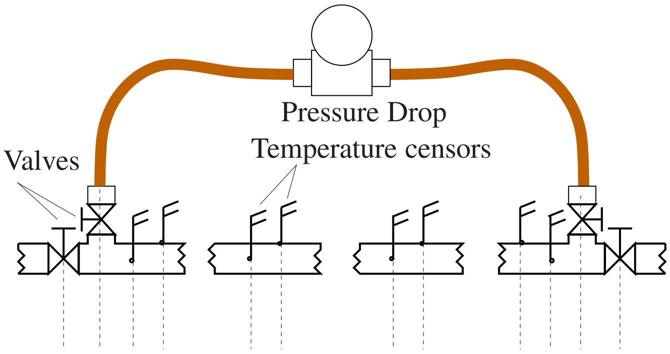


Fig. 5: Sketch of the experimental setup.

6 THE DATA

The preliminary analysis of the experimental setup yield that the transient flow will appear for Reynolds numbers in the interval $[2100\ ;\ 2500]$. This corresponds to a flow in the interval $[65\ ;\ 75]\ [kg/h]$ when the supply temperature

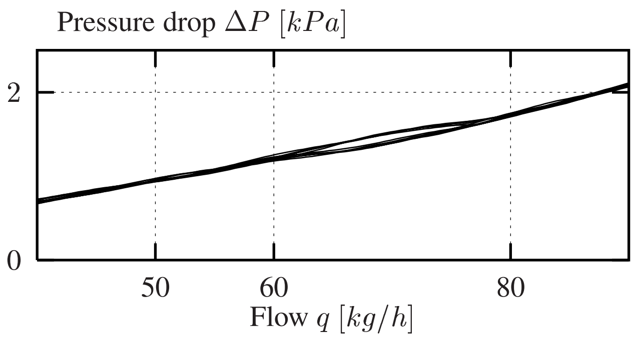


Fig. 6: Phase diagram of the pressure drop ΔP vs. the flow q.

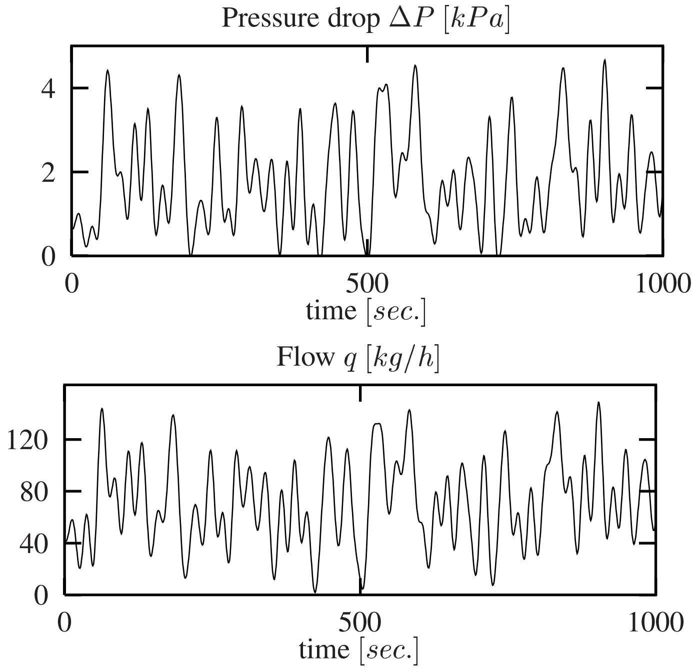


Fig. 7: The measured pressure drop and corresponding flow in a type (iii) experiment.

of the water, T_s is 30 $[^\circ C]$. This assumption is supported by the phase diagram in Fig. 6, obtained from a preliminary steady state experiment. Furthermore, a hysteresis effect is seen in this interval, corresponding to that the transition from laminar to turbulent flow occurs at a higher flow rate than the transition from turbulent to laminar flow, respectively.

Several controlled dynamical experiments were performed using different levels of flow and constant supply temperature. The experiments can be categorized into the following three groups:

(i) q : 10 20 30 40 $[kg/h]$ and T_s : 30 $[^\circ C]$.

(ii) q : 90 110 130 150 $[kg/h]$ and T_s : 30 $[^\circ C]$.

(iii) q : 20 60 100 140 $[kg/h]$ and T_s : 30 $[^\circ C]$.

where (i) is a dynamical experiment in the laminar region, (ii) is a dynamical experiment in the turbulent region and finally the dynamical experiment (iii) is covering both regions, including the transient region. The sample frequency of the experiments, cf. above, was 1 Hz. In Fig. 7 a sample time series of the pressure drop and the flow is shown for the type (iii) experiment, i.e. both laminar, transient and turbulent flow occurs.

7 RESULTS

The sub-models for laminar and turbulent flow were estimated using type (i) and (ii) experimental data. Eq. (4) and Eq. (6) were applied for λ_l and λ_t, respectively, and the first order non-linear model Eq. (3) was found adequate in both regions. Finally, the total model was written on the form Eq. (1) and type (iii) experimental data was used for parameter estimation. Since most of the model parameters, such as the pipe diameter, the length and the minor loss constant were either known or estimates directly in the sub-models, only the (most interesting) parameters a, b, the initial state q_0 and the noise terms had to be estimated. To ensure that the parameter estimates converges towards the global supremum for the likelihood function, several estimations using different initial values of the parameters were performed. In Fig. 8 the estimated autocorrelation function for the model residuals are shown. It is concluded that the residuals can be considered as being Gaussian white noise and thus

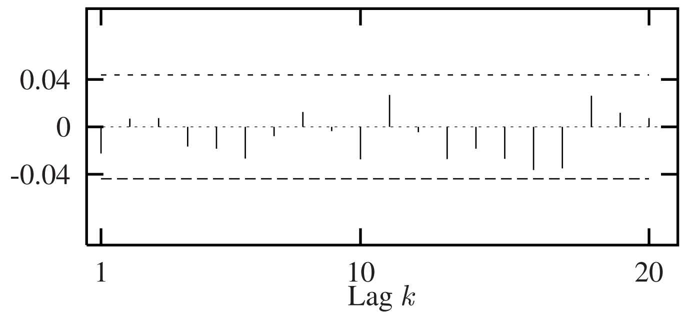


Fig. 8: The estimated autocorrelation function for the residuals.

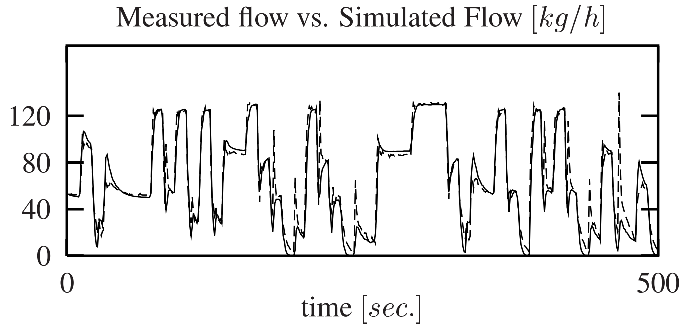


Fig. 9: Cross validation. The model is used to simulate the system using a stochastic independent time series. Simulated (solid line) and measured (dotted line).

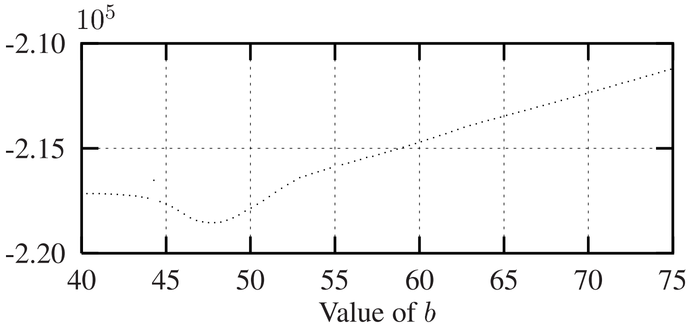


Fig. 10: The profile likelihood for the parameter b. The optimal value is found for $b = 48$ $[kg/h]$.

that the model describes the dynamics of the system. Hence, statistically it is concluded that a reasonable model is found. The reasonable performance of the model is supported by the cross validation study, shown in Fig. 9. The optimum value of the parameter b, which is the value where $w(\mathbf{x}) = 0.5$, is found to be $b = 48$ $[kg/h]$. The profile likelihood is shown in Fig. 10. The estimated value of b is smaller than first expected compared with the phase diagram in Fig. 6, which was obtained for steady state experiments. It can be argued, as in Bottin and Chaté (1998), that all turbulence has to vanish before the flow is purely laminar. Thus, during the dynamical experiment the flow has

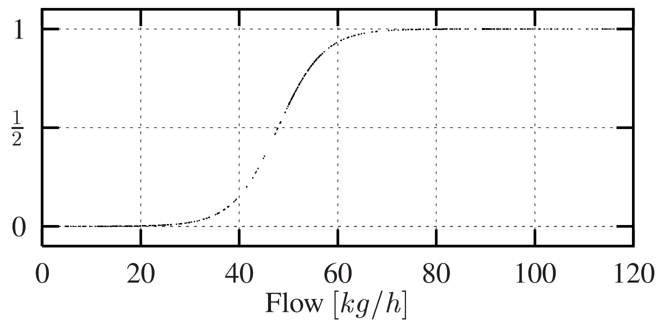


Fig. 11: The estimated weight function $w(\mathbf{x})$ as a function of the flow q.

to decrease below 40 $[kg/h]$ before the flow can be regarded as being purely laminar. The estimated weight function is shown in Fig. 11. It is seen that the transient flow is modeled to be in the region $[40\ ;\ 60]$ $[kg/h]$.

8 CONCLUSION

A smooth threshold model of the flow in a pipe has been presented. The model structure was formed using physical knowledge. A statistical framework was applied to estimate the model parameters including the shape and the location of the threshold using experimental data. Prediction error analysis and physical interpretation of the system were used to determine that the model structure was reasonable. The estimated center of the threshold was at 48 $[kg/h]$. The transient flow was modeled to be in the region $[40\ ;\ 60]$ $[kg/h]$. The estimated smooth threshold seems very reasonable compared to the physical interpretation of the system. However, the advantage of using a smooth threshold model would be more clear if the sub-models were even more different. Furthermore, the model has to be modified if the temperature of the water changes, i.e. the location and scaling parameter then should also depend upon temperature. Since the threshold may be sharpened it follows that also more/less sharp thresholds could be applied. Only one threshold was considered in this work but more than one threshold could easily be applied, provided that the model is identifiable.

Paper D

A MODEL FOR THE HEAT DYNAMICS OF A RADIATOR

D

ABSTRACT

A model of the heat dynamics of a panel radiator is presented. The model is obtained from physical reasoning and experimental data. The model is lumped into two compartments and the size of these compartments depends on the actual flow. By introduction of variable compartment sizes, it is shown that it is possible to keep the model order low and obtain a very accurate model when compared to empirical data. The model is formulated in continuous time, using stochastic differential equations which make it directly physical interpretable. Statistical methods are used for parameter estimation and model validation. The simulation performance of the model is illustrated.

Keywords: Radiators, heat dynamics, lumped parameter modelling, grey box modelling method.

1 INTRODUCTION

There is a need for development of dynamic models of water based central heating systems in order to analyze and interpret the physical interaction between system components, such as radiators, valves, heat exchangers etc. Such models can be useful e.g. in the design of control systems containing feedback loops. A radiator is a typical component in a water based central heating system. It is a distributed system for which the dynamics in principle may be described by physical laws concerning mass, energy and momentum. In this paper a model describing the heat dynamics of a radiator is proposed combining physical knowledge and statistical methods.

Different modeling techniques have been applied to model the dynamics of components in water based central heating systems, see e.g. Jonsson (1990); Sejling (1993); Hansen (1997); Bourdouxhe et al. (1998). The dynamics of the components and the interactions between these can be described by considering the thermal and hydraulic characteristics of the system. However, the amount of literature on dynamic modelling of radiators is sparse.

Various steady state models of radiators can be found in the literature, see e.g. Benonysson (1991). Such steady state models are in Madsen and Holst (1995) and in Andersen et al. (2000) found adequate in deriving a total model of the heat dynamics in a building. A possible approach towards obtaining a dynamic model is to lump the radiator into compartments, and formulate a dynamic model based upon laws of heat transfer. The energy balance of each compartment is then expressed in terms of ordinary differential equations. This approach is used in Paulsen and Grundtoft (1985), where the radiator is divided into N equally sized compartments. The energy balance of each compartment is calculated using 'handbook' parameters. Based on experiments and simulations it is found, that the best performance is obtained when selecting N in the range 10-20 compartments. A drawback using this approach is that the model order becomes rather high. The large number of compartments is found necessary to avoid numeric diffusion, i.e., the temperature 'moves' faster than the flow, which is a non-physical property. In Hansen (1997), a stochastic modelling approach, the grey box modelling method, is applied to model the heat dynamics of a radiator. The grey box modeling method is characterized by using both physical knowledge and information from data in determining an adequate model structure. The proposed model is lumped into a number of equally sized compartments. The model structure is identified using physical laws as well as measurements from an experimental setup. Statistical methods are used to determine the model parameterization and the optimal number of compartments is found to be two for the particular experimental setup. However, applying statistical test indicates that the model parameterization is inadequate. It is concluded in Hansen (1997) that the model structure has to be improved.

In this paper a new model parameterization is proposed using the grey box modeling approach. The grey box modelling approach has earlier been applied for various other components in water based central heating systems, see e.g. Jonsson et al. (1992); Jonsson and Palsson (1992, 1994); Hansen (1997); Andersen et al. (2000). For detailed discussion on the grey box modeling method, see Tulleken (1992, 1993); Melgaard (1994); Ljung and Glad (1994). The model presented in this paper is also a lumped parameter model, but the basic idea is that the compartmental sizes may vary, depending on the flow. This corresponds better to the distribution of the temperature of the water and radiator material, which can be observed using thermo-vision (Adunka (1992)). The heat dynamics of the compartments are modelled by considering the energy balances. The heat transfer to the air surrounding the radiator surface is modelled using a similar dependency. Finally, the variation of the section sizes is modelled. The suggested model describes the heat dynamics when excited by different input signals, such as changes in the flow controlled by a thermostatic valve and variations in the ambient air temperature. Statistical methods are used to determine estimate the significant model parameters and to validate the model structure.

2 THE MODELLING APPROACH

In this section the applied modelling approach is briefly described. The grey box modelling method is selected since it is a suitable approach in combining physical knowledge and experimental data in the modelling process, see e.g. Tulleken (1992, 1993); Melgaard (1994); Ljung and Glad (1994). Physical knowledge is used in forming the model structure, but the modelling is kept in a stochastic framework and statistical methods are applied in determining the parameters of the model. The model is formulated in continuous time and expressed in terms of stochastic differential equations. Both the states and the parameters of the model are estimated using a Maximum Likelihood (ML) method. Finally, the model is validated using both statistical methods and physical knowledge. This modelling approach may be preferable compared to a deterministic modelling approach because the physical system to be modelled is only known partly and model approximations are inevitable. The grey box modelling method complies with these difficulties, using statistical methods to determine whether or not the model parameterization is adequate, and if not, how to improve the model.

To model the impact from the assumingly most important variables on the heat dynamics of the radiator, well known thermo dynamic relationships are used, and formulated in terms of a system of ordinary differential equations:

$$\frac{d\mathbf{X}}{dt} = \mathbf{f}\left(\mathbf{X}, \mathbf{U}, \theta, t\right), \tag{1}$$

where $\mathbf{X}$ is a vector of system states and the vector $\mathbf{U}$ contains the known inputs. Finally, θ is a vector of parameters. The lumped model, Eq. (1), provides a deterministic description of the evolution in time of the states of the system. It is obvious that any description of the form (1) only yields an approximation of the evolution of the true system. In order to use Eq. (1) as the foundation for a description of the true variation of the states a stochastic term is included in Eq. (1) leading to the following system of stochastic differential equations:

$$\frac{d\mathbf{X}}{dt} = \mathbf{f}\left(\mathbf{X}, \mathbf{U}, \theta, t\right) + \mathbf{G}(\theta, t)d\mathbf{W}, \tag{2}$$

where $\mathbf{W}$ is a Wiener process with $E[d\mathbf{W}]^2 = dt$ and $\mathbf{G}(\theta, t)$ is a function describing how the disturbance is entering the system. For a discussion on stochastic differential equations, see e.g. Kloeden et al. (1997); Øksendal (1985); Risken (1996). There are several reasons for introducing a noise term in Eq. (1), referring to Melgaard (1994):

- Modeling approximations. The system described by Eq. (1) might be an approximation to the true system.
- Unrecognized and unmodelled inputs may affect the evolution of the states.
- Measurements of the input are noise-corrupted.

Having measured some function of the state variables, a state space representation can be formulated:

$$\frac{d\mathbf{X}}{dt} = \mathbf{f}\left(\mathbf{X}, \mathbf{U}, \theta, t\right) + \mathbf{G}(\theta, t)d\mathbf{W}, \tag{3}$$

$$\mathbf{Y} = \mathbf{h}\left(\mathbf{X}, \mathbf{U}, \theta, t\right) + \mathbf{e}_t, \tag{4}$$

where $\mathbf{Y}$ is a vector of the actually observed variables. Eq. (3) is a continuous-time system equation while Eq. (4) is a discrete time observation equation. The function $\mathbf{h}$ describes the relationship between the state variables $\mathbf{X}$ and the measurements $\mathbf{Y}$ and $\mathbf{e}_t$ is a vector describing the measurement noise which is assumed to be Gaussian distributed. The parameter vector θ contains the equivalent thermal components i.e. capacitances, resistances etc., and is estimated by a ML method. Descriptions of the ML method can be found in Melgaard (1994); Nielsen et al. (2000). Later on in the paper, examples of a states space description, eqs. (3-4), will be given.

Finally, the presented model is validated using physical interpretation and statistical methods, such as residual analysis, simulation studies and cross validation.

3 THE RADIATOR

The considered panel radiator has a nominal power of 395 W and has the dimensions shown in Fig. 1. The volume of the radiator is approximately 5 l. A thermostatic valve is placed in the inlet of the radiator. The radiator is used in an experimental setup in a low energy test building at the Technical University of Denmark. The experimental setup consists of the test building with a water based central heating system where the radiator is acting as a component. The experimental setup is discussed in Hansen (1997). The variables measured during experiments are listed in Tab. 1. Data from these experiments are used to estimate and validate the model in the following.

Table 1: Measured variables during the experiment.

Symbol		Unit
T_f	The inlet water temperature	$^\circ$C
T_r	The outlet water temperature	$^\circ$C
q	The flow	l/h
T_a	The ambient air temperature	$^\circ$C

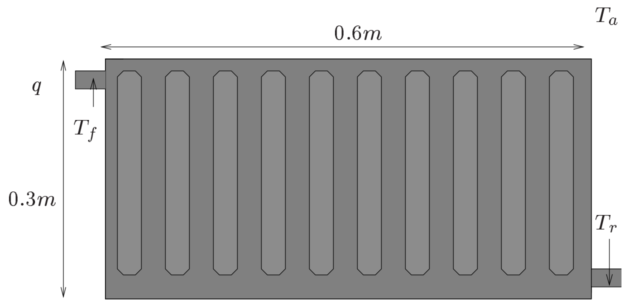


Fig. 1: The radiator.

4 FORMULATION OF A MODEL

By using thermo-vision it is possible to observe the temperature distribution on the surface of the radiator, see Adunka (1992). Two observations regarding radiators are:

(i) The flow forces the water through the radiator. Simultaneously, there is a natural drift of the water, since cold water is more heavy than hot water and this causes the water to split into some sort of boundary layers as sketched in Fig. 2.

(ii) The temperature distribution on the radiator surface, and thus the heat transfer to the surroundings, depends upon the flow. The water will cool down while passing through the radiator, i.e. the main heat transfer to the surroundings takes place at the top of the radiator.

The following paragraph describes the model formulation and the resulting differential equations, based on the observations (i) and (ii). A radiator is a distributed system, a fact which prescribes that the heat transfer has to be described by partial differential equations. However, the distributed character, cf. (i), of the heat transfer can be approximated by introducing compartments in the system (lumping). Furthermore the problem is simplified by restricting the system, i.e. only the heat capacities of the radiator and the surrounding air are

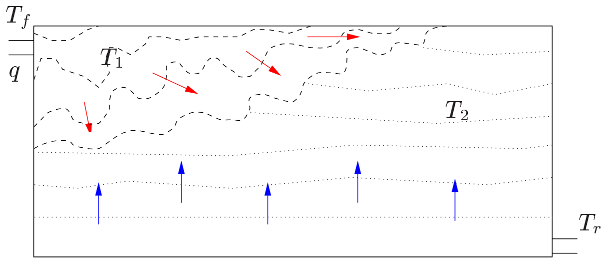


Fig. 2: A sketch of the assumed drift of water in the radiator.

modelled. An interpretation of the modelling approach is sketched in Fig. 3. The radiator is lumped into 2 compartments where the temperature is assumed constant. The energy balance of each element is considered and this includes a description of the surrounding air. Finally, the size of each compartment depends on the flow.

Referring to Fig. 3 it is assumed that the water in each compartment is perfectly mixed and that the fluid properties are constant. The assumption that the radiator can be lumped into perfectly mixed compartments can be be validated using steady state measurements Paulsen and Grundtoft (1985). The assumption that the fluid properties are constant is reasonable since the variations in the inlet

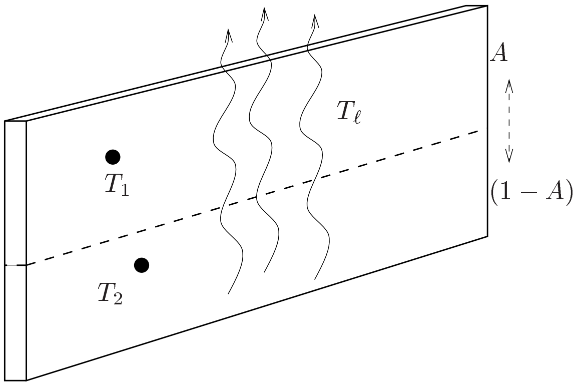


Fig. 3: The lumped model. The radiator is divided into two compartments. The size of the compartments depend on the flow.

temperature of such a radiator is typically in the interval 40-60 °C. Thus the density and the specific heat of the water are estimated to vary relatively less than 1 and 0.1 %, respectively. Finally, it is assumed that the water and the radiator material are incompressible. The temperatures of the two compartments are denoted T_1 and T_2, respectively. The temperature of the surrounding air is denoted T_ℓ. The functional area of the upper compartment is denoted A and the lower compartment $(1 - A)$, where $A \in]0\ 1[$. Furthermore it is assumed, that the temperature of the metal in each compartment equals the temperature of the water, i.e. T_1 and T_2, respectively. The energy balance is written in terms of ordinary differential equations. In general as:

$$\frac{dQ_{stored}}{dt} = \sum \Phi_{in} - \sum \Phi_{out}, \tag{5}$$

where $\Phi_{in/out}$ denotes the input/output power, and dQ/dt denotes the time derivative of the energy $Q(t)$.

The heat capacity of each section is denoted C_1 and C_2, respectively. By introducing the heat capacity of the entire radiator, C_r, implies that:

$$C_1 = C_r A \quad \text{and} \quad C_2 = C_r(1 - A). \tag{6}$$

The heat balance Eq. (5) of the upper compartment becomes:

$$C_r A \frac{dT_1}{dt} = kq(T_f - T_1) - kq(T_1 - T_2) - H_r A(T_1 - T_\ell)^n, \tag{7}$$

where k is the product of the specific heat capacity, the density of the water and the radiator material. H_r is the heat transfer coefficient from the radiator to the surrounding air with temperature T_ℓ and n denotes the radiator exponent. Similar, the heat balance of the lower section becomes:

$$C_r(1 - A)\frac{dT_2}{dt} = kq(T_1 - T_2) - H_r(1 - A)(T_2 - T_\ell)^n, \tag{8}$$

where the temperature T_2 is taken to be equal to the return temperature. Considering the heat transfer from the radiator to the room air, an approximation to the energy balance of the surrounding air becomes:

$$C_\ell \frac{dT_\ell}{dt} = H_r A (T_1 - T_\ell)^n + H_r (1 - A)(T_2 - T_\ell)^n - B_\ell (T_\ell - T_a), \quad (9)$$

where B_ℓ is the heat transmission coefficient to the ambient air with temperature T_a. Finally the variation in time of the section areas need to be modelled. The upper section area A, and thus the lower section area $(1 - A)$, will be modelled as:

$$\frac{dA}{dt} = aq - bA, \quad (10)$$

where a and b are parameters to be estimated, and q is the flow. Equation (10) states that when the flow increases the area of the upper compartment increase as well. The latter term on the right hand side of (10) leads to an exponential decrease of the upper area if the flow at some time instant becomes small. This corresponds to the physical interpretation of the drift of the hot water in the radiator. Furthermore it should be noted, that since the flow is restricted, constraints are put on the parameters a and b so that the area A is restricted to be in the interval $A \in]0\ 1[$. Equation 10 is a very simple expression for how the compartment sizes may vary. If the statistical analysis indicates that this description is not sufficient, Eq. 10 can be extended to include the fluid properties and the structure of the radiator as well as the flow. The appropriateness of Eq. 10 will be examined in the subsequent section.

In summary, the heat dynamics of the radiator is modelled by the four coupled non-linear ordinary differential equations, Eq. (7-10). The temperature in each compartment, the surrounding air temperature and the compartment sizes are taken as state variables. The flow, the inlet temperature and the ambient air temperature are input variables. By introducing additive noise, and regarding the states and parameters as stochastic variables, the set of equations, Eq. (7-10), forms the system equation, Eq. (3). Having measured the outlet temperature, T_r, the observation equation, Eq. (4) becomes:

$$T_r = T_2 + e_1. \quad (11)$$

The total model is given in the appendix, eqs. (12-16).

The system equation $d\mathbf{X}_t = \mathbf{f}(\mathbf{X}(t), \mathbf{U}(t), \theta, t)dt + \mathbf{G}(\theta, t)d\mathbf{W}_t$:

$$\frac{dT_1}{dt} = \frac{kq}{C_r A}(T_f - 2T_1 + T_2) - \frac{H_r}{C_r}(T_1 - T_\ell)^n + d\omega_1 \tag{12}$$

$$\frac{dT_2}{dt} = \frac{kq}{C_r(1-A)}(T_1 - T_2) - \frac{H_r}{C_r}(T_2 - T_\ell)^n + d\omega_2 \tag{13}$$

$$\frac{dT_\ell}{dt} = \frac{H_r}{C_\ell}A(T_1 - T_\ell)^n + \frac{H_r}{C_\ell}(1-A)(T_2 - T_\ell)^n - \frac{B_\ell}{C_\ell}(T_\ell - T_a) + d\omega_3 \tag{14}$$

$$\frac{dA}{dt} = aq - bA + d\omega_4. \tag{15}$$

The observation equation $\mathbf{Y}_k = \mathbf{h}(\mathbf{X}(t), \mathbf{U}(t), \theta, t) + \mathbf{e}(t)$:

$$T_r = T_2 + e_t. \tag{16}$$

5 RESULTS

In this section the results from the modelling procedure are given. The results are based on 1000 observations with sampling time of 1 min. from the experiment described in section 3. The state variables and the parameters have been estimated using the ML method. The model is validated using statistical methods to support the physical interpretation of the model states. Finally, the simulation performance of the system is illustrated and the model is validated on a stochastic independent time series.

In Tab. 2 the estimated model parameters and their standard deviation are presented. The parameter estimates, which are directly physically interpretable, seems reasonable validated to handbook parameters. Since the model parameters are directly physical interpretable they can be used in simulation studies, e.g. as design parameters for control strategies. Furthermore, the model states may also be given a physical interpretation. The water and the air temperatures are shown in Fig. 4. The estimated values of the states agree with the physical

Table 2: Estimates of model parameters.

Symbol	Estimate	Std. deviation	Unit
C_r	3.01	0.05	kJ/K
H_r	49.06	4.36	$kJ/(Kh)$
C_ℓ	0.90	0.08	kJ/K
B_ℓ	9.73	0.42	$kJ/(Kh)$
a	6.4 10^{-2}	6.44 10^{-3}	$m^2/(l/h)$
b	1.4 10^{-2}	6.24 10^{-3}	$1/h$

interpretation, namely that the water is cooling down and the surrounding air is heated. Thus, the model may be used to simulate the unobserved states of the radiator.

The estimated compartment sizes and the measured flow are shown in Fig. 5. The variations in the flow are inducted by a thermostatic valve (slow variation) as well as a pump switching between two operating points. It is seen that the

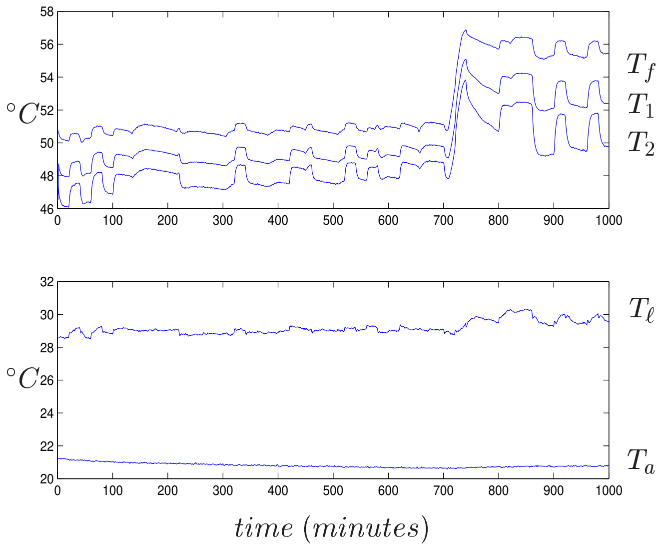


Fig. 4: Top: Water temperatures. Bottom: Air temperatures.

compartment sizes as expected depend on the variations in the flow.

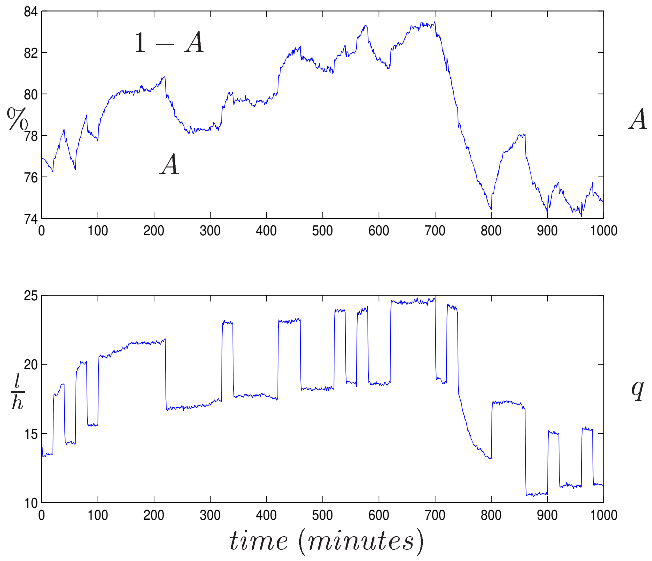


Fig. 5: Top: Estimated sizes of the compartments. Bottom: Measured flow.

It is found that the model is able to predict the return temperature with a high accuracy, i.e. the prediction errors, or the model residuals, are very small. The predicted return temperature are not shown, since it is not possible to distinguish the predicted temperature from the true observations in a plot. In Fig. 6 the cumulated periodogram of the residuals is shown. It is concluded, that the model residuals can be accepted as being white noise, i.e. the model describes all the observed dynamics of the system. The simulation performance of the model is presented in Fig. 7. It is seen that the model is able to simulate the system very accurately and the simulation performance is therefore considered satisfactory. Finally, a cross validation on an independent time series has been performed and is presented in Fig. 8. It is seen that the model agrees nicely with the measured time series. From this statistical analysis it is concluded that it is not necessary to extend the model further in order to model the heat dynamics of the radiator.

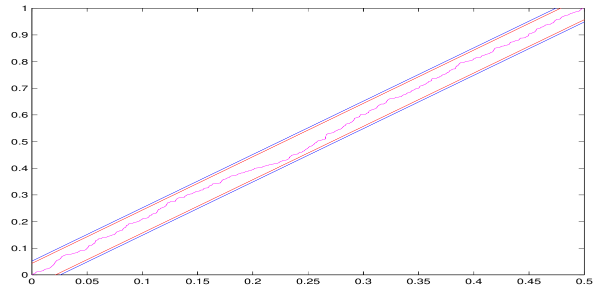


Fig. 6: Normalized accumulated periodogram

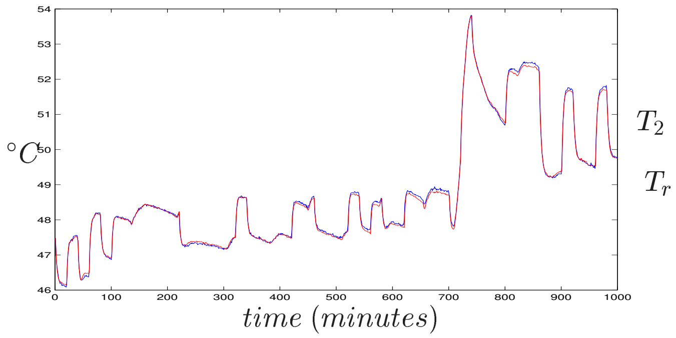


Fig. 7: Simulation of the return temperature T_r.

6 CONCLUSION

A lumped parameter model for the heat dynamics of a radiator has been proposed. Both physical knowledge and experimental data have been used in formulating the model structure. It was argued, that the compartment sizes in a lumped parameter model should be a function of the actual flow. The variation in the size of the compartments could be interpreted as how the flow influences the temperature distribution and the drift of the water in the radiator, a fact which can be observed using thermo-vision. Introducing the compartment size as a state variable, it was possible to model the variation of the compartment sizes. This further implied, that since the lumping of the radiator was dependent upon the flow, it was possible to reduce the model order signifi-

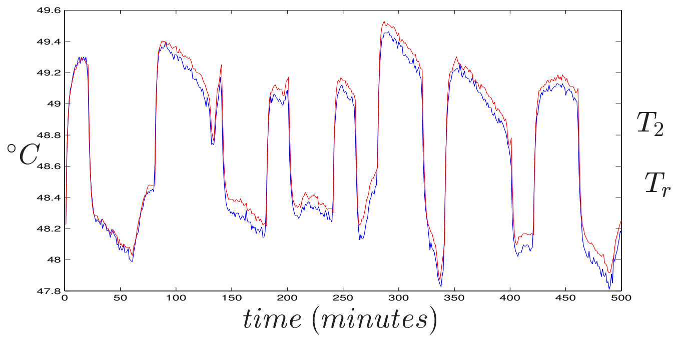


Fig. 8: Cross validation.

cantly compared to lumped parameter models with fixed compartment sizes. The varying sizes of the compartments were modelled in terms of a simple first order differential equation, thus the heat dynamics of two compartments of the radiator and the surrounding air was considered. The proposed model was non-linear and formulated in continuous time with directly physical interpretable parameters. Statistical methods are used in estimating and validating the model. It was shown, that the state variables correspond nicely with experimental data from the system. Furthermore, the prediction performance of the model was satisfactory and the residuals could be accepted as being Gaussian white noise. Simulation and cross validation studies also showed a nice agreement with experimental data. Although the model is simple there was found no statistical evidence that the model structure should be extended. It must be emphasized that the model is developed for one particular radiator type under typical operating conditions. It still has to be examined if the model is valid for other types of radiators and under extreme operating conditions as well. An interesting aspect would be to validate the variations of the compartment sizes using a time series of thermo-vision snapshots. However, it is believed that introducing a state variable as a time varying variable could reduce model order in other lumped parameter models of physical systems.

NOMENCLATURE

Symbol		Unit
A	Area of upper compartment	$\% \, m^2$
B_ℓ	Conductivity coeff.	$kJ/(Kh)$
C_ℓ	Heat capacity of ambient air	kJ/K
C_r	Heat capacity of radiator	kJ/K
H_r	Conductivity coeff.	$kJ/(Kh)$
T_a	Ambient air temperature	$^\circ C$
T_f	Supply water temperature	$^\circ C$
T_1	Temperature of upper compartment	$^\circ C$
T_2	Temperature of lower compartment	$^\circ C$
T_ℓ	Ambient air temperatur	$^\circ C$
T_r	Return temperature	$^\circ C$
a	Radiator parameter	$m^2/(l/h)$
b	Radiator parameter	l/h
$dw_{(\cdot)}$	System error	$^\circ C$
$e_{(\cdot)}$	Measurement error	$^\circ C$
k	Constant	
n	Radiator exponent	$-$
q	Flow	l/h

Paper E

MODELLING OF A THERMOSTATIC VALVE WITH HYSTERESIS EFFECTS

E

ABSTRACT

This paper presents a model of a thermostatic valve based on first principles and where the hysteresis effect is modeled using an adaptive model for friction compensation. The grey box modeling approach is applied, i.e. both physical interpretation and statistical methods are used to build and validate the suggested model. The model performance is illustrated using empirical data and issues concerning the modeling of hysteresis in valves are discussed.

Keywords: Thermostatic Valve, Hysteresis effects, Friction, Grey box modeling, Statistical methods.

1 INTRODUCTION

The use of mathematical models as a means for simulation of heating systems has improved both the understanding and the capability of controlling such systems. The importance of developing and applying adequate mathematical models has increased since the focal point in product development has changed from focusing on the component (sub optimization) to focusing on the interactions of the components (system optimization). Today, a large part of product development can be done by computer simulation as a replacement for the traditional in situ testing, which for complex systems is both expensive and time consuming.

The introduction of thermostatic valves in residential buildings during the last decades has significantly reduced the energy consumption and improved the thermal comfort. Today the thermostatic valve is a standard component in most water based heating systems. In this study a model for a thermostatic valve is presented. A major difficulty in the modeling thermostatic valves is the highly non-linear effect from hysteresis. We propose to use a dynamic model based on physical interpretation and first principles (force balances) in combination with an adaptive friction model to deal with this problem. The resulting model is characterized as a grey box model due to the combination of using first principles and empirical methods. The unknown terms of the model are estimated using statistical methods and experimental data, whereas the model is validated using both physical interpretation and statistical methods.

2 THE MODELLING OF THERMOSTATIC VALVES

The thermostatic valve considered in this study is typically used for temperature control of small hot water cylinders (e.g. storage tanks) or heat exchangers in radiator heating systems. It is a self-acting thermostatic valve, where a gas ampoule in the thermostat will increase/decrease the pressure on the valve cone as the surrounding temperature increases/decreases and force the valve to close/open.

Since the thermostatic valve is common in heating system installations it is important to develop adequate models, e.g. for design of control strategies and product development. There exist extensive literature on mathematical models of both thermostats and valves separately as well as models of thermostatic valves, see e.g. Hansen (1997); Bourdouxhe et al. (1998); Zou et al. (1999); Kayihan (2000). The purpose of these models is primarily to optimize the control of heating and cooling systems. The mathematical modeling, however, is not straight-forward due to strong non-linear factors such as hysteresis. The purpose of this study is to develop a model that is directly physical interpretable and able to handle the non-linearities.

To obtain an adequate physical model we propose to use the grey box modeling approach in modeling the thermostatic valve. The grey box modeling approach is characterized by using both prior knowledge, such as physical laws, as well as information from empirical data in the identification procedure. The advantage of this approach is that the model may be given a physical interpretation as well as non-linear effects may be more effectively described compared to a black box model. On the other hand, the use of empirical data provides for more adequate models compared to a pure deterministic or white-box model. The grey box method will not be discussed in detail in this paper. For a discussion on grey box modeling approaches, see Tulleken (1992); Ljung and

Glad (1994); Melgaard (1994). For grey box applications of components in heating systems, see e.g. Jonsson and Palsson (1994); Weyer et al. (2000) for heat exchanger modeling, Madsen and Holst (1995) for modeling building heat dynamics, and Gordon et al. (2000) for modeling of air conditioners.

3 THE MODEL FORMULATION

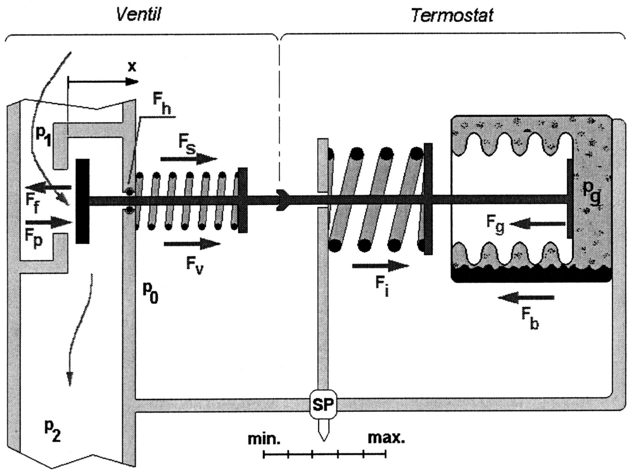
In this section a model of the valve will be proposed. The model is based on physical interpretation and first principles. A simplified sketch of the thermostatic valve is shown in Fig. 1, emphasizing the mechanical parts. Arrows indicate the direction of the forces, which influences on the mechanical parts. Note that only forces that are considered significant are sketched. These are in particular the force from the thermostat (via the spring), friction in the valve and the forces due to the differential pressure as well as the velocity of the liquid through the valve opening.

3.1 Formulation of the force balance

In the following each of the considered forces that is assumed to affect the valve will be formulated in mathematical terms. Based on this a force balance will be proposed. It should be emphasized that the sign of the various forces are indicated by the arrows in Fig. 1, assuming that the forces are positive when considered from the right. The force balance is given as the sum of the considered forces:

$$-F_f + F_p + F_s - F_g + F_v + F_i - F_b \pm F_h = 0, \tag{1}$$

where the terms are explained in Fig. 1. In the following each term will be considered more closely. The flow forces, F_f can be interpreted as a spring force $K_f \Delta p_k$ that are trying to close the valve:



p_0	Atmospheric pressure
p_1	Pressure at inlet
p_2	Pressure at outlet
SP	Fix point
x	Valve opening
F_f	Flow forces
F_p	Differential pressure at valve cone
F_s	Differential pressure at stem/spindle
F_h	Friction
F_v	Valve spring
F_i	Setting spring
F_g	Differential pressure at bellows
F_b	Bellows spring

Fig. 1: Sketch of the thermostatic valve. The shaded area surrounding p_q indicates the gas chamber, while the arrows at p_1 and p_2 indicates the upstream and downstream pressure. The nomenclature for the various parameters are given in the table below the sketch.

$$F_f = K_f \cdot x \cdot \Delta p_k, \tag{2}$$

where Δp_k is the differential pressure at the cone, K_f is a constant and x is the valve opening. The forces due to the differential pressure are defined from the area on which they are affecting:

$$F_p = A_k(p_1 - p_2) = A_k \cdot \Delta p_k. \tag{3}$$
$$F_s = A_s(p_1 - p_0) = A_s \cdot \Delta p_s. \tag{4}$$
$$F_g = A_b(p_g - p_0) = A_b \cdot \Delta p_b(T_g). \tag{5}$$

A_k, A_s and A_b are the area of the cone, the spindle and the bellows, respectively. The differential pressure is dependent of the temperature of the gas inside the thermostat, T_g, which determines the pressure of the gas p_b. The spring forces, i.e. the forces from the valve spring, the setting spring and the bellows spring, are defined from Hooks law, which is a linear function of the valve position:

$$F_v = F_{v0} - C_v \cdot x. \tag{6}$$
$$F_i = F_{i0} - C_i \cdot x. \tag{7}$$
$$F_b = F_{b0} + C_b \cdot x. \tag{8}$$

$F_{v0,i0,b0}$ and $C_{v,i,b}$ are constants. The force due to friction, $\pm F_h$, is working against the velocity of the valve and hereby a cause to hysteresis. Since it is not straight-forward to define this friction force, an adaptive model will be introduced in Section 3.2 to model for the friction force. If the valve opening is set equal to $x = 0$ in (10), four fix-point values (SP) are given as:

$$A_k \cdot \Delta p_{k,SP} + A_s \cdot \Delta p_{s,SP} + F_{v0} + F_{i,SP} - A_b \cdot \Delta p_b(T_{g,SP}) - F_{b0} = 0. \tag{9}$$

In this expression the force from fiction is neglected. The fix-point values are constant and depend only on the actual calibration (set-points) of the thermostatic valve. Now, the force balance can be written as Eq. (10):

$$
\begin{aligned}
-K_f \cdot x \cdot \Delta p_k + A_k \cdot \Delta p_k + A_s \cdot \Delta p_s - A_b \cdot \Delta p_b(T_g) \\
+\big(F_{v0} - C_v \cdot x\big) + \big(F_{i0} - C_i \cdot x\big) - \big(F_{b0} + C_b \cdot x\big) \pm F_h = 0.
\end{aligned} \tag{10}
$$

By isolating the valve position x we obtain:

$$
x = \frac{A_k\big(\Delta p_k - \Delta p_{k,SP}\big) + A_s\big(\Delta p_s - \Delta p_{s,SP}\big) - A_b\big(\Delta p_b(T_g) - \Delta p_b(T_{g,SP})\big) \pm F_h}{C_v + C_i + C_b + K_f \Delta p_k}. \tag{11}
$$

In the following the modeling of the hysteresis force $\pm F_h$ will be discussed.

3.2 Models of hysteresis

This section discusses the modeling of the hysteresis force, $\pm F_h$. Hysteresis is a phenomenon that appears in many mechanical and electrical systems, including thermostatic valves. The hysteresis force is characterized by the fact that the valve position is a function of local phenomena such as the gradient of the velocity of the stem. The hysteresis effect is a significant factor in the modeling of thermostatic valves and it is considered important to account for this in the model which objective is for prediction and control purposes. In section 4 the effects of hysteresis on the thermostatic valve will be visualized using empirical data. In the following some modeling approaches for hysteresis will be discussed. Emphasis will be put on a model proposed in Dahl (1968) that will be included in the total model of the thermostatic valve.

The play model

A very simple model of hysteresis is the *'play model'*, see e.g. (Visintin (1994); Cahlon et al. (1997)). It is defined as two functions that separates the hysteresis region, which typically looks like the phase diagram shown in Fig. 2. If the temperature of the valve is decreasing the valve position will be described by the slope of the left side of the region. If the temperature is increasing the position is determined by the slope on the right side of the region. If the gradient of the temperature is changing the model will switch to the appropriate slope. In such a situation the valve position will not change and the movement between the two slopes will be horizontal.

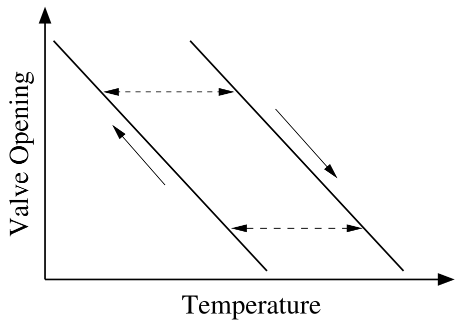


Fig. 2: Sketch of the *'play model'*.

It is obvious that the play model is a rough approximation and a pure black box model of the hysteresis, i.e. no physical interpretation is accounted for.

Dahl's model

A more detailed model is the Dahl model. It is a commonly used model for describing forces arising from friction F_h and described in Dahl (1968). The model states the relationship between the stretch and strain of a given material. The Dahl model was originally developed for simulation of control systems with friction. The model may however be used for other applications, e.g. in adaptive friction compensation, see Olsson et al. (1997). This is also the purpose of the Dahl model in this study. Let the friction force be defined as F, the Coulomb's friction force as F_c and the displacement as x. Dahl's model has then the form:

$$\frac{dF}{dx} = \sigma\Big(1 - \frac{F}{F_c} sgn(v)\Big)^{\alpha}, \tag{12}$$

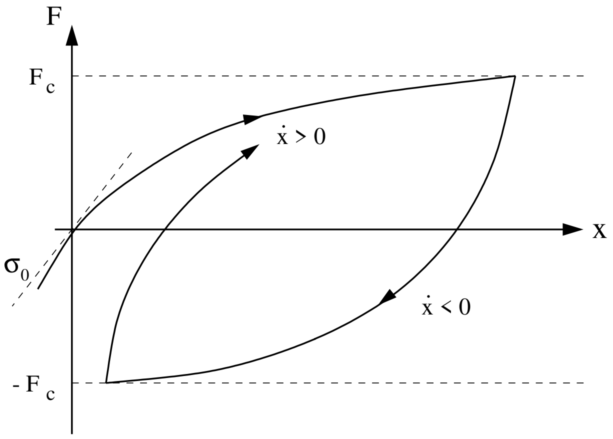


Fig. 3: Sketch of the friction force F as a function of the dissaplacement x.

where σ denotes the elasticity, and α is the slope of the movement. A sketch of the model is shown in Fig. 3. It is seen that the friction force is solely defined from the length of the displacement and its direction, $sgn(v)$. Hereby the possibility of static friction nor the Stribeck effect is not considered in the model (Olsson et al. (1997)). In Fig. 3 it is shown that as the displacement increase the friction force will equal the Coulomb's friction force F_c, and thus the maximal friction force for both negative and positive displacement, i.e. $|F| \leq F_c$. By changing the direction of the movement the curve is flipped over and the maximal friction force is then $-F_c$. The slope for $F = 0$ is indicated as σ_0.

The Dahl model can be reformulated for the time domain, see Olsson et al. (1997):

$$\frac{dF}{dt} = \frac{dF}{dx}\frac{dx}{dt} = \frac{dF}{dx}v = \sigma\Big(1 - \frac{F}{F_c}sgn(v)\Big)^{\alpha} v, \tag{13}$$

which results in a first order non-linear differential equation.

Using the Dahl model for the valve hysteresis

This section describes how the Dahl model will be used to model the friction force, F_h, in the model of the valve Eq. (11). Since Dahl's model directly gives the friction force as a function of the displacement it seems obvious to use the the valve position as x. This gives some difficulties, however, since the purpose of the model is to predict the valve position, but then one would not know the friction force before the valve position is known. There is a clear contradiction in using F_h to determine the valve position x and at the same time know the value of F_h. To overcome this problem we have modified Dahl's model to the valve problem by letting the gas temperature T_g be the parameter which determines the friction force. It is seen from empirical data in Section 4 that it is mainly the gas temperature of the valve, T_g, that can be correlated to the effects of hysteresis. By determining a friction force using Dahl's model and T_g it is possible to model the hysteresis of the valve. Hereby the adaptive model of the friction force becomes:

$$\frac{dF_h}{dt} = v\sigma\big(1 - \frac{F_h}{F_{h,max}} sgn(v)\big)^{\alpha}, \tag{14}$$

where $F_{h,max}$ is the maximum of the friction force and $v = \frac{dT_g}{dt}$ is the gradient of the gas temperature. In Olsson et al. (1997) it is argued that the slope α may be set equal to 1 and this will be assumed in the following.

In summary the steady-state model for the valve position can be formulated as a continuous time state space model, combining the dynamic friction model Eq. (14) with the force balance for the valve position Eq. (11). By introducing noise terms, the state space formulation becomes:

$$\frac{dF_h}{dt} = v\sigma\big(1 - \frac{F_h}{F_{h,max}} sgn(v)\big) + dw_t. \tag{15}$$

$$x_k = \frac{A_k\big(\Delta p_k - \Delta p_{k,SP}\big) + A_s\big(\Delta p_s - \Delta p_{s,SP}\big) - A_b\big(\Delta p_b(T_g) - \Delta p_b(T_{g,SP})\big) + F_h}{C_v + C_i + C_b + K_f \Delta p_k} + e_k. \tag{16}$$

Here dw_t is assumed to be a standard Wiener process and e_k is assumed to be a sequence of independent normally distributed variables. The introduction of the noise terms indicates that the model is an approximation to the system and that noise and unmodelled input may affect the system. Furthermore, the assumption about the noise process makes it possible to estimate unknown parameters in the model using a Maximum Likelihood method Melgaard (1994). The results of applying the model Eq. (15-16) and estimating the unknown parameters will be presented in Section 5.

4 THE EXPERIMENT AND THE DATA

This section presents empirical data that will be used for the subsequent parameter estimation and model validation. An experimental setup has been established in order to obtain a steady state characteristic of the valve position and to investigate the influence from the assumingly significant factors. Basically, the valve is put down in a small tank filled with water, mimicking the surrounding temperature, and the following variables are measured:

- The flow, q through the valve $[kg/h]$.
- The temperature of the surroundings, T_{tank} (water tank) $[^oC]$.
- Temperature of the water, T_{med}, through the valve $[^oC]$.
- The static pressure, p_s $[bar]$.
- The differential pressure, p_{dif} $[mbar]$.

The surrounding temperature (temperature of the water tank) may only be varied slowly so that the temperature in the valve can be assumed to be in steady state. Due to the slow variations it is assumed that the temperature of the gas in the thermostatic valve equals the temperature of the surroundings, i.e. the temperature of the water tank:

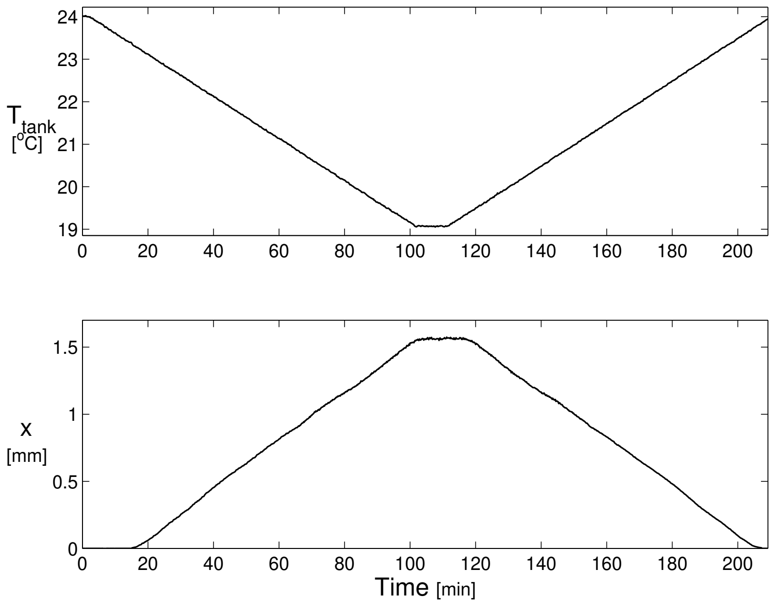


Fig. 4: Experimental data from run 1.

$$T_g = T_{tank}. \tag{17}$$

For the subsequent model identification only the temperature T_{tank} has been purposely varied by introducing steps and the resulting flow has been measured. Note that the variables p_s and p_{dif} here are kept constant, while T_{med} is assumed not to influence on the valve position since this is totally dominated by T_{tank}.

Two experimental runs have been applied to the experimental setup in order to collect data. The experimental runs are characterized by the excitation (steps) in the water temperature in order to determine the response on the valve position. During such an experiment the valve will be both closed and almost fully open. Plots of the collected data (time series) from two experimental runs are shown in Fig. 4 and 5. In Fig. 6 a phase diagram of the valve opening and the temperature of the valve is shown. The hysteresis effect is seen very clearly.

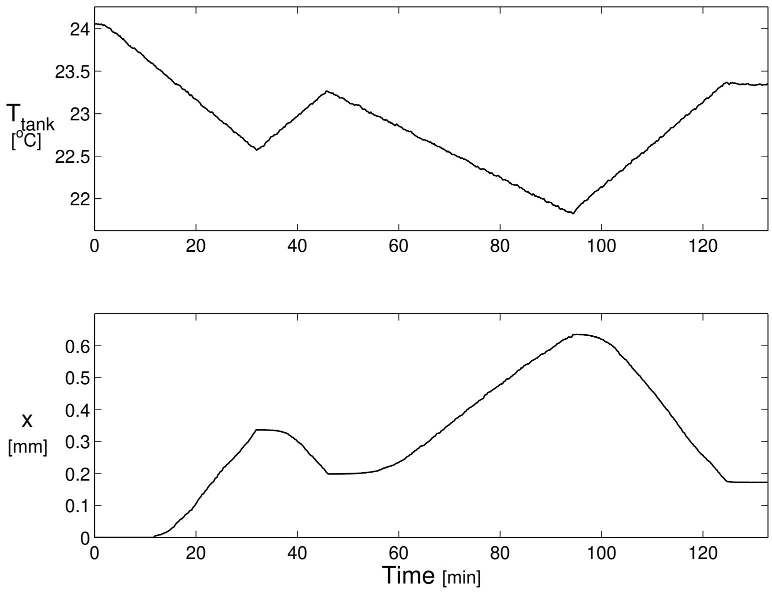


Fig. 5: Experimental data from run 2.

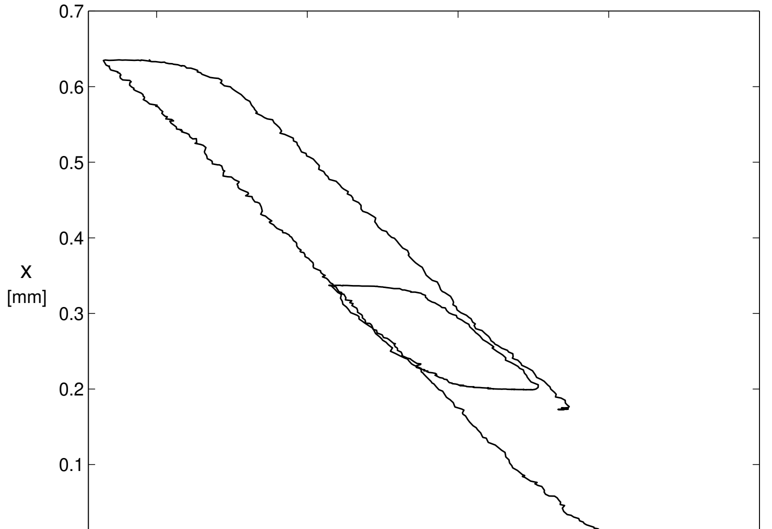


Fig. 6: Phase diagram of valve opening vs. valve temperature (run 2).

The uncertainties of the measured data used for model identification are listed in Tab. 1. The uncertainty in the flow and the various pressures are given relatively and the uncertainty of the temperatures are in absolute units.

Table 1: Measurement uncertainties

Variable	Symbol	Uncertainty σ	Unit
flow	q	$0.015 \cdot q$	kg/h
surrounding temperature	T_{tank}	0.05	^oC
differential pressure	p_{dif}	$0.015 \cdot p_{dif}$	$mbar$
static pressure	p_s	$0.015 \cdot p_s$	bar

5 RESULTS

The measured variables listed in Tab. 1 toghether with the model Eq. (15-16) has been applied and the unknown parameters have been estimated. Note that the pressure of the thermostat gas, $\Delta p_b(T_g)$, has been calculated using a higher order polynomia for the relation between $\Delta p_b(T_g)$ and T_g. Also, it should be noted that the valve position x is calculated by a formula (table values) for the relation between the flow, differential pressure and the corresponding valve position.

The parameters of the model that are known a priori, listed in Tab. 2. The remaining parameters are unknown and have been estimated. First, the parameters listed in Tab. 3 are found using data from a preliminary experiment where the temperature is constant, i.e. the hysteresis term $\pm F_h$ is set to zero. These parameters are determined using the least squares criterion. Finally, the estimated parameters of the parameters for the hysteresis function are found using the ML method, keeping the remaining parameters of Tab. 2 and 3 constant. These estimated parameters of the hysteresis function are listed in Tab. 4.

As a first check the estimated model parameters are considered. Compared to physical interpretation the estimates seem very reasonable indicating that the model formulation is realistic from a physical point of view. From an

Table 2: Known parameter values

Parameter	Value	Unit	Description
A_b	1032	mm^2	Area of Bellows
C_v	1.8	N/mm	Valve spring
C_i	13.7	N/mm	Setting spring
C_v	3.2	N/mm	Bellows spring

Table 3: Estimated parameter values found by least squares

Parameter	Value	Unit	Description
A_s	2.45	mm^2	Area of spindle
A_k	52.0	mm^2	Area of cone
K_f	99.7	mm	Flow spring

empirical point of view, a comparison between the measured and modelled output is shown in Fig. 7. It is seen that the model is able to predict the system quite accurately.

The model residuals shows no systematic patterns, and this is supported by the estimated autocorrelation function for the model residuals. The test in the estimated autocorrelation function indicates a satisfactory model fit since no information is left in the model residuals, see e.g. (Holst et al. (1992)). As another comparision the estimated standard deviation of the residuals is $\hat{\sigma}_\epsilon = 2.4 \cdot 10^{-3} mm$ which corresponds to a relative uncertainty of 1%. This is less than the measurement uncertainty, listed in Tab. 1. As a comparison to the play model the estimated variance on the prediction errors using the play model is larger by a factor 3, and also larger than the measurement uncertainty.

Table 4: Estimated parameters found by ML

variable	estimate	std.dev.
σ	6.41	0.589
$F_{h,max}$	1.31	0.067
$p_{b,set}$	0.0776	$0.892 \cdot 10^{-3}$

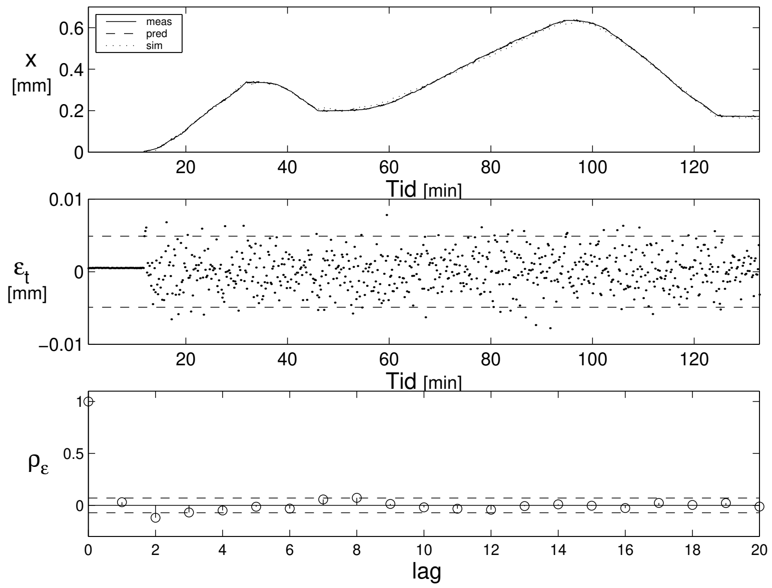


Fig. 7: Estimated model performance. Top: Measured, predicted and simulated model output. Middle: Model residuals. Bottom: Estimated autocorrelation function for the residuals.

Thus, the proposed model is significantly better than the simple play model. Furthermore the model may be interpreted physically. Finally, a cross validation study using the model estimated model on the independent experimental data from run 1 is shown in Fig. 8. It is seen that the model is still able to predict the system with high accuracy.

6 SUMMARY AND DISCUSSION

A grey box model of a thermostatic valve has been presented. The model is based on physical interpretation and statistical methods has been used to estimate and validate the model. The combination of physical knowledge and statistical methods implies that the model can be interpreted physically and that non-linear effects, such as hysteresis, can be handled effectively. Hysteresis is a phenomena that makes the modeling of thermostatic valves difficult at

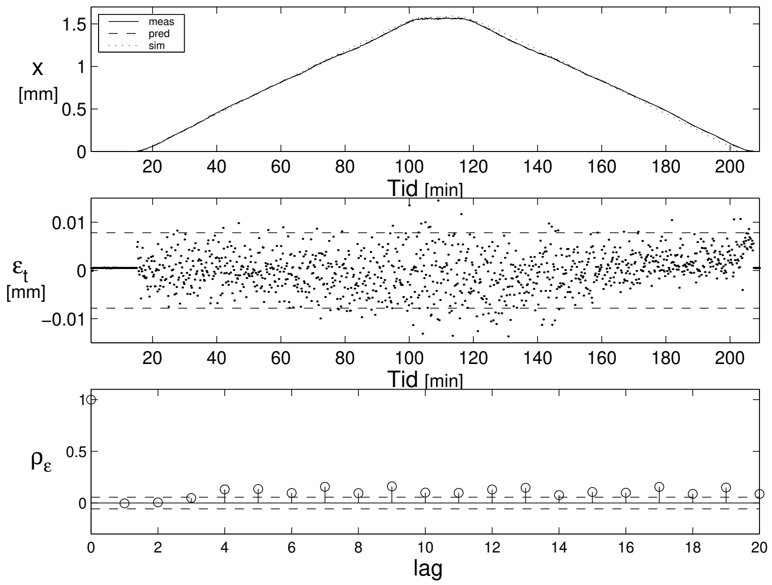


Fig. 8: Cross validation of the model. Top: Measured, predicted and simulated model output. Middle: Model residuals. Bottom: Estimated autocorrelation function for the residuals

least if the approach is solely based on empirical data. To overcome this problem an adaptive model for friction was applied to overcome for the hysteresis effect. Using statistical methods it was shown that the model is able to describe empirical data with high accuracy. If the model should be applied for system analysis it might be necessary to model the heat dynamics of the thermostat as well. However, the model for the valve position would probably still be valid. Since the Dahl model does not account for the static friction, a possibility is to use Bliman-Sorine models, in order to compensate for the lacking term Olsson et al. (1997). The Bliman-Sorine models can be interpreted as two parallel Dahl models, one fast and one slow. The fast has a larger (maximal friction force - steady state friction) than the slow one. Another aspect is to apply the model for dynamic simulations, e.g. in simulaiton of building dynamics. This would imply that the dynamics of the thermostat has to be considered as well. Preliminary results show that the model presented in this study is still useful and adequate. The results of the dynamic model will be given in a subsequent paper.

Paper F

MODELLING THE HEAT DYNAMICS OF A BUILDING USING STOCHASTIC DIFFERENTIAL EQUATIONS

ABSTRACT

This paper describes the continuous time modelling of the heat dynamics of a building. The considered building is a residential like test house divided into two test rooms with a water based central heating. Each test room is divided into thermal zones in order to describe both short and long term variations. Besides modelling the heat transfer between thermal zones, attention is put on modelling the heat input from radiators and solar radiation. The applied modelling procedure is based on collected building performance data and statistical methods. The statistical methods are used in parameter estimation and model validation, while physical knowledge is used in forming the model structure. The suggested lumped parameter model is thus based on thermodynamics and formulated as a system of stochastic differential equations. Due to the continuous time formulation the parameters of the model are directly physical interpretable. Finally, the prediction and simulation performance of the model is illustrated.

Keywords: Stochastic differential equations, heat dynamics, parameter estimation, ML method.

1 INTRODUCTION

The use of models for the heat dynamics in a building is a perceptive and practicable method to reduce the energy consumption and to improve the thermal comfort. A model can serve as a useful tool in e.g. selecting insulation materials, analyzing control strategies or in the design of a suitable heating system. Along with the varieties of applications, two different modelling procedures may be used. The traditional approach is to use knowledge of the physical building characteristics and models of subprocesses and by those means achieve a deterministic model. An alternative method is to use building performance data and statistical methods.

Various deterministic approaches are described in the literature, see e.g. Mitchell and Beckman (1995), while the literature on statistical approaches is less. A statistical approach, *the grey box modelling method*, is applied by Madsen and Holst (1995) in order to derive a total model for the heat dynamics of a building with a single test room. The proposed model is formulated as a system of stochastic differential equations and a Maximum Likelihood (ML) method is used for estimation of the parameters in the continuous-time model, based on discrete-time building performance data. As argued in Madsen and Holst (1995) there are in particular two main differences between the grey box modelling approach and the traditional deterministic approach. Firstly, while the traditional approach uses knowledge about the physical characteristics only, the grey box model structure can be identified using both knowledge about the physical characteristics of the building and information from building performance data. Secondly, while the traditional approach is kept in a deterministic framework, the grey box approach is kept in a stochastic framework. Hereby statistical methods can be applied in identifying a suitable parameterization. This is often a very difficult task using the traditional approach.

In this paper the grey box modelling method is applied. The attention is focused on modelling the heat dynamics of the indoor air temperature in a residential like building when excited by various heat inputs, such as solar radiation and heat from radiators. The building considered is a low energy test building from where building performance data from a planed experiment is used. The modelling task is extended compared to the model presented in Madsen and Holst (1995). It consists of modelling the heat dynamics in various rooms of the test building, each differently affected by solar radiation. Furthermore, the power from the radiators is not known completely and has to be modelled as well. As in Madsen and Holst (1995), the test house is not inhabited. The grey box modelling method, introduced in Section 2, is used since both information from the physics as well as information from the data can be used in determining a reasonable parameterization. The test building is described in Section 3 followed by a description of the building performance data in Section 4. Using this information, a total model is presented in Section 5. The model is formulated in continuous time, in terms of stochastic differential equations. The continuous-time formulation ensures that the parameters of the model are directly physical interpretable, while the stochastic frame work permits a description of the accuracy of the model. Statistical methods are used to identify, estimate and validate the suggested model. The results are presented in Section 6, and finally, a conclusion with discussion is given.

2 THE MODELLING APPROACH

In this section the applied modelling approach is introduced. The modelling procedure may be described by a flowchart, sketched in Fig. 1. It consists of three stages: identification, estimation and validation. The loop continues until an adequate parameterization of the model is found. In the following each stage will be further discussed.

2.1 Identifying the model structure

The first step in the modelling procedure is to select a model structure. Both information from physics and information from measurements are used to identify a suitable model parameterization. The most important variables can be recognized, and insight of the most important dynamics can be achieved by examination of measurements from the system. To model the impact from

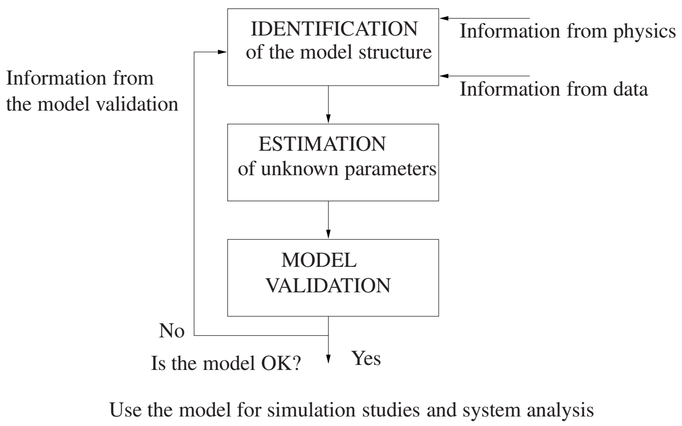


Fig. 1: The modelling procedure.

the assumingly most important variables, well known thermodynamic relationships are used, and formulated in terms of a system of ordinary differential equations:

$$d\mathbf{X}_t = \mathbf{f}\left(\mathbf{X}_t, \mathbf{U}_t, \theta, t\right) dt, \tag{1}$$

where $\mathbf{X}_t \in \mathbb{R}^{nx}$ is a vector of system states and the vector $\mathbf{U}_t \in \mathbb{R}^{nu}$ contains the known inputs. Finally, $\theta \in \mathbb{R}^{np}$ is a vector of parameters. The lumped model Eq. (1) provides a deterministic description of the evolution in time of the states of the system. It is obvious that any description of the form Eq. (1) gives only an approximation of the evolution of the true system. In order to use Eq. (1) as the foundation for a description of the true variation of the states a stochastic term is included in Eq. (1) leading to the following system of stochastic differential equation:

$$d\mathbf{X}_t = \mathbf{f}\left(\mathbf{X}_t, \mathbf{U}_t, \theta, t\right) dt + \mathbf{G}(\theta, t) d\mathbf{W}_t, \tag{2}$$

where $\mathbf{W}_t$ is a standard Wiener process and $\mathbf{G}(\theta, t)$ is a function describing how the disturbance is entering the system. There are several reasons for introducing a noise term in Eq. (1), referring to Madsen and Holst (1996):

- Modelling approximations. The system described by Eq. (1) might be an approximation to the true system.
- Unrecognized and unmodelled inputs may affect the evolution of the states.
- Measurements of the input are noise-corrupted.

For a discussion on the application of stochastic differential equations in physics, see e.g. Øksendal (1985).

2.2 Applied estimation method

Having measured some function of the state variables, a state space representation can be formulated:

$$d\mathbf{X}_t = \mathbf{f}\left(\mathbf{X}_t, \mathbf{U}_t, \theta, t\right) dt + \mathbf{G}(\theta, t) d\mathbf{W}_t, \tag{3}$$

$$\mathbf{Y}_t = \mathbf{h}\left(\mathbf{X}_t, \mathbf{U}_t, \theta, t\right) + \mathbf{e}_t, \tag{4}$$

where $\mathbf{Y}_t$ is a vector of the actually observed variables at the test building. Eq. (3) is a continuous-time system equation and Eq. (4) is the discrete time observation equation. The function $\mathbf{h}$ describes the relationship between the state variables $\mathbf{X}_t$ and the measurements $\mathbf{Y}_t \in \mathbb{R}^{ny}$. $\mathbf{e}_t \in \mathbb{R}^{ny}$ is a vector describing the measurement noise which is assumed to be Gaussian distributed. The parameter vector θ contains the equivalent thermal components i.e. capacitances, resistances etc., and is estimated by a ML method. Descriptions of the ML method can be found in (Madsen and Melgaard (1993); Nielsen et al. (2000)). The software package (Madsen and Melgaard (1993)), using the ML method, is used in estimating the parameter vector θ. Later on in the paper examples of a state space description will be illustrated.

2.3 Model validation

In the same manner as both building characteristics and information from data are used in identifying the model structure, the same information is used in validating the model. Since the model parameters are directly physical interpretable the behavior of the model should agree with the characteristics of the system, i.e. the estimated parameter vector $\hat{\theta}$ should be close to the expected value. Simultaneously, the model should be able to predict the system with a reasonable accuracy. However, it is difficult to determine the goodness of fit using this criteria only. Therefor, various statistical tests are applied in validating the model, such as residual analysis, test for model order and parameter significance. In the modelling procedure information from the model validation is a useful tool in improving the model performance. If the model validation indicates an inadequate model performance, information from the model validation can indicate how to improve the model structure, as shown in Fig. 1, et cetera.

3 THE TEST BUILDING

This section yields a description of the test building and its heating system. The purpose is to emphasize on the most important building characteristics and the influence from the heating system. This information is used in selecting a model structure for the heat dynamics in the test house. A more detailed description of the test building, the experimental design and the experiment can be found in Hansen (1997).

The test building is located at the Technical University of Denmark (DTU), near Copenhagen, and was built around 1980. The term *low energy* refers to the nominal heat consumption, which is about 2.5 KW at $-12\,^{\circ}$C outside and $20\,^{\circ}$C inside, i.e. a nominal heat loss of $\sim$ 78 W/°C, Hansen (1996). The test building is a single-storeyed building with a non-ventilated crawl space and a

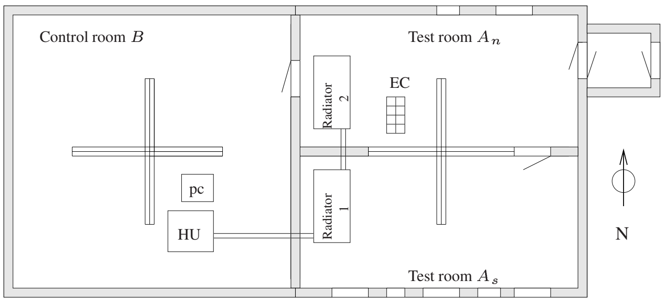


Fig. 2: A sketch of the test building and its interior.

roof space used as an office. The ground floor conforms a test area of 120 m^2 and is divided into three rooms, as sketched in Fig. 2. The test rooms, A_s and A_n, are used as test zones. The rooms are almost symmetrical and conforms each an area of 30 m^2. Test room A_s has 3.6 m^2 window area facing south while test room A_n has 1.9 m^2 window area facing north. Room B is 60 m^2 and is used as a control zone. It contains some of the equipment for the experiment. All the windows in the control room are sheiled from solar radiation. The ceiling and the outer walls of the test building are light sandwich constructions based on a masonite beam insulated with 300 mm mineral wool, while the walls separating the rooms are insulated with 95 mm mineral wool. The floor in each room are layer constructions which makes the building thermal heavy and the use of insulation makes the building extremely tight. A water based central heating system is used for the heat supply. It consists of a heating unit, HU, placed in room B, and a radiator in each of the test rooms. The two radia-

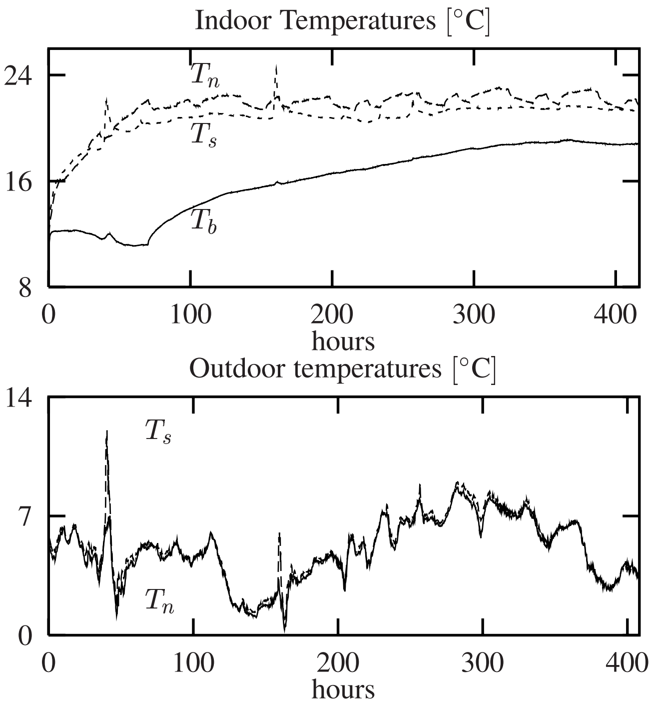


Fig. 3: Measured temperatures from an experiment at the test building. The indexes are as follows: s denotes south, n north and b the control room.

tors are identical with a nominal power of 395 W. The flow in each radiator is controlled by a thermostatic valve. An additional heat input is placed in room A_n, denoted EC. The EC-unit is remote controlled, i.e. a 0-5 V on/off signal controls an electrical on/off heat input of 0 or 150 W governed by a PRBS (Godfrey (1980)). The idea of the EC is to compensate for the sparse solar radiation on the northern side of the test building, Hansen (1996). Finally, the data acquisition system (PC) performs measurements of pressure, flow, temperature, light sensors, and consists of converter equipment and a computer, which is used to store the measured data obtained from experiments.

4 THE DATA

In this section the dynamics of the most important variables will be discussed. The aim is to use this information in identifying the model structure. The experiment was carried out in the period 22/11-19/12 1996, and discussed in Hansen (1997). Measurements from this experiment are used in the modelling procedure. The following variables were measured at a frequency of 0.1 Hz:

T_i	[°C]	indoor air temperature
T_u	[°C]	outdoor air temperature
I	[W/m^2]	solar radiation
T_s	[°C]	supply temperature to radiator
T_r	[°C]	return temperature from radiator
q	[l/h]	flow in radiator

The air temperatures were measured at several points in each of the three rooms, while the outdoor temperature and the solar radiation were measured on the southern and northern side of the test building. Finally, the supply and return temperatures as well as the flow were measured at the radiator in each test room. Measurements of the air temperatures are plotted in Fig. 3. Considering these time series, the dynamics of the air temperatures in the three rooms

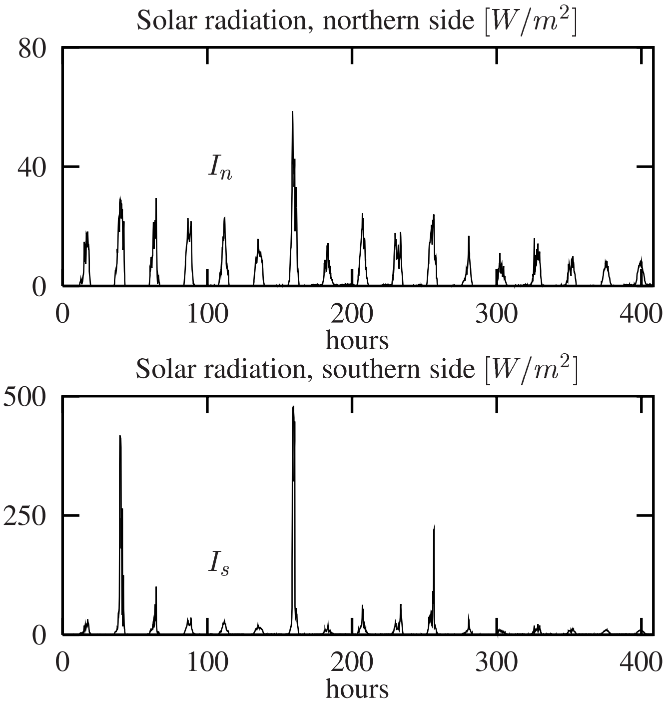


Fig. 4: Measurements of solar radiation.

are excited quite differently. This is as planed due to the experimental design, which has been performed to obtain the most informative data, see Hansen (1997). In order to excite the system, the experiment is started while the temperature in the building is low, i.e., the heating system has been shut down for a while in advance. The temperatures move towards a stationary temperature in the test rooms, due to the different heat inputs. The stationary temperature is determined by the set points of thermostatic valves, controlling the flow in the radiators. In the control room, B, the temperature T_b is increasing slowly. The main heat input to the control room is probably the heat loss from the heating unit and the conductive heat transfer from the test rooms. Since the test building is extremely tight, this explains the slow increment of the temperature. The solar radiation and the radiator flows, depicted in Fig. 4 and 5, respectively, affect the dynamics of the test room temperatures very differently. In the southern test room, A_s, the air temperature T_s is increasing rapidly when the solar radiation is high. Due to the thermostatic valve, the flow of water in the radiator q_s is reduced each time the solar radiation is high. Hence, the radi-

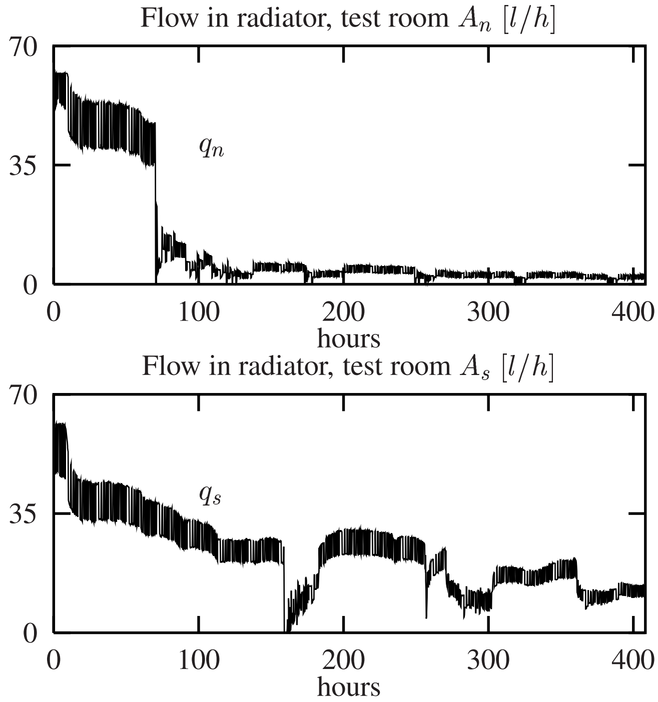


Fig. 5: Measurements of radiator flow.

ator flow and the solar radiation are strongly correlated. On the northern side of the test building the measured solar radiation I_n is severely limited compared to the radiation on the southern side of the test building. Only a part of the diffuse solar radiation is measured on the northern side. The power from the EC-unit (150 W) is the main heat input for the northern test room A_n, and only a very little flow, q_n, is required to maintain a stationary temperature. Thus, the main heat inputs and hereby the excitation of the test rooms temperatures are different. The main heat source in the northern test room is the EC-unit while the main heat source in the southern test room is the solar radiation and the power from the radiator.

5 FORMULATION OF A MODEL

In this section the formulation of a linear model for the variations of the test room air temperatures is proposed. The following items are considered important:

- The model should be able to describe all the dynamics of interest in the room air temperatures, including the impact from the solar radiation and the radiator power.
- The model should be formulated in continuous time in order to have parameters which are physical interpretable.
- The model order should be as low as possible.

Besides the information from the measurements in Fig. 3-5, additionally information from the measurements is obtained. In Fig. 6 the output from the EC-unit and the temperature T_n in the northern test room, A_n, is shown. Considering the variations in the room air temperature, both a quick response after the input is shifted and a more persistent response are recognized. This indicates that at least two time constants are required for describing the short-term and the long-term variations of the room air temperatures in each test room, i.e. a second order model. In order to model both the short-term and the long-term variations, the heat dynamics of the room air and of the thermal heavy constructions are considered. This approached is described in Madsen and Holst (1995) and it is found useful if the short-term variations of the room air temperature are important, as for instance when controllers containing a feedback from the room air temperature are considered. While the information from the data indicates that at least a second order model is required, the attention is turned towards the test building. The goal is to obtain some equations describing the heat dynamics of the room air and the thermal heavy constructions, which in this particular case is the floor. Hereby, each test room are regarded as two different thermal zones, the air and the floor. The heat equations for each zone are formed by considering the heat transfer, i.e. conduction, convection and radiation, that is assumed to have greatest influence on the heat

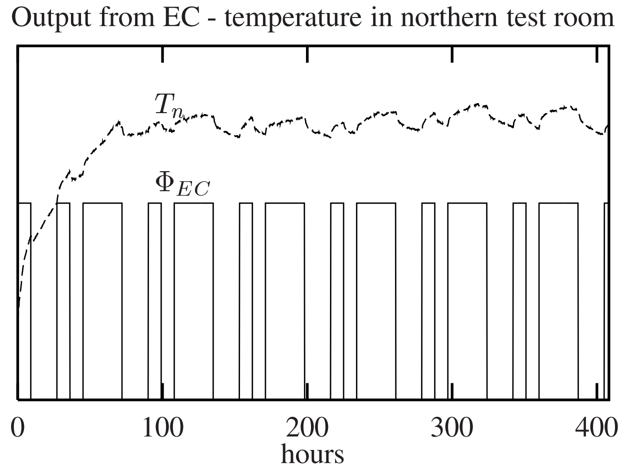


Fig. 6: The heat output from the EC-unit and the temperature in the northern test room.

dynamics in the test building. As argued in Section 4, it is assumed that the most important heat transfer that influences on the room air temperature are heat input from solar radiation, the EC-unit, the radiators and conductive heat transfer through the walls and floor in each room, as sketched in Fig. 7. Similarly, the most important heat transfer that directly affects the temperature of the floor are assumed to be heat input from solar radiation and conductive heat transfer from the room air. Taking the model approximations into account, i.e. neglecting further heat transfer, the heat dynamics can be expressed as a system of ordinary differential equations, see e.g. Ljung and Glad (1994):

$$\frac{d(\text{Heat stored})}{dt} = \sum \text{Power in} - \sum \text{Power out}, \tag{5}$$

$$\Rightarrow \quad C_i \frac{dT_i}{dt} = \sum \Phi_{in} - \sum \Phi_{out}, \tag{6}$$

where C (J/K°C) denotes heat capacity, T denotes temperature (°C) and Φ (W) the heat transfer that is appraised to influence on the heat dynamics. To derive a parameterization for Φ, well known deterministic expressions for convective, conductive and radiate heat transfer are applied. The conductive heat transfer through the walls and floor are modelled by a simple first order differential

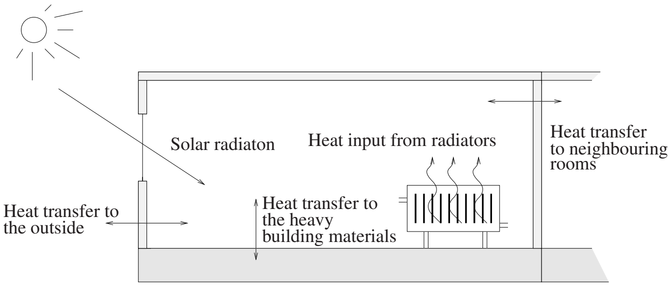


Fig. 7: A sketch of the heat transfer in the test house. The arrows symbolize the heat transfer that is assumed to be most important for the heat dynamics.

equation:

$$C\frac{dT}{dt} = \Phi_w = B_w\left(T_j - T_i\right), \tag{7}$$

where B_w (W/°C) is the thermal conductivity for a specific wall or floor w, and T_i (°C) and T_j are the temperatures at each side of the wall or floor. When using Eq. (7) it is assumed, that the heat transfer is one-dimensional. The assumed heat conduction is sketched in Fig. 8. The heat conduction between the air and the floor is sketched in Fig. 9. Note that only the heat capacity of the floor and the air are modelled. Since the walls of the test building are light, it is assumed that the heat capacities of the walls can be neglected. Hence, the heat capacities of the walls are included in the 'heat capacity of the air', and only the thermal resistance of the walls are modelled.

To model the transmitted power from solar radiation, Φ_s (W), the following relationship is used:

$$\Phi_s = IA_e, \tag{8}$$

where A_e (m^2) denotes an effective window area and I ($\mathrm{W/m}^2$) is the measured solar radiation. In Eq. (8) it is expected, that A_e for the considered building will be about 60 % of the measured window size, since only a fraction of the solar radiation will be transmitted to the inside of the test building. The remaining part will be reflected and/or absorbed by the window, Hansen et al.

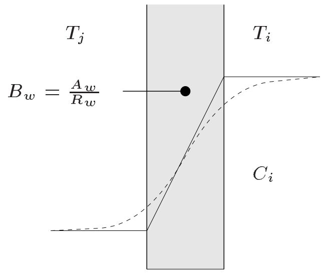


Fig. 8: Model approximations for the conductive heat transfer through a wall.

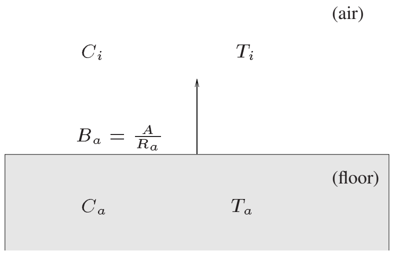


Fig. 9: The conductive heat transfer between the air and the floor.

(1990). Furthermore, it is assumed, that both the air temperature and the floor temperature are affected by the solar radiator, i.e. a fraction p of the solar radiation Φ_s is transmitted directly to the floor while the remaining part, $(1-p)$, will directly heat the air. Hereby, the input from solar radiation is modelled as:

$$\Phi_{s,air} = (1-p)IA_e \qquad \Phi_{s,floor} = pIA_e. \tag{9}$$

Now, Eq. (5) can be used to describe the heat dynamics for the temperature of the air and floor in a single test room, respectively:

$$C_i \frac{dT_i}{dt} = \sum_{walls} B_w (T_j - T_i) + B_a (T_a - T_i) + (1-p)\Phi_s + \Phi_r + \Phi_{EC}. \tag{10}$$

$$C_a \frac{dT_a}{dt} = B_a (T_i - T_a) + p\Phi_s. \tag{11}$$

In Eq. (10) the power from the radiators is denoted Φ_r. The heat output from the EC-unit is known quite well, and is not modelled. In the following different models for Φ_r will be suggested. A simple model, see e.g. Albrechtsen (1992), is given by:

$$\Phi_r = c_p \rho q (T_f - T_r), \tag{12}$$

where c_p (W/K) is the specific heat capacity of the water in the radiator, ρ (kg/m^3) is the mass density of the water, q (m^3/s) is the flow and T_f and T_r ($^\circ$C) is the supply and return water temperature. Eq. (12) simply states, that the power from the radiators is equal to the change in energy of the water, when it runs through the radiator. Another simple model is given by:

$$\frac{dQ}{dt} = cT_d \quad \text{where} \quad T_d = \frac{T_f + T_r}{2} - T_i, \tag{13}$$

where c is a constant. Eq. (13) is reported to be useful in situations where the flow is not too varying, Benonysson (1991).
If it is important to model the dynamics of the radiator as well, the following first order equation can be used:

$$C_r \frac{dT_{rad}}{dt} = c_p \rho q \left(T_f - T_r\right) - B_r \left(T_{rad} - T_\ell\right)^n, \tag{14}$$

where T_{rad} is the mean temperature of radiator surface and n is the radiator exponent, Hansen (1997). When using Eq. (12) or (13) to model the radiator power each test room is modelled by two differential equations, as in Eq. (10-11). Using Eq. (14) implies three equations for each room. Thus, using the radiator model Eq. (12) or (13) a model for the heat dynamics of the test building can be formulated:

$$C_{i,s}\frac{dT_{i,s}}{dt} = B_{u,s}\left(T_{u,s} - T_{i,s}\right) + B_b\left(T_b - T_{i,s}\right) + B_n\left(T_{i,n} - T_{i,s}\right) + B_{a,s}\left(T_{a,s} - T_{i,s}\right) + \Phi_{r,s} + A_{e,s}\left(1 - p_s\right) I_s \quad (15)$$

$$C_{a,s}\frac{dT_{a,s}}{dt} = B_{a,s}\left(T_{i,s} - T_{a,s}\right) + p_s A_{e,s} I_s \quad (16)$$

$$C_{i,n}\frac{dT_{i,n}}{dt} = B_{u,n}\left(T_{u,n} - T_{i,n}\right) + B_b\left(T_b - T_{i,n}\right) + B_n\left(T_{i,s} - T_{i,n}\right) \quad (17)$$

$$+ B_{a,n}\left(T_{a,n} - T_{i,n}\right) + \Phi_{r,n} + \Phi_{EC} + A_{e,n}\left(1 - p_n\right) I_n \quad (18)$$

$$C_{a,n}\frac{dT_{a,n}}{dt} = B_{a,n}\left(T_{i,n} - T_{a,n}\right) + p_n A_{e,n} I_n \quad (19)$$

The parameters in Eq. (15-19) are specifized in the nomenclature. The system of differential equations, Eq. (15-19), is obvious an approximation to the true system. According to the discussion in Section 2, the system is formulated in terms of stochastic differential equations, yielding the linear state space description:

$$d\mathbf{X} = \mathbf{A}\mathbf{X}dt + \mathbf{B}\mathbf{U}dt + d\mathbf{W}, \quad (20)$$

$$\mathbf{Y} = \mathbf{C}\mathbf{X} + \mathbf{D}\mathbf{U} + \mathbf{e}. \quad (21)$$

Eq. (15-19) written matrix form implies:

$$\mathbf{A} = \begin{bmatrix} \frac{-(B_{u,s}+B_b+B_n+B_{a,s})}{C_{i,s}} & \frac{B_{a,s}}{C_{i,s}} & \frac{B_n}{C_{i,s}} & 0 \\ \frac{B_{a,s}}{C_{a,s}} & \frac{-B_{a,s}}{C_{a,s}} & 0 & 0 \\ \frac{B_n}{C_{i,n}} & 0 & \frac{-(B_{u,n}+B_b+B_n+B_{a,n})}{C_{i,n}} & \frac{B_{a,n}}{C_{i,n}} \\ 0 & 0 & \frac{B_{a,n}}{C_{a,n}} & \frac{-B_{a,n}}{C_{a,n}} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} \frac{B_{u,s}}{C_{i,s}} & \frac{B_b}{C_{i,s}} & \frac{1}{C_{i,s}} & \frac{A_{e,s}(1-p_s)}{C_{i,s}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{p_s A_{e,s}}{C_{a,s}} & 0 & 0 & 0 & 0 \\ 0 & \frac{B_b}{C_{i,n}} & 0 & 0 & \frac{B_{u,n}}{C_{i,n}} & \frac{1}{C_{i,n}} & \frac{1}{C_{i,n}} & \frac{(1-p_n)A_{e,n}}{C_{i,n}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{p_n A_{e,n}}{C_{a,n}} \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} T_{i,s} \\ T_{a,s} \\ T_{i,n} \\ T_{a,n} \end{bmatrix} \quad \mathbf{U} = \begin{bmatrix} T_{u,s} \\ T_b \\ \Phi_{r,s} \\ I_s \\ T_{u,n} \\ \Phi_{r,n} \\ \Phi_{EC} \\ I_n \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \quad d\mathbf{W} = \begin{bmatrix} d\omega_1 \\ d\omega_2 \\ d\omega_3 \\ d\omega_4 \end{bmatrix} \quad \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

6 RESULTS

In this section the model performance is presented. The model parameters are estimated using the ML method and building performance data. The applied data covers the period of 17 days from the test building, discussed in Section 4. Before estimating the parameters, the data has been subsampled to a sampling time of 10 minutes. This sampling time is chosen based on the apriori estimate of the smallest time constant, which is 20 minutes, along with Shannon's theorem.

By using the modelling procedure discussed in Section 2 it is found that the radiator power in the test rooms have to be modelled differently due to different excitation of the input variables. It is found that the diffuse solar radiation on the northern side of the test building doesn't influence the heat dynamics in the northern test room significantly. Contrary, the solar radiation on the southern side of the test building has a major impact on the heat dynamics in the southern test room. It is found that only the air temperature is directly affected by solar radiation, i.e. the estimate of the fraction p is zero. Due to the fact that the radiators in the test rooms are controlled by thermostatic valves, the radiator power and the solar radiation are strongly correlated. This implies that Eq. (13) are suitable in modelling the radiator power in the northern test room while Eq. (12) and (14) are found adequate in the southern test room. The estimates and standard deviations of the remaining model parameters of the model are shown in Tab. 1.

It should be noted that only the the thermal equivalent parameters of the model Eq. (15-19), which are found significant, are tabulated. The parameters values seems reasonable compared to their expected values. However, the estimated values of the capacity of the indoor air are quite large. This can be explained by the fact, that not only the capacity of the air is modelled, but also the heat capacity of the interior in the test rooms and probably the inner part of the walls, as also reported in Madsen and Holst (1996). From that point of view,

Symbol	Estimate	Unit	Std. deviation
Bu, s	12.65	$kJ/(Kh)$	0.60
Bb	32.39	$kJ/(Kh)$	0.93
Bn	38.95	$kJ/(Kh)$	0.91
Ba, s	542.18	$kJ/(Kh)$	11.65
Bu, n	28.35	$kJ/(Kh)$	0.25
Ba, n	624.34	$kJ/(Kh)$	1.91
Ci, s	421.28	kJ/K	7.76
Ca, s	1531.85	kJ/K	67.46
Ci, n	810.19	kJ/K	15.75
Ca, n	3314.74	kJ/K	180.02
$A_{e,s}$	1.67	m^2	0.02

Table 1: Parameter estimates of the model Eq. (15-19).

	Northern room	Southern room	Unit
τ_{air}	25	42	min.
τ_{floor}	23	26	hours

Table 2: Estimated time constants.

the estimates seem very reasonable and indicates in that the model parameterization is suitable. Furthermore, the estimates of the time constants, $\hat{\tau}_i$ are calculated by the eigen-values, λ_i of the system matrix $\mathbf{A}$, i.e. $\hat{\tau}_i = -1/\lambda_i$. The estimated time constants, $\hat{\tau}_i$, are listed in Tab. 2. These estimates are close to their expected values. The heat dynamics of the air is fast compared to the slow heat dynamics of the floor, which was expected.

The model residuals have been analyzed to test if the residuals from the model can be accepted as being white noise. If this is the case, the model describes all the information given in the measurements. Test for white noise is applied through estimation of the cumulated periodogram Eq. (22) and the autocorrelation function Eq. (23):

$$\hat{C}(v_j) = \frac{\sum_{i=1}^{j} \hat{I}(v_i)}{\sum_{i=1}^{N/2} \hat{I}(v_i)} \quad \text{where } \hat{I} \text{ equals:} \tag{22}$$

$$\hat{I}(v_i) = \frac{1}{N}\left(\sum_{t=1}^{N} \varepsilon(t) cos(2\pi v_i t)\right)^2 + \frac{1}{N}\left(\sum_{t=1}^{N} (\varepsilon(t) sin(2\pi v_i t)\right)^2,$$

$$\hat{\rho}(k) = \frac{1}{N\hat{\sigma}_\varepsilon^2} \sum_{t=1}^{N-|k|} (\varepsilon(t) - \hat{\varepsilon}_\mu)(\varepsilon(t+|k|) - \hat{\varepsilon}_\mu), \tag{23}$$

where ε_t denotes the prediction error or residual at time t. Hence, the residuals are analyzed in both the frequency and time domain. The predicted states and the equivalent measurements can not be distinguished in a graph, and is therefor not plotted. However, the statistical tests, Eq. (22) and Eq. (23) may be shown visually. In Fig. 10 the residuals from each test room are compared

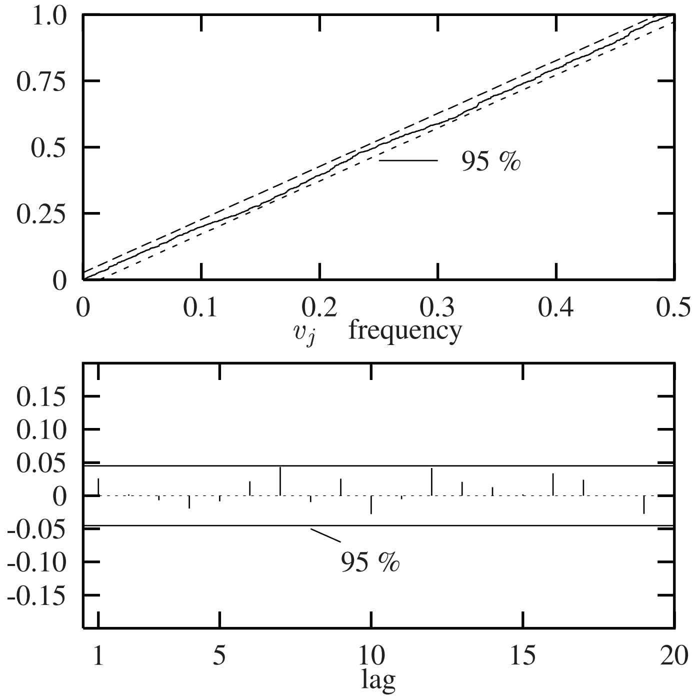


Fig. 10: Residual analysis of the prediction error. Top: Cumulated Periodogram for the residuals of the southern air temperature. Bottom: Autocorrelation function for the residuals of the northern air temperature. In both cases, the residuals can be assumed to be white noise.

with the approximated 95% interval of significance. The statistical tests show, that the residuals of each test room can be regarded as white noise. This implies, that the model order is sufficient. Finally, a simulation performance of the model using the estimated parameters is shown in Fig. 11. Considering that the temperatures were measured with an accuracy of 0.25 °C and that the maximum difference between the simulated and measured temperatures, i.e. the simulation error, is 0.40 °C during the period of 17 days, the simulation performance is considered excellent.

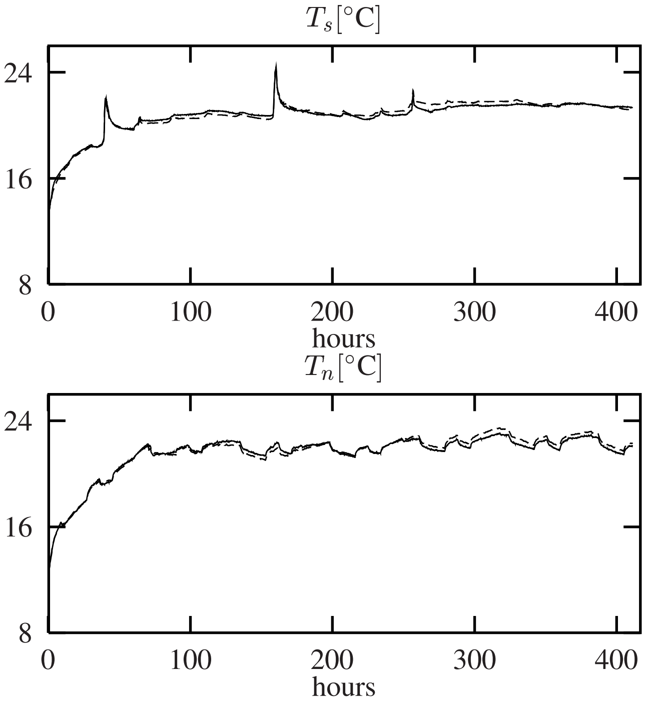


Fig. 11: Simulation performance during 17 days. Top: Simulation performance of the air temperature in the southern test room. Bottom: Simulation performance of the air temperature in the northern test room.

7 CONCLUSION

In this paper a lumped parameter model describing the heat dynamics of a residential like building has been proposed. The selected modelling method was the grey box modelling approach, i.e. the model was kept in a stochastic framework and the model structure was identified using both information from building performance data as well as known thermal properties. The model was formulated as a system of stochastic differential equations and statistic methods were applied for identifying, estimating and validating the model.

The heat dynamics in two test rooms were modelled. The two test rooms were facing opposite directions, north and south, respectively, and were hereby very differently affected by the solar radiation. Each test room was divided into two thermal zones in order to model both the short term and the long term variations. Furthermore, submodels for the radiator power and the solar radiation were presented. Finally, a linear state space model was formulated, in terms of stochastic differential equations.

The model parameters were estimated using a ML method. Through statistical tests and physical interpretation of the system, it was argued, that the model gives a good description of the heat dynamics of the considered building. Different models for the radiator power were found suitable in each test room due to the very different influence from the solar radiation on the northern and southern side of the building, respectively. Furthermore the radiator power, which was controlled by thermostatic valves, were strongly correlated with the solar radiation and affected the choise of radiator models. The estimated parameters seemed reasonable compared to the expected values of the equivalent thermal components. Furthermore, the residuals from the model could be accepted as being white noise, which indicated that the model describes all the variation in the data. Finally, it was shown that the model was able to simulate the system with a very high accuracy.

Nomenclature

Symbol	
$T_{i,s}$	Air temperature, room A_s ($°C$)
$T_{a,s}$	Floor temperature, room A_s ($°C$)
$T_{i,n}$	Air temperature, room A_n ($°C$)
$T_{a,n}$	Floor temperature, room A_n ($°C$)
T_b	Air temperature, room B ($°C$)
$T_{u,s}$	Outside air temperature on the southern side ($°C$)
$T_{u,n}$	Outside air temperature on the northern side ($°C$)
I_s	Solar radiation at the southern side (W/m^2)
I_n	Solar radiation at the northern side (W/m^2)
$\Phi_{r,s}$	Power from radiator, room A_s (W)
$\Phi_{r,n}$	Power from radiator, room A_n (W)
Φ_{EC}	Power from EC-unit, room A_n (W)
$A_{e,s}$	Effective window area facing south (m^2)
$A_{e,n}$	Effective window area facing north (m^2)
$B_{u,s}$	Conductivity coeff. from room A_s to the outside ($kJ/(Kh)$)
$B_{u,n}$	Conductivity coeff. from room A_n to the outside ($kJ/(Kh)$)
B_b	Conductivity coeff. from room A_s and A_n to room B ($kJ/(Kh)$)
B_n	Conductivity coeff. from room A_s to A_n ($kJ/(Kh)$)
$B_{a,s}$	Conductivity coeff. from the air to the floor, room A_s ($kJ/(Kh)$)
$B_{a,n}$	Conductivity coeff. from the air to the floor, room A_n ($kJ/(Kh)$)
$C_{i,s}$	Heat capacity of the air, room A_s (kJ/K)
$C_{a,s}$	Heat capacity of the floor, room A_s (MJ/K)
$C_{i,n}$	Heat capacity of the air, room A_n (kJ/K)
$C_{a,n}$	Heat capacity of the floor, room A_n (MJ/K)
p_s	Fraction of solar radiation transmitted to the floor, room A_s ($-$)
p_n	Fraction of solar radiation transmitted to the floor, room A_n ($-$)
$e_{(\cdot)}$	Measurement error ($°C$)
$dw_{(\cdot)}$	System error ($°C$)
Y_1	Air temperature, room A_s ($°C$)
Y_2	Air temperature, room A_n ($°C$)

G

Paper G

THE ERROR IN VARIABLES (EIV) REGRESSION APPROACH AS A MEANS OF IDENTIFYING UNBIASED PHYSICAL PARAMETER ESTIMATES: APPLICATION TO CHILLER PERFORMANCE

ABSTRACT

Classical regression analysis by the method of Ordinary Least Squares (OLS) considers errors to impact the dependent variable only, while it is implicitly assumed that the independent or regressor variables are free of error. This assumption is frequently unrealistic, since measurement inaccuracies in the independent variables are sometimes very large. In such cases, model coefficients identified will be biased. While this may be an issue of secondary concern if the model is used for predictive purposes only, this would be problematic if the model coefficients are to be used for diagnostic purposes. A modeling approach that can overcome this deficiency is the Error-In-Variable (EIV) regression approach that can provide unbiased parameter estimates even in the presence of errors in the regressor variables (under the assumption that the errors are unbiased and normally distributed). This paper describes this method, and based on very precise chiller performance data measured in a laboratory, illustrates the benefit and advantage of the EIV method over the OLS method. An important finding of this study is that biased parameter estimates are very likely to occur when OLS is used for identification even when well-maintained field instrumentation is used.

Keywords: Physical parameter estimation, Error In Variables Models, Ordinary Least Squares, Chiller Performance Modeling, Fault Detection and Diagnosis.

1 INTRODUCTION

Regression analysis is widely used in the engineering science to describe empirical relationships between measured data. The most widely used method of regression analysis is the Ordinary Least Squares (OLS) method that minimizes the sum of the squared residuals (see any statistical textbook, e.g. Draper and Smith (1981)). However, one of the basic assumptions in OLS, or in any classical regression analysis, is that the errors are present only in the dependent variables while the independent or regressor variables are free of error.

This assumption is frequently unrealistic, since sampling errors, measurement errors, modeling errors etc. may imply random inaccuracies in the independent variables as well, (van Huffel and Vandewalle (1991)). A consequence of neglecting the errors in the independent variables is that model parameter estimates obtained by OLS are biased; see e.g. Draper and Smith (1981). A possible remedy that could remove this bias may be the use of Error in Variable (EIV) models, discussed in Draper and Smith (1981); J. and Fuller (1989); Fuller (1987); Casella and Berger (1990).

If the purpose of the model is purely predictive, then a model determined by OLS is probably as good as the EIV model for the mean predictive value, (van Huffel and Vandewalle (1991)). However, the estimation of the associated uncertainty or confidence intervals or bands is an equally important issue. Since the prediction errors are calculated as the sum of the squared distances between the dependent variable and the model predictions, the sum of squared residuals will be unrealistically small when using OLS. This will also imply that prediction intervals estimated by the OLS method will be misleadingly tighter, an issue of concern in certain types of applications, for example in fault detection using control chart techniques, (Box and Lucenco (1997)). How to estimate prediction confidence levels in such cases is treated in advanced statistical textbooks, see e.g. Fuller (1987).

In many applications the accuracy of model parameters is important, especially when the model parameters have be interpreted physically. An obvious example is the use of physical models for Fault Detection and Diagnosis (FDD) of engineering systems. This type of model-based diagnosis has reached a certain maturity in the engineering field, see e.g. Gertler (1998); Chen and Patton (1999) and is also being increasingly adopted to the FDD of HVAC&R components and systems, (Grimmelius et al. (1995); Han et al. (1999); Haves et al. (1996); Karki and Karjalainen (1999); Lee et al. (1996); Ng et al. (1997); Rossi and Braun (1997); Stylianou and Nikanpour (1996)).

The objective of this study is to evaluate the benefit in using the EIV model approach as compared to the OLS method. Specifically, we propose to perform the evaluation in the framework of a steady state chiller performance model proposed by Gordon and Ng (1994, 1995) and later refined by Ng et al. (1997).

The Gordon-Ng model is derived from the laws of thermodynamics to which realistic approximations have been made in modeling the various heat transfer phenomena occurring in actual chillers. The great benefit of the model is that it is not only linear, but its formulation is such that the model parameters have a direct physical interpretation.

Comparison of both parameter identification approaches requires that we have the "correct" physical parameters values to start with. We propose to overcome this difficulty by using published data by Ng et al. (1997) who made very careful and meticulous performance measurements on a chiller in the lab, and also determined the physical parameters of the same chiller by direct measurement. Subsequently, we propose to add different noise levels to the chiller performance data to mimic the larger uncertainties present in field measurements. Such data would allow us to investigate the extent to which OLS and EIV parameter estimates differ from the "correct" values as the noise level is increased. Since the physical parameters are measured explicitly we can directly evaluate whether or not the parameter estimates are biased. Aside from comparing the parameter estimates themselves, tests with different noise levels can provide guidelines of how accurate the instrumentation for field monitoring should be in order to obtain reasonable accurate physical parameter estimates. Furthermore, simulation studies with both the OLS and the EIV methods can provide an indication of the extent to which sensor uncertainty influences the variance of the parameter estimates.

The paper is organized as follows: The EIV method is described in section 2, followed by that of the Gordon-Ng model in section 3. In section 4 we describe the lab chiller data used for the comparison. Different methods to obtain the EIV estimates are discussed in section 5, and estimation and simulation results are presented in section 6. The final section provides a summary of the paper.

2 REGRESSION WITH EIV

Regression with errors in variables (EIV) is also known as the measurement error model, see for example Draper and Smith (1981); Fuller (1987). The model is a generalization of the simple linear regression of the form:

$$Y_i = \alpha + \beta X_i + \epsilon_i, \tag{1}$$

where Y_i is the dependent variable(s), X_i is the regressor(s) or independent variable(s), α and β are the model parameters, and ϵ_i is the error term, here assumed to be normally distributed with mean zero and variance σ_ϵ^2.

The main difference between OLS and EIV is that in the EIV model does not assume the X_i to be known with full certainty. Instead the regressor variable is taken to be a random variable whose mean is X_i. The general EIV model assumes that observed pairs (x_i, y_i) sampled are discrete realizations of random variables (X_i, Y_i) whose means satisfy the linear relationship:

$$E[Y_i] = \alpha + \beta E[X_i]. \tag{2}$$

Posing $E[Y_i] = \eta_i$ and $E[X_i] = \xi_i$, the relationship given by Eq. (2) becomes:

$$\eta_i = \alpha + \beta \xi_i, \tag{3}$$

which is a linear relationship between the means of the random variables. The variables η_i and ξ_i are sometimes called latent variables, a term that refers to quantities that cannot be directly measured. Latent variables may be not only impossible to measure directly, but impossible to measure at all. If we are regressing on Y on X, i.e. modeling as the response and as the regressor, we can define the EIV model as:

$$Y_i = \alpha + \beta x_i + \epsilon_i \qquad \epsilon_i \in N(0, \sigma_\epsilon^2), \tag{4}$$

$$X_i = \xi_i + \delta_i \qquad \delta_i \in N(0, \sigma_\delta^2), \tag{5}$$

where $N(0, \sigma^2)$ implies a normally distributed random number with mean 0 and variance σ^2.

There are two different types of relationships that can be specified in the EIV model, one that specifies a linear functional relationship, and one describing the structural linear relationship, following Casella and Berger (1990). The linear functional relationship is as described in Eq. (4). The parameters of interest are α and β, and inference on these parameters is made using the joint distribution of (X_i, Y_i) conditional on ξ_i.

For the structural model, the dependent variables are modeled as $\xi_i \in N(\xi, \sigma_\xi^2)$. As in the functional model, the parameters of interest are α and β but inference on these parameters is now made using the joint distribution of (X_i, Y_i) unconditional on ξ_i. For physical systems a functional relationship will often seem most appealing, and in this paper we will only consider models of this type.

3 A PHYSICAL MODEL FOR CHILLER PERFORMANCE

In this section we will briefly describe the second-generation complete chiller model proposed by Gordon and Ng (1994, 1995) and later extended Ng et al. (1997), which will henceforth be referred to as GN model. We shall apply both the OLS and EIV method to obtain estimates of the physical parameters. The GN model is a simple, analytical, universal model for chiller performance based on first principles of thermodynamics and linearized heat losses. The model predicts the dependent chiller COP (defined as the ratio of chiller (or evaporator) thermal cooling capacity Q_{ch} by the electrical power E consumed by the chiller (or compressor) with specially chosen independent (and easily measurable) parameters such as the fluid (water or air) inlet temperature from the condenser T_{cdi}, fluid temperature entering the evaporator (or the chilled water return temperature from the building) T_{chi}, and the thermal cooling capacity of the evaporator. The GN model is a three-parameter model which, for parameter identification, takes the following form:

$$\left(\frac{1}{COP}+1\right)\frac{T_{chi}}{T_{cdi}}-1=a_1\frac{T_{chi}}{Q_{ch}}+a_2\frac{(T_{cdi}-T_{chi})}{T_{cdi}Q_{ch}}+a_3\frac{(1/COP+1)Q_{ch}}{T_{cdi}}, \tag{6}$$

where the temperatures are in absolute units, and Q_{ch} and E are in consistent units.

The parameters of the model given by Eq. (6), have physical meaning:

$a_1 = \Delta S$, the total internal entropy production in the chiller,
$a_2 = Q_{leak}$, the heat losses (or gains) from (or in to) the chiller,
$a_3 = R$, the total heat exchanger thermal resistance ($= \frac{1}{C_{cd}} + \frac{1}{C_{ch}}$), where C is the effective thermal conductance of either the condenser cd or the evaporator ch.

The authors point out that Q_{leak} is typically an order of magnitude smaller than the other terms. Though small, it is not negligible for accurate modeling. The authors suggest that it should be retained in the model if the other two parameters being identified are to be used for chiller diagnostics.

Let us introduce:

$$x_1=\frac{T_{chi}}{Q_{ch}},\quad x_2=\frac{(T_{cdi}-T_{chi})}{T_{cdi}Q_{ch}},\quad x_3=\frac{(1/COP+1)Q_{ch}}{T_{cdi}},\quad \text{and} \tag{7}$$

$$y=\left(\frac{1}{COP}+1\right)\frac{T_{chi}}{T_{cdi}}-1. \tag{8}$$

It is easily seen that the model given by Eq. (6) becomes:

$$y = a_1x_1 + a_2x_2 + a_3x_3. \tag{9}$$

The model has been extended to apply to the case of variable condenser flow rate in Gordon et al. (2000). Though the final functional form is non-linear, the parameter estimation can be done using methods such as OLS and EIV.

4 DATA

The data in this study, taken from Ng et al. (1997), are from a semi-hermatic reciprocating chiller with a nominal cooling capacity of 10.5 kW. It consists of 30 measurements of $(T_{chi}, T_{cdi}, Q_{ch}, E)$ taken with high accuracy in a laboratory test loop. The units of these variables and their measurement accuracy are shown in Tab. 1. It is realistic to assume that the measurement errors of the different variables share no common elemental source, i.e. that the errors are uncorrelated.

The data are plotted in Fig. 1 to provide the reader with a sense of the range of variation in the performance variables. Besides measuring the four performance variables, Ng et al. (1997) have also been able to separately measure the individual physical parameters of the chiller, namely $\Delta S, Q_{leak}$ and R by means of intrusive component-level high-grade instrumentation. Based on the 30 measurements, the mean and standard deviation of these measurements are shown in Tab. 2. The corresponding values obtained by OLS to the origi-

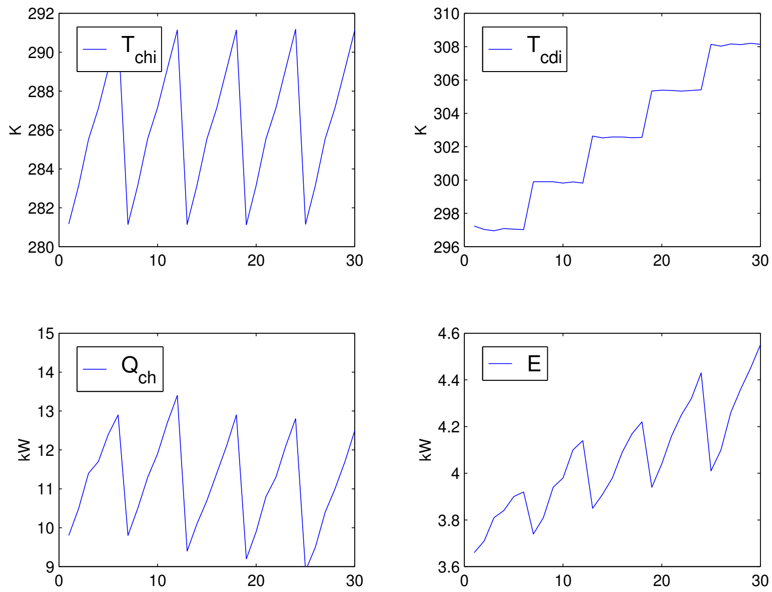


Fig. 1: The data.

nal performance data are: $R = 2.505 KkW^{-1}$, $\Delta S = 5.55\ 10^{-3} kWK^{-1}$, $Q_{leak} = 4.38 kW$. Note that these values are within the one standard deviation (except for Q_{leak}).

Table 1: Experimental uncertainty of the various performance variables for the chiller tested in a laboratory (taken from Ng et al., 1997).

Variable	T_{cdi}	T_{chi}	Q_{ch}	E
Uncertainty	$\pm 0.05\ K$	$\pm 0.05\ K$	$\pm 0.2\ kW$	$\pm 0.01\ kW$

Table 2: Mean value (E) and standard deviation (S) of the measurements of the physical parameters in the Gordon-Ng model.

$E[R]$	$S[R]$	$E[\Delta S]$	$S[\Delta S]$	$E[Q_{leak}]$	$S[Q_{leak}]$
2.689	0.177	5.518 10^{-3}	2.16 10^{-4}	2.723	1.016

5 ESTIMATION BASED ON MOMENTS

In Section 2, the difference between the OLS and EIV methods was described. In this section we shall discuss how the model parameters can be estimated, suggest different ways of determining the influence of the measurement error on the independent variables, and how to account for it.

In OLS estimation, the regression parameters are determined by:

$$\beta_{OLS} = (X^T X)^{-1} X^T Y, \tag{10}$$

where X is a matrix of size $n \times p$, i.e. p independent variables each containing n observations and Y is a matrix of size $n \times q$, where q is the number of dependent variables. Again, it must be emphasized that using OLS (10), the independent variables, X, are assumed to be free of error. If the variances of the measurement errors are known, the bias of the OLS estimator can be removed and a consistent estimator, called Corrected Least Squares (CLS) can be derived, see e.g. van Huffel and Vandewalle (1991). The CLS corrects for the extra source of error in an EIV model and can be considered to be a method-of-moments. Assuming that the correlation between errors in the dependent variables and the errors in the independent variables are uncorrelated, the estimator becomes:

$$\beta_{CLS} = (X^T X - S_{xx}^2)^{-1}(X^T Y - S_{xy}^2), \tag{11}$$

where S_{xx}^2 is a $p \times p$ matrix with the covariance of the measurement errors and $S_{xy}^2$is a $p \times 1$ vector with the covariance between the regressor variables and the dependent variable.

A simple conceptual explanation is that Eq. (11) performs on the estimator matrix an effect essentially the opposite of what ridge regression, see any statistical textbook, for example, Chatterjee (1991) does. While ridge regression 'jiggles' or randomly enhances the dispersion in the numerical values of the dependent variables in order to reduce the adverse effect of multi-collinearity on the estimated parameter bias, CLS tightens the variation in an attempt to reduce the effect of random error on the dependent variables.

We will compare the estimates obtained by β_{OLS} and β_{CLS}, respectively for the GN model. The computation of β_{OLS} is straightforward, i.e. we can use Eq. (10) directly. In order to determine β_{CLS}we will have to calculate S_{xx}^2 and S_{xy}^2. This requires knowledge of the measurement errors, and assuming that we have this insight, one of three methods can be adopted:

1. Computation of S_{xx}^2 and S_{xy}^2 by propagation of errors.
2. Computation of S_{xx}^2 and S_{xy}^2 by applying stochastic calculus.
3. Computation of S_{xx}^2 and S_{xy}^2 by stochastic simulation.

The three methods will be discussed and compared with each other below.

5.1 Computation of and by propagation of errors

Consider a model (or a function f) such that $\Phi = f(\delta, \gamma, \lambda)$ and assume that δ, γ and λ are measured with known uncertainties σ_δ^2, σ_γ^2 and σ_λ^2 respectively. Then the uncertainty σ_Φ^2 in Φ can be approximated by the error propagation formula assuming the measurement errors of different variables to be uncorrelated, i.e., assuming the covariance terms to be zero, see for example, Coleman and Steele (1999):

$$\sigma_\Phi^2 = \sigma_\delta^2(\frac{\partial f}{\partial \delta})^2 + \sigma_\gamma^2(\frac{\partial f}{\partial \gamma})^2 + \sigma_\lambda^2(\frac{\partial f}{\partial \lambda})^2. \tag{12}$$

Equation (12) can be used to compute the correction factors for the model. Since there are 3 independent variables and one dependent variable in the GN model, we have to compute a total of 9 correction factors, namely the elements in the following:

$$S_{xx}^2 = \begin{bmatrix} S_{x1,x1}^2 & S_{x1,x2}^2 & S_{x1,x3}^2 \\ S_{x2,x1}^2 & S_{x2,x2}^2 & S_{x2,x3}^2 \\ S_{x3,x1}^2 & S_{x3,x2}^2 & S_{x3,x3}^2 \end{bmatrix}, \quad S_{xy}^2 = \begin{bmatrix} S_{x1,y}^2 \\ S_{x2,y}^2 \\ S_{x2,y}^2 \end{bmatrix}, \tag{13}$$

where $S_{x1,x1}^2 = \sigma_{x1}^2$ etc. To illustrate how the elements of the matrix in Eq. (13) are determined, we consider the GN model and the calculation of $S_{x1,x1}^2$, element (1,1) of S_{xx}^2 which is the correction factor for the term $x1x1$. Recall from the GN model given by Eq. (6) that the regressor variable x_1, was defined as:

$$x_1 = \frac{T_{chi}}{Q_{ch}}.$$

Applying Eq. (12), the following correction factor is obtained:

$$S^2_{x1,x1} = \sigma^2_{x1} = \sigma^2_T(\frac{1}{Q_{ch}})^2 + \sigma^2_Q(\frac{T_{chi}}{Q^2_{ch}})^2, \tag{15}$$

where σ^2_T and σ^2_Q are the variance of the measurement error for the temperature and the thermal load, respectively.

5.2 Computation of and by applying stochastic calculus

The calculation of S^2_{xx} and S^2_{xy} can also be done by comparing the expressions of X^TX and X^TY for both noisy and noise-free variables and then applying stochastic calculus. To illustrate how this is done, let us again consider the GN model and the calculation of $S^2_{x1,x1}$, the correction factor for the term $x1x1$. Since the variables in this expression are measured and thus contain error, it is to be realized that we are measuring $\hat{T}_{cdi} = T_{cdi} + \epsilon_1$, $\hat{T}_{chi} = T_{chi} + \epsilon_2$, $\hat{Q}_{cd} = Q_{cd} + \epsilon_3$ and $\hat{E} = E + \epsilon_4$, where ϵ_i denotes the measurement error for the different variables, assumed normally distributed with mean 0 and variance σ^2_i. Thus, in the model the expression for the $\hat{x1}$ variable can be written as:

$$\hat{x1} = \frac{(T_{chi} + \epsilon_2)}{(Q_{cd} + \epsilon_3)}. \tag{16}$$

For parameter estimation, the element $\hat{x1}\hat{x1}$ (element 1,1 in) becomes:

$$\hat{x1}\hat{x1} = \left(\frac{(T_{chi} + \epsilon_2)}{(Q_{cd} + \epsilon_3)}\right)^2. \tag{17}$$

The noise-free counterpart to Eq. (17) becomes:

$$x1x1 = \left(\frac{T_{chi}}{Q_{cd}}\right)^2. \tag{18}$$

By comparing Eq. (17) where noise is included with the noise-free expression given by Eq. (18) we can find the correction factor $S^2_{x1,x1}$. We have:

$$x1x1 = \hat{x1}\hat{x1} - S^2_{x1,x1}. \tag{19}$$

The calculation of $S^2_{x1,x1}$ can be done by applying stochastic calculus, namely that in the computation we have assumed that the errors are independent with mean zero and variance equal to the square of the measurement error. In other words, for two sources of error, e.g. ϵ_1 and ϵ_2 we have:

$$E[\epsilon_1] = 0,\ E[\epsilon_2] = 0,\ E[\epsilon_1\epsilon_2] = 0,\ E[\epsilon_1\epsilon_1] = \sigma_1^2,\ E[\epsilon_2\epsilon_2] = \sigma_2^2, \tag{20}$$

and all higher order terms (equal or higher than 3) are equal to zero. By doing so we find:

$$S^2_{x1,x1} = \frac{\sigma_Q^2 T_{chi}^2 + \sigma_T^2 Q_{ch}^2}{Q_{ch}^2(Q_{ch}^2 + \sigma_Q^2)}, \tag{21}$$

where σ_T^2 and σ_Q^2 are the variance of the measurement error for the temperature and the thermal load, respectively. Although Eq. (21) may seem complicated it can easily be computed with standard computer software such as MapleV or Mathematica. Furthermore, the method can provide answers to which measurement errors that will contribute the most to the correction factor and thus are most important to reduce. Finally, it should be noted that the expression given by Eq. (21) is slightly different from the approximation found be the method of propagation of errors, i.e. Eq. (15). However, the numerical values are almost identical, and in this study the influence of the approximation made by propagation of errors has been found to be negligible.

5.3 Computation of and by applying stochastic simulation

Finally, an estimate of the correction factors may be found by stochastic (or Monte Carlo) simulation, also known as bootstrapping, see Davidson and Hinkley (1997). The method may be useful if the data at hand is assumed noise-free and one needs to investigate how increasing sensor error will affect the parameter estimates. The following equation holds:

$$x1 = \hat{x1} - S_{x1}, \tag{22}$$

where $x1$ and $\hat{x1}$ are noise-free and noisy values of the variable $x1$ respectively, and $S_{x1}$is the measurement error.

The correction factor can be estimated by adding noise to the noise-free data and then repeating this simulation to obtain a large data set (say, 1000 values) of noisy data $\hat{x1}$. The correction factor for each series is then found by:

$$S^2_{x1,x1} = (\hat{x1} - x1)^2. \tag{23}$$

Finally, a consistent estimate of the correction factor may be obtained by averaging over the total number of simulation runs.

5.4 Comparison of the correction methods

The three different methods have both advantages and disadvantages, which will be discussed below.

The method of stochastic simulation is probably best if the data at hand is close to being noise free and the purpose of the study is to investigate the affect of neglecting measurement errors in the parameter estimation. The method is less useful if the data at hand is considered noisy and the purpose of the study is to obtain corrected and less biased parameter estimates. This is not an issue for the two other methods, namely the propagation of errors and stochastic calculus methods. These may be used to investigate the affect of neglecting measurement errors in the parameter estimation and to correct for parameter bias in noisy data.

The method of propagation may seem the most attractive from a practical viewpoint. Many standard programs exist which can provide estimates of the uncertainty due to errors in the variables and thus an estimate of the correction factors. However, a drawback is, as in this case, that if the same variables enter in the expression for the different regressors, the errors in the regressors are correlated and one has to be cautious in calculating the sign for the correction factor by propagation of errors. Say, we want to calculate the correction factor $\S^2_{x1,x2} = S_{x1}S_{x2} = \sigma_{x1}\sigma_{x2}$. Both S_{x1} and S_{x2} can be calculated easily by the propagation of errors, but will always be positive. Thus if the correction factor were to be negative, special attention has to be given when using the method of propagation of errors. However, for many applications, the errors in the regressors are uncorrelated which reduces the problem to the calculation of only the diagonal elements in S^2_{xx}, and, in such instances, the method by propagation of errors may be most suitable.

The problem with negative correction factors from using the method of propagation of errors is avoided when applying stochastic calculus. The method gives the correct results but is not as well known, more tedious, and not as straightforward as the propagation of errors method.

In general, we suggest applying the method using stochastic calculus when the measurement errors in the regressors are correlated. If this is not the case, it may be more attractive to use the method of propagation of errors. Finally, if the purpose is to get a rough estimate of the effect of measurement error on the parameter estimates, the method of stochastic simulation may be the fastest and most attractive method.

6 EXAMPLE: THE GN MODEL

In this example, we will assume that the data presented in Section 4 are noise free and illustrate the effect of neglecting measurement errors in the parameter estimation. The data has been measured with high accuracy in the lab and although there may be some measurement error we will neglect this in order to show the principles of the EIV method.

Table 1 assembled the uncertainty in the four performance variables as dictated by the instrumentation used. These accuracies are representative of the best, i.e. lowest uncertainties, one could hope to achieve in the laboratory. In the field, the associated uncertainties are bound to be higher, and consequently, different levels of errors need to be introduced to this basic data set. We have done so by generating normally distributed random numbers with different levels of standard deviation (to mimic precision error). Note that we have not introduced bias to the measurements themselves. Following Coleman and Steele (1999), one can assume the error which we propose to introduce effectively include both bias and precision measurement errors.

In an effort to be realistic in how we corrupt the data, we have chosen to maintain the same magnitude of relative uncertainties of the four variables shown in Tab. 1, and to use the same multiplier for all four variables. Hence a value of 1.2 would imply that the uncertainties of all four variables have been increased by 20% from those given in Tab. 1. For the purposes of the evaluation performed, we have taken the basic measured data and added normally distributed noise ranging from a magnitude from 0 to 2 (with step of 0.1). The selection of the upper value of 2 is based on what we consider to be realistic uncertainty levels to be found in well-maintained field instrumentation relating to chillers.

We have estimated the parameters for the data at different noise levels using both the OLS method and the EIV method. For the EIV parameter estimates we have calculated the correction factors using all the three suggested methods to ensure that all three methods yield equivalent results. In Fig. 2-4 the parameter estimates in the GN model, ΔS, Q_{leak} and R, are shown when both the OLS and the EIV method are used for different magnitudes of noise added to the data. The estimation for each parameter and noise level has been repeated 1000 times using different noise sequences. The mean and corresponding 95% confidence intervals are shown in the plots. It is seen that as the noise level increases, the OLS estimates drift away from their true values. When the magnitude of the errors are greater than about 1.4 the true parameter values fall outside the 95% confidence level and may be regarded as being statistically different. Since a multiplier of 2.0 is representative of the uncertainty level to be found even in well-maintained field instrumentation, it is clear that unbiased parameter estimates are very likely to occur if OLS is used for identification. This is not the case for the EIV method. It is seen that the mean value of the parameter estimates are very close to the true values even when the magnitude of error is as large as 2.0. Thus this example illustrates, that when magnitude of the measurement error is known, consistent estimates can be obtained by the EIV method which avoids the biased estimates problem that plagues the OLS method.

The practical implication towards, say, chiller FDD work, is wthat physical parameters estimated by the OLS method and used in a control chart, see e.g. Box and Lucenco (1997), setting to detect the onset of deviations or faults in chiller performance, are likely to be biased and lead to mis-interpretation and false alarms. Parameters estimated by the EIV method are likely to yield physical parameters much more consistent with those determined by the chiller manufactures under careful in-house lab tests using high-grade sensors.

We have also investigated how the width of the uncertainty bands on the parameter estimates (not to be confused with those of model prediction) of the OLS and EIV models in the figures above are affected by larger sample size, and whether the widths differ between both methods. We found that, as expected, a larger sample size will lead to more accurate estimates and tighter parameter confidence bands, but that the width of the bands by both methods were essentially similar.

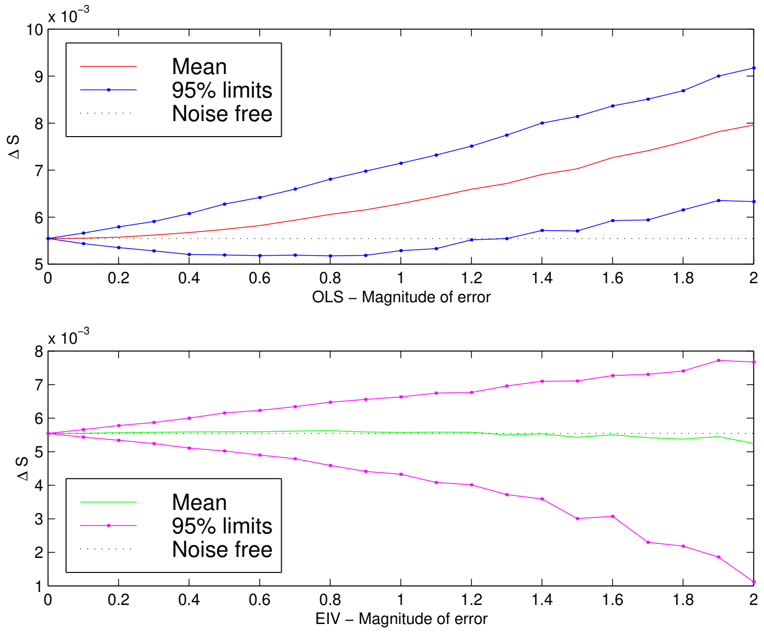


Fig. 2: Comparison of OLS (top) and EIV (bottom) estimates of the entropy.

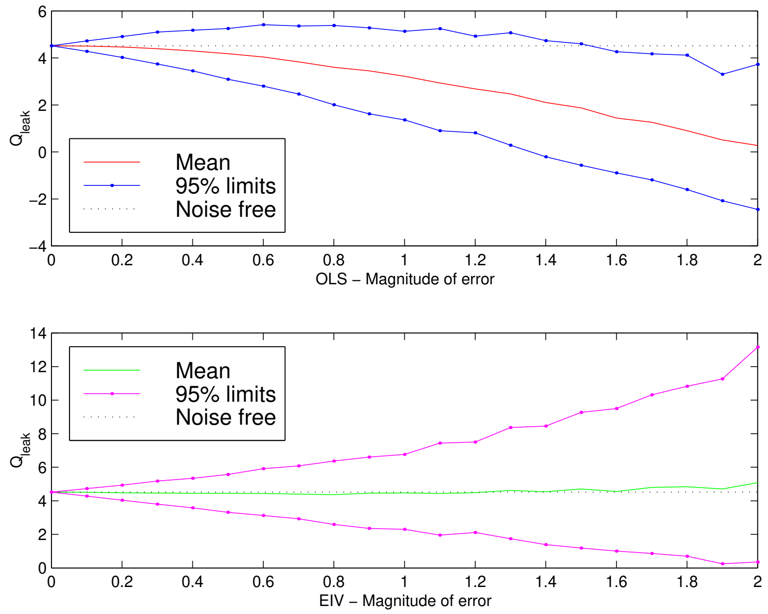


Fig. 3: Comparison of OLS (top) and EIV (bottom) estimates of the heat losses.

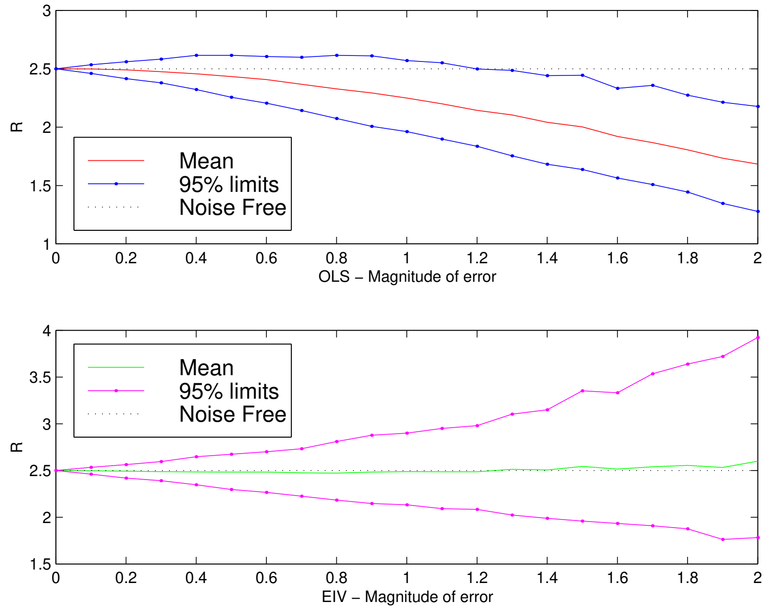


Fig. 4: Comparison of OLS (top) and EIV (bottom) estimates of the heat exchanger thermal resistance.

7 SUMMARY

The objective of this paper was to demonstrate that unbiased parameter estimates can be identified by the EIV regression approach when errors are present in the regressor variables, while OLS estimation can lead to biased estimates. However, in order for the EIV method to yield fully unbiased estimates, the levels or magnitude of the measurement errors, assumed to be random and normally distributed only, should be fairy well known in advance. Otherwise the method may suffer from the same type of problem which the OLS method does, namely biased estimates. It is obvious that practically even if one may not know the exact magnitude of such errors, even a realistic guess-estimate would be better than using OLS, the degree of lack of bias being in proportion to how accurately one is able to estimate the measurement errors in the regressors.

The paper also discussed the need for such a procedure in the framework of chiller FDD. If the purpose of the chiller model identified from actual performance data is to predict future chiller performance, model parameter bias is not a critical issue. A model identified by OLS will probably predict the mean value of the response as well as that of an EIV model, though, the prediction uncertainty bands will be under-estimated leading to the occurrence of high false alarms during fault detection. However, when one wishes to interpret the model parameters in terms of physical equipment health or performance, as provided by say the Gordon-Ng chiller model, then model parameter bias assumes a new importance, and the EIV regression approach is clearly superior to the OLS method. The evaluation of both identification methods is based on very precise chiller performance data measured in a laboratory to which different magnitudes of noise has been added to mimic field conditions. An important finding of this study is that biased parameter estimates are very likely to occur when OLS is used for identification even when well-maintained field instrumentation is used.

NOMENCLATURE

C	=	thermal conductance of the heat exchangers of the chiller
E	=	electric power consumed by the chiller
Q_{ch}	=	thermal cooling capacity of the chiller
Q_{leak}	=	heat losses or gains from or into the chiller
R	=	total heat exchanger thermal resistance
ΔS	=	total internal entropy production in the chiller
S	=	standard deviation
T_{cdi}	=	fluid inlet temperature to the condenser
T_{chi}	=	fluid inlet temperature to the chiller or evaporator
$X,\ x$	=	regressor variable(s)
$Y,\ y$	=	dependent variable(s)
$a_1,\ a_2, ..$	=	model parameters
n	=	number of observations
p	=	number of regressor variables
α, β	=	model parameters
ϵ, δ, Φ	=	error
σ_2	=	variance
ξ	=	latent variable, or unobservable state variable
η	=	latent variable, or unobservable state variable
COP	=	coefficient of performance
CLS	=	corrected least squares
EIV	=	error in variables
FDD	=	fault detection and diagnosis
OLS	=	ordinary least squares

BIBLIOGRAPHY

Adunka, F. (1992). Heizkostenverteilung bei heikorpern mit start gedrosselten durchflussen. *HLH*.

Albrechtsen, O. (1992). Central heating systems with water (in danish).

Andersen, K. K., Madsen, H., and Hansen, L. H. (2000). Modeling the heat dynamics of a building using stochastic differential equations. *Energy and Buildings*, **1**(31), 13–24.

Antoulas, A. C. and Anderson, B. D. O. (1999). On the choise of inputs in identification for robust control. *Automatica*, **35**, 1009–1031.

Asprey, S. P. and Macchietto, S. (2000). Statistical tools for optimal dynamic model building. *Computers and Chemical Engineering*, **24**, 1261–1267.

Bally, V. (1998). Construction of asymptotically optimal controls for control and game problems. *Probability Theory Related Fields*, **111**, 453–467.

Benonysson, A. (1991). Dynamic modelling and operational optimization of district heating systems. ET, DTU.

Bianchi, C. and Cleur, E. M. (1996). Indirect estimation of stochastic differential equation models: some computational experiments. *Computational Economics*, **9**, 257–274.

Bottin, S. and Chaté, H. (1998). Statistical analysis of the transition to turbulence in plane couette flow. *The European Physical Journal B*, **6**, 143–155.

Bourdouxhe, J.-P., Grodent, M., and Lebrun, J. (1998). Reference guide for dynamic models of hvac equipment. ASHRAE Research Project, Laboratory of Thermodynamics, University of Liége.

Box, G. E. P. and Jenkins, G. M. (1970). *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco.

Box, G. E. P. and Lucenco, A. (1997). *Statistical Control by Monitoring and Feedback Adjustment*. John Wiley & Sons, New York.

Brohlin, T. and Graebe, S. F. (1995). Issues in nonlinear stochastic grey-box identification. In Blanke and Söderström, editors, *SYSID '94*, volume 3, pp. 213–218. IFAC.

Cahlon, B., Schmidt, D., Shillor, M., and Zou, X. (1997). Analysis of thermostat models. *Euro. Jnl of Applied Mathematics*, **8**, 437–457.

Casella, G. and Berger, R. L. (1990). *Statistical Inference*. Duxbury Press.

Chatterjee, S. (1991). *Regression Analysis By Example Regression*. John Wiley & Sons Ltd.

Chen, J. and Patton, R. J. (1999). *Robust Model-Based Fault Detection for Dynamic Systems*. Kluwer Academic Publishers, Boston.

Cleveland, W. S. (1993). *Visualizing data*. Hobart Press.

Coleman, H. W. and Steele, W. G. (1999). *Experimentation and Uncertainty Analysis for Engineers*. John Wiley & Sons.

Dahl, P. (1968). A solid friction model. Technical report, Aerospace Corp., El Segundo, CA, USA.

Davidson, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and their Applications*. Cambridge University Press.

Draper, N. and Smith, H. (1981). *Applied Regression Analysis*. John Wiley & Sons, New York, 2 edition.

Emery, A. F., Nenarokomov, A. V., and Fadale, T. D. (2000). Uncertainties in parameter estimation: the optimal experiment design. *International Journal of Heat and Mass Transfer*, **43**, 3331–3339.

Forssell, U. and Ljung, L. (2000). Some results on optimal design. *Automatica*, **36**, 749–756.

Fox, R. W. and McDonald, A. T. (1985). *Introduction to Fluid Mechanics*. John Wiley & Sons.

Franck, G. and Rake, H. (1985). Identification of a large water-heated cross-flow heat exchanger with binary multifrequency signals. *IFAC Identification and System Parameter Estimation*, pp. 1859–1864.

Freidlin, M. and Pfeiffer, R. (1998). A threshold estimation problem for processes with hysteresis. *Statistics & Probability Letters*, **36**, 337–347.

Fuller, W. A. (1987). *Measurement Error Models*. John Wiley & Sons, New York.

Gard, T. C. (1988). *Introduction to stochastic differential equations*. Marcel Dekker.

Gertler, J. J. (1998). *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker Inc., New York.

Godfrey, K., editor (1993). *Perturbation signals for system identification*. Prentice Hall.

Godfrey, K. R. (1980). Correlation methods. *Automatica*, **16**, 527–534.

Goodwin, G. C. (1982). Experiment design. In *IFAC Identification and System Parameter Estimation*.

Goodwin, G. C. and Payne, R. L. (1977). *Dynamic system identification: Experiment design and data analysis*. Academic Press.

Gordon, J. M. and Ng, K. C. (1994). Thermodynamic modeling of reciprocating chillers. *Journal of Applied Physics*, **75**(6), 2769.

Gordon, J. M. and Ng, K. C. (1995). Predictive and diagnostic aspects of a universal thermodynamic model for chillers. *International Journal of Heat Mass Transfer*, **38**(5), 807.

Gordon, J. M., Ng, K. C., Chua, H. T., and Lim, C. K. (2000). How varying condenser flow rate affects chiller performance: Thermodynamic modeling and experimental confirmation. *Applied Thermal Eng.*, **20**, 1149–1159.

Granger, C. and Lin, J. L. (1994). Using the mutual information coefficient to identify lags in nonlinear models. *Journal of time series analysis*, **15**(4), 371–384.

Grimmelius, H. T., Woud, J. K., and Been, G. (1995). On-line failure diagnosis for compression refrigeration plants. *International Journal of Refrigeration*, **18**(1), 31–41.

Haerdle, W. (1990). *Applied nonparametric regression*. Cambridge University Press.

Han, C. Y., Xiao, Y., and Ruther, C. J. (1999). Fault detection and diagnosis of hvac systems. *ASHRAE Transactions*, **105**.

Hansen, Kjerulf-Jensen, and Stampe, editors (1990). *Varme- og klimateknik, Grundbog*. DANVAK ApS.

Hansen, L. H. (1996). Dynamic analysis of a low energy test house and a central heating system. technical report 20, IMM, DTU.

Hansen, L. H. (1997). *Stochastic Modelling of Central Heating Systems*. Ph.D. thesis, IMM, DTU.

Harremoes, P. and Madsen, H. (1999). Fiction and reality in the modelling world-balance between simplicity and complexity, calibration and identifiability, verification and falsification. *Wat. Sci. Tech.*, **39**(9), 1–8.

Haves, P., Salsbury, T., and Wright, J. A. (1996). Condition monitoring in hvac subsystems using first principle models. *ASHRAE Transactions*, **102**, 519–527.

Hensen, J. (1996). Application of modelling and simulation to hvac systems. In *30th International Conference MOSIS*.

Holst, J., Holst, U., Madsen, H., and Melgaard, H. (1992). Validation of grey box models. In *IFAC Adaptive Systems in Control and Signal Processing*, pp. 53–60, Grenoble, France.

Hurn, A. S. and Lindsay, K. A. (1997). Estimating the parameters of stochastic differential equations by monte carlo methods. *Mathematics and computers in simulation*, **43**, 495–501.

Hurn, A. S. and Lindsay, K. A. (1999). Estimating the parameters of stochastic differential equations. *Mathematics and computers in simulation*, **48**, 373–384.

J., B. P. and Fuller, W. A., editors (1989). *Statistical Analysis of Measurement Error Models and Applications*. AMS-IMS-SIAM. Proceedings of the Joint Summer Research Conference.

Johansson, R. (1993). *System modeling identification*. Prentice Hall.

Jonsson, G. (1990). *Parameter estimation in models of heat exchangers and geothermal reservoirs*. Ph.D. thesis, Department of Matematical Statistics, Lund Institute of Technology.

Jonsson, G. and Palsson, O. (1992). Use of empirical relations in the papameters of heat-exchanger models. *The AMSE transaction, Journal of Dynamic Systems, Measurements and Control*, **30**, 1193–1199.

Jonsson, G. and Palsson, O. (1994). An application of extended kalman filtering to heat exchanger models. *The AMSE transaction, Journal of Dynamic Systems, Measurements and Control*, **116**, 257–264.

Jonsson, G., Palsson, O., and Sejling, K. (1992). Modeling and parameter estimation of heat exchangers - a statistical approach. *The AMSE transaction, Journal of Dynamic Systems, Measurements and Control*, **114**, 673–679.

Jørgensen, S. B. and Hangos, K. M. (1995). Grey-box modeling for identification and control: An emerging discipline or an established technology? In Blanke and Söderström, editors, *SYSID '94*, volume 3, pp. 231–236. IFAC.

Karki, S. H. and Karjalainen, S. (1999). Performance factors as a basis of practical fault detection and diagnostic methods for air-handling units. *ASHRAE Transactions*, **105**.

Kayihan, A. (2000). Friction compensation for a process control valve. *Control Engineering Practice*, **8**, 799–812.

Khasminskii, R., Krylov, N., and Moshchuk, N. (1999). On the estimation of parameters for linear stochastic differential equations. *Probability Theory Related Fields*, **113**.

Kim, Y. T. (1999). Parameter estimation in infinite-dimensional stochastic differential equations. *Statistics & Probability Letters*, **45**, 195–204.

Kinder, M. and Wiechert, W. (1996). Stochastic simulation of biotechnical processes. *Mathematics and Computers in Simulation*, **42**, 171–178.

Kloeden, P. E., Platen, E., and Schurz, H. (1997). *Numerical Solution of SDE Through Computer Experiments*. Springer-Verlag. Berlin, 2 edition.

Lee, W. Y., Park, C., and Kelly, G. E. (1996). Fault detection in an air-handling unit using residual and recursive parameter identification methods. *ASHRAE Transactions*, **102**, 528–539.

Li, T. H. (1998). Time-correlation analysis of nonstationary time series. *Journal of time series analysis*, **19**(1), 47–67.

Li, W. K. and Hui, Y. V. (1994). Robust residual cross correlation tests for lagged relations in time series. *Journal of Statist. Comput. Simul.*, **49**, 103–109.

Lindskog, P. and Ljung, L. (1995). Tools for semi-physical modeling. In Blanke and Söderström, editors, *SYSID '94*, volume 3, pp. 237–242. IFAC.

Lindskog, P. and Ljung, L. (2000). Ensuring monotonic gain characteristics in estimated models by fuzzy model structures. *Automatica*, **36**, 311–317.

Ljung, L. (1982). Model validation. In *IFAC Identification and System Parameter Estimation*, pp. 73–75, Wahsington D.C., USA.

Ljung, L. and Glad, T. (1994). *Modeling of dynamic systems*. Prentice Hall.

Madsen, H. and Holst, J. (1995). Estimation of continuous-time models for the heat dynamics of a building. *Energy and Buildings*, **1**(22), 67–79.

Madsen, H. and Holst, J. (1996). Modelling non-linear and non-stationary time series. IMM, DTU.

Madsen, H. and Melgaard, H. (1993). *CTLSM - version 2.6*. IMSOR, DTU.

McQuarrie, A., Shumway, R., and Tsai, C. L. (1997). The model selection criterion aicu. *Statistics and Probability Letters*, **34**, 285–292.

Melgaard, H. (1994). *Identification of physical models*. Ph.D. thesis, IMM, DTU.

Mitchell, J. W. and Beckman, W. A., editors (1995). *Building Simulation '95*, volume Fourth International Conference Proceedings, Madison, Wisconsin U.S.A. International Building Performance Simulation Association.

Montgomery, D. C. (1991). *Design and Analysis of Experiments*. John Wiley & Sons Ltd., third edition.

Ng, K. C., Hua, H. T., Ong, W., Lee, S. S., and Gordon, J. M. (1997). Diagnostics and optimisation of reciprocating chillers: Theory and experiment. *Applied Thermal Engineering*, **17**(3), 263–276.

Nielsen, B. and Madsen, H. (1998). Identification of a linear continuous time stochastic model for the heat dynamics of a greenhouse. *Journal of agric. Engng. Res.*, **71**, 249–256.

Nielsen, H. and Madsen, H. (2001). A generalization of some classical time series tools. *Computational Statistics & Data Analysis*. accepted for publication.

Nielsen, J. N., Madsen, H., and Young, P. C. (2000). Parameter estimation in stochastic differential equations: an overview. *Annual Reviews in Control*, **24**, 83–94.

Norberg, R. (1995). Differential equations for moments of present values in life insurance. *Insurance: Mathematics and Economics*, **17**, 171–180.

Nowak, R. and Veen, B. V. (1993). Nonlinear system identification with pseudorandom multilevel excitation. *IEEE International Conference on Acoustics, Speech and Signal Processing*, **4**, 456–459.

Øksendal, B. (1985). *Stochastic Differential Equations*. Springer-Verlag.

Olsson, H., Aastrom, K. J., de Wit, C. C., Gafvert, M., and Lischinsky, P. (1997). Friction models and friction compensation. Technical report, LTH.

Palsson, O. P. (1993). *Stochastic modeling, control and optimization of district heating systems*. Ph.D. thesis, IMSOR, DTU.

Paulsen, O. and Grundtoft, S. (1985). Dynamic testing of heating systems (in danish). Danish Institute of Technology.

Pazman, A. (1986). *Foundations of optimal experimental design*. Kluwer Academic Publishers.

Point, N., Wouwer, V., and Remy, M. (1996). Practical issues in distributed parameter estimation: gradient computation and optimal experiment design. *Control Eng. Practice*, **4**(11), 1553–1562.

Ray, A. and Patankar, R. (1999). A stochastic model of fatigue crack propagation under variable-amplitude loading. *Engineering Fracture Mechanics*, **62**, 477–493.

Risken, H. (1996). *The Fokker-Planck Equation*. Springer, 2 edition.

Rossi, T. M. and Braun, J. E. (1997). A statistical rule-based fault detection and diagnostic method for vapor compression air conditioners. *International Journal of Heating, Ventilating, Air-Conditioning and Refrigeration Research*, **3**(1), 19–37.

Sadegh, P., Hansen, L. H., Madsen, H., and Holst, J. (1998). Input design for linear dynamic systems using maxmin criteria. *Journal of Information and Optimization Sciences*, **19**(2), 223–240.

Sejling, K. (1993). *Modelling and prediction of load in district heating systems*. Ph.D. thesis, IMSOR, DTU.

Shao, J. (1993). Linear model selection by cross-validation. *Journal of the American Statistical Association, Theory and methods*, **88**(422), 486–494.

Sowa, J. (1998). Comparison of methods of including stochastic factors into deterministic models of indoor air quality. *Energy and Buildings*, **27**, 301–308.

Stylianou, M. and Nikanpour, D. (1996). Performance monitoring, fault detection, and diagnosis of reciprocating chillers. *ASHRAE Transactions*, **102**, 615–627.

Szilder, K., Lozowski, E. P., and Reuter, G. W. (1998). A stochastic model of global atmospheric response to enchanced greenhouse warming with cloud feedback. *Atmospheric Research*, **47-48**, 475–489.

Tenno, R. and Uronen, P. (1995). State and parameter estimation for wastewater treatment processes using a stochastic model. *Control Eng. Practice*, **3**, 793–804.

Timmer, J. (2000). Parameter estimation in nonlinear stochastic differential equations. *Chaos, Solitons and Fractals*, **11**, 2571–2578.

Tong, H. (1990). *Non-linear Time Series - A Dynamical System Approach*. Clarendon Press Oxford.

Tulleken, H. (1992). *Grey-box Modelling and Identification Topics*. Ph.D. thesis, DELFT University of Technology.

Tulleken, H. (1993). Grey box modelling and identification using physical knowledge and bayesian techniques. *Automatica*, **29(2)**, 285–308.

van Huffel, S. and Vandewalle, J. (1991). *The Total Least Squares Problem - Computational Aspects and Analysis*, volume 9. Frontiers in Applied Mathematics.

Visintin, A. (1994). *Differential models of hysteresis*. Springer-Verlag.

Weyer, E., Szederkenyi, G., and Hangos, K. (2000). Grey box fault detection of heat exchangers. *Control Engineering Practice*, **8**, 121–131.

Wong, W. (1997). Frequency domain tests of multivariate gaussianity and linearity. *Journal of time series analysis*, **18**(2), 181–194.

Yarmolik, V. N. and Demidenko, S. N. (1988). *Pseudorandom Sequences for Random Testing*. John Wiley & Sons Ltd.

Zarrop, M. B. (1979). *Optimal experiment design for dynamic system identification*. Springer-Verlag.

Zou, X., Jordan, J., and Shillor, M. (1999). A dynamical model for a thermostat. *Journal of Engineering Mathematics*, **36**, 291–310.

Ph. D. theses from IMM

1. **Larsen, Rasmus.** (1994). *Estimation of visual motion in image sequences. xiv* + 143 pp.
2. **Rygard, Jens Moberg.** (1994). *Design and optimization of flexible manufacturing systems. xiii* + 232 pp.
3. **Lassen, Niels Christian Krieger.** (1994). *Automated determination of crystal orientations from electron backscattering patterns. xv* + 136 pp.
4. **Melgaard, Henrik.** (1994). *Identification of physical models. xvii* + 246 pp.
5. **Wang, Chunyan.** (1994). *Stochastic differential equations and a biological system. xxii* + 153 pp.
6. **Nielsen, Allan Aasbjerg.** (1994). *Analysis of regularly and irregularly sampled spatial, multivariate, and multi-temporal data. xxiv* + 213 pp.
7. **Ersbøll, Annette Kjær.** (1994). *On the spatial and temporal correlations in experimentation with agricultural applications. xviii* + 345 pp.
8. **Møller, Dorte.** (1994). *Methods for analysis and design of heterogeneous telecommunication networks.* Volume 1-2, *xxxviii* + 282 pp., 283-569 pp.
9. **Jensen, Jens Christian.** (1995). *Teoretiske og eksperimentelle dynamiske undersøgelser af jernbanekøretøjer. viii* + 174 pp.
10. **Kuhlmann, Lionel.** (1995). *On automatic visual inspection of reflective surfaces.* Volume 1, *xviii* + 220 pp., (Volume 2, *vi* + 54 pp., fortrolig).
11. **Lazarides, Nikolaos.** (1995). *Nonlinearity in superconductivity and Josephson Junctions. iv* + 154 pp.
12. **Rostgaard, Morten.** (1995). *Modelling, estimation and control of fast sampled dynamical systems. xiv* + 348 pp.
13. **Schultz, Nette.** (1995). *Segmentation and classification of biological objects. xiv* + 194 pp.

14. **Jørgensen, Michael Finn.** (1995). *Nonlinear Hamiltonian systems. xiv* + 120 pp.

15. **Balle, Susanne M.** (1995). *Distributed-memory matrix computations. iii* + 101 pp.

16. **Kohl, Niklas.** (1995). *Exact methods for time constrained routing and related scheduling problems. xviii* + 234 pp.

17. **Rogon, Thomas.** (1995). *Porous media: Analysis, reconstruction and percolation. xiv* + 165 pp.

18. **Andersen, Allan Theodor.** (1995). *Modelling of packet traffic with matrix analytic methods. xvi* + 242 pp.

19. **Hesthaven, Jan.** (1995). *Numerical studies of unsteady coherent structures and transport in two-dimensional flows.* Risø-R-835(EN) 203 pp.

20. **Slivsgaard, Eva Charlotte.** (l995). *On the interaction between wheels and rails in railway dynamics. viii* + 196 pp.

21. **Hartelius, Karsten.** (1996). *Analysis of irregularly distributed points. xvi* + 260 pp.

22. **Hansen, Anca Daniela.** (1996). *Predictive control and identification - Applications to steering dynamics. xviii* + 307 pp.

23. **Sadegh, Payman.** (1996). *Experiment design and optimization in complex systems. xiv* + 162 pp.

24. **Skands, Ulrik.** (1996). *Quantitative methods for the analysis of electron microscope images. xvi* + 198 pp.

25. **Bro-Nielsen, Morten.** (1996). *Medical image registration and surgery simulation. xxvii* + 274 pp.

26. **Bendtsen, Claus.** (1996). *Parallel numerical algorithms for the solution of systems of ordinary differential equations. viii* + 79 pp.

27. **Lauritsen, Morten Bach.** (1997). *Delta-domain predictive control and identification for control. xxii* + 292 pp.

28. **Bischoff, Svend.** (1997). *Modelling colliding-pulse mode-locked semiconductor lasers. xxii* + 217 pp.

29. **Arnbjerg-Nielsen, Karsten.** (1997). *Statistical analysis of urban hydrology with special emphasis on rainfall modelling.* Institut for Miljøteknik, DTU. *xiv* + 161 pp.

30. **Jacobsen, Judith L.** (1997). *Dynamic modelling of processes in rivers affected by precipitation runoff. xix* + 213 pp.

31. **Sommer, Helle Mølgaard.** (1997). *Variability in microbiological degradation experiments - Analysis and case study. xiv* + 211 pp.

32. **Ma, Xin.** (1997). *Adaptive extremum control and wind turbine control. xix* + 293 pp.

33. **Rasmussen, Kim Ørskov.** (1997). *Nonlinear and stochastic dynamics of coherent structures. x* + 215 pp.

34. **Hansen, Lars Henrik.** (1997). *Stochastic modelling of central heating systems. xxii* + 301 pp.

35. **Jørgensen, Claus.** (1997). *Driftsoptimering på kraftvarmesystemer.* 290 pp.

36. **Stauning, Ole.** (1997). *Automatic validation of numerical solutions. viii* + 116 pp.

37. **Pedersen, Morten With.** (1997). *Optimization of recurrent neural networks for time series modeling. x* + 322 pp.

38. **Thorsen, Rune.** (1997). *Restoration of hand function in tetraplegics using myoelectrically controlled functional electrical stimulation of the controlling muscle. x* + 154 pp. + Appendix.

39. **Rosholm, Anders.** (1997). *Statistical methods for segmentation and classification of images. xvi* + 183 pp.

40. **Petersen, Kim Tilgaard.** (1997). *Estimation of speech quality in telecommunication systems. x* + 259 pp.

41. **Jensen, Carsten Nordstrøm.** (1997). *Nonlinear systems with discrete and continuous elements.* 195 pp.

42. **Hansen, Peter S.K.** (1997). *Signal subspace methods for speech enhancement. x* + 226 pp.

43. **Nielsen, Ole Møller.** (1998). *Wavelets in scientific computing. xiv* + 232 pp.

44. **Kjems, Ulrik.** (1998). *Bayesian signal processing and interpretation of brain scans. iv* + 129 pp.

45. **Hansen, Michael Pilegaard.** (1998). *Metaheuristics for multiple objective combinatorial optimization. x* + 163 pp.

46. **Riis, Søren Kamaric.** (1998). *Hidden markov models and neural networks for speech recognition. x* + 223 pp.

47. **Mørch, Niels Jacob Sand.** (1998). *A multivariate approach to functional neuro modeling. xvi* + 147 pp.

48. **Frydendal, Ib.** (1998.) *Quality inspection of sugar beets using vision. iv* + 97 pp. + app.

49. **Lundin, Lars Kristian.** (1998). *Parallel computation of rotating flows. viii* + 106 pp.

50. **Borges, Pedro.** (1998). *Multicriteria planning and optimization. - Heuristic approaches. xiv* + 219 pp.

51. **Nielsen, Jakob Birkedal.** (1998). *New developments in the theory of wheel/rail contact mechanics. xviii* + 223 pp.

52. **Fog, Torben.** (1998). *Condition monitoring and fault diagnosis in marine diesel engines. xii* + 178 pp.

53. **Knudsen, Ole.** (1998). *Industrial vision. xii* + 129 pp.

54. **Andersen, Jens Strodl.** (1998). *Statistical analysis of biotests. - Applied to complex polluted samples. xx* + 207 pp.

55. **Philipsen, Peter Alshede.** (1998). *Reconstruction and restoration of PET images. vi* + 132 pp.

56. **Thygesen, Uffe Høgsbro.** (1998). *Robust performance and dissipation of stochastic control systems.* 185 pp.

57. **Hintz-Madsen, Mads.** (1998). *A probabilistic framework for classification of dermatoscopic images. xi* + 153 pp.

58. **Schramm-Nielsen, Karina.** (1998). *Environmental reference materials methods and case studies. xxvi* + 261 pp.

59. **Skyggebjerg, Ole.** (1999). *Acquisition and analysis of complex dynamic intra- and intercellular signaling events.* 83 pp.

60. **Jensen, Kåre Jean.** (1999). *Signal processing for distribution network monitoring. xv* + 199 pp.

61. **Folm-Hansen, Jørgen.** (1999). *On chromatic and geometrical calibration. xiv* + 238 pp.

62. **Larsen, Jesper.** (1999). *Parallelization of the vehicle routing problem with time windows.xx* + 266 pp.

63. **Clausen, Carl Balslev.** (1999). *Spatial solitons in quasi-phase matched struktures. vi* + (flere pag.)

64. **Kvist, Trine.** (1999). *Statistical modelling of fish stocks. xiv* + 173 pp.

65. **Andresen, Per Rønsholt.** (1999). *Surface-bounded growth modeling applied to human mandibles. xxii* + 125 pp.

66. **Sørensen, Per Settergren.** (1999). *Spatial distribution maps for benthic communities.*

67. **Andersen, Helle.** (1999). *Statistical models for standardized toxicity studies. viii* + (flere pag.)

68. **Andersen, Lars Nonboe.** (1999). *Signal processing in the dolphin sonar system. xii* + 214 pp.

69. **Bechmann, Henrik.** (1999). *Modelling of wastewater systems. xviii* + 161 pp.

70. **Nielsen, Henrik Aalborg.** (1999). *Parametric and non-parametric system modelling. xviii* + 209 pp.

71. **Gramkow, Claus.** (1999). *2D and 3D object measurement for control and quality assurance in the industry. xxvi* + 236 pp.

72. **Nielsen, Jan Nygaard.** (1999). *Stochastic modelling of dynamic systems. xvi* + 225 pp.

73. **Larsen, Allan.** (2000). *The dynamic vehicle routing problem. xvi* + 185 pp.

74. **Halkjær, Søren.** (2000). *Elastic wave propagation in anisotropic inhomogeneous materials. xiv* + 133 pp.

75. **Larsen, Theis Leth.** (2000). *Phosphorus diffusion in float zone silicon crystal growth. viii* + 119 pp.

76. **Dirscherl, Kai.** (2000). *Online correction of scanning probe microscopes with pixel accuracy.* 146 pp.

77. **Fisker, Rune.** (2000). *Making deformable template models operational. xx* + 217 pp.

78. **Hultberg, Tim Helge.** (2000).*Topics in computational linear optimization. xiv* + 194 pp.

79. **Andersen, Klaus Kaae.** (2001).*Stochastic modelling of energy systems. xiv* + 191 pp.

80. **Thyregod, Peter.** (2001). *Modelling and monitoring in injection molding. xvi* + 132 pp.

81. **Schjødt-Eriksen, Jens.** (2001). *Arresting of collapse in inhomogeneous and ultrafast Kerr media.*

82. **Bennetsen, Jens Christian.** (2000). *Numerical simulation of turbulent airflow in livestock buildings. xi* + 205 pp + Appendix.

83. **Højen-Sørensen, Pedro A.d.F.R.** (2001). *Approximating methods for intractable probabilistic models: - Applications in neuroscience. xi* + 104 pp + Appendix.

84. **Nielsen, Torben Skov.** (2001). *On-line prediction and control in non-linear stochastic systems. xviii* + 242 pp.

85. **Öjelund, Henrik.** (2001). *Multivariate calibration of chemical sensors. xviii* + 184 pp.

86. **Adeler, Pernille Thorup.** (2001). *Hemodynamic simulation of the heart using a 2D model and MR data. xv* + 180 pp.

87. **Nielsen, Finn Årup.** (2001). *Neuroinformatics in functional neuroimaging.* 330 pp.

88. **Kidmose, Preben.** (2001). *Blind separation of heavy tail signals. viii* + 136 pp.

89. **Hilger, Klaus Baggesen.** (2001). *Exploratory analysis of multivariate data. xxiv* + 186 pp.

90. **Antonov, Anton.** (2001). *Object-oriented framework for large scale air pollution models.* 156 pp. + (flere pag).

91. **Poulsen, Mikael Zebbelin.** (2001). *Practical analysis of DAEs.* 130 pp.