

Reducing Prediction Uncertainty of Weather Controlled Systems

Timotheus Gerardus Doeswijk

Promotor: Prof. dr. ir. G. van Straten
Hoogleraar Meet-, Regel-, en Systeemtechniek
Wageningen Universiteit

Co-promotor: Dr. Ir. K.J. Keesman
Universitair hoofddocent, leerstoelgroep
Meet-, Regel-, en Systeemtechniek
Wageningen Universiteit

Promotie-commissie: Prof. Dr. A.A.M. Holtslag
Wageningen Universiteit
Prof. Dr. J. Molenaar
Wageningen Universiteit
Prof. Dr. A. Bagchi
Universiteit Twente
Dr. Ir. L.J.S. Lukasse
Agrotechnology & Food Innovations

Dit onderzoek is uitgevoerd binnen de onderzoekschool Production Ecology & Resource Conservation

Reducing Prediction Uncertainty of Weather Controlled Systems

Timotheus Gerardus Doeswijk

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit
Prof. Dr. M.J. Kropff
in het openbaar te verdedigen
op dinsdag 24 april 2007
des namiddags om half twee in de aula

Reducing Prediction Uncertainty of Weather Controlled Systems, 2007

Timotheus Gerardus Doeswijk

Ph.D. thesis Wageningen University, Wageningen, The Netherlands - with summary in Dutch - 174 p.

Keywords: uncertainty analysis, weather forecasts, storage, parameter estimation, model prediction, total least squares

ISBN: 90-8504-613-0

Voor Chantal
Tosca, Blanca, Milo

Abstract

In closed agricultural systems the weather acts both as a disturbance and as a resource. By using weather forecasts in control strategies the effects of disturbances can be minimized whereas the resources can be utilized. In this situation weather forecast uncertainty and model based control are coupled. In this thesis a model of a storage facility of agricultural produce is used as an example for a weather controlled system.

The first step in reducing prediction uncertainty is taken by reducing the uncertainty in the weather forecast itself. A Kalman filter approach is used for this purpose. Weather forecast uncertainty is significantly reduced up to 10 hours for temperature, up to 32 hours for wind speed and up to 3 hours for global radiation by using this approach.

For a linearized model of the storage facility error propagation rules have been derived. The uncertainty of the output can therefore be analytically calculated. The medium range weather forecast, *i.e.* up to ten days ahead, consists of an ensemble of 50 forecasts. The mean and variance of these forecasts are used for model prediction and model output uncertainty prediction. Furthermore, by using optimal control in conjunction with a cost criterion, the uncertainty of the system state is incorporated into the cost criterion. As a result the control inputs shift towards parts with less uncertainty in the weather forecast. Finally, a numerical risk evaluation showed that if feedback is applied, as in receding horizon optimal control, the cost increase is limited to 5% for a 24 hour feedback interval.

Mathematical models are always an approximation of the real system. Model uncertainties arise in the model structure and/or in unknown parameter values. If measurements from the system are present, the model is fitted to the data by changing parameter values. Generally, parameters that are nonlinear in system model output are estimated by nonlinear least-squares (NLS) optimization algorithms. For systems that rational in their parameters a reparameterization method is proposed such that the new parameters can be estimated with ordinary least squares. As a result a modified predictor appears. In the noise free case this leads to the exact parameter estimates whereas a nonlinear least squares approach might end up in a local minimum. If noise is present, however, the linear estimates might be biased and the modified predictor has only a limited range. Because after linear reparameterization the data structure generally becomes errors-in-variables a bias compensated total least squares approach is used. The predictive performance of the modified predictor in this case largely improves and is regarded as a powerful alternative to the existing least squares methods.

Contents

Abstract	vii
Table of Contents	1
1 Introduction	5
Part I Uncertainty Analysis	17
2 Adaptive Weather Forecasting using Local Meteorological Information	19
3 Improving Site-Specific Weather Forecasts for Agricultural Applications	31
4 Uncertainty Analysis of a Storage Facility under Optimal Control	43
5 Impact of Weather Forecast Uncertainty in Optimal Climate Control of Storehouses	59
Part II Parameter Estimation	73
6 Parameter Estimation and Prediction of a Rational Storage Model: an ordinary least squares approach	75
7 Parameter Estimation and Prediction of a Rational Storage Model: a total least squares approach	89
8 Linear Parameter Estimation of Rational Biokinetic Functions	101
9 Conclusions and general discussion	117
Bibliography	128
Appendices	139
A Nomenclature	141

B Storage model	147
C Stability of filter	153
D Bias compensation	157
Summary	161
Samenvatting	165
Dankwoord	169
Curriculum Vitae	171
List of Publications	173

Chapter 1

Introduction

1.1 Background

Many systems in agricultural practice are influenced by the weather. When these systems are closed they are usually controlled in order to keep the climate inside the system on a reference trajectory. In this context the weather is a disturbance of the system. In practice, the weather does not only act as a disturbance but is also necessary as a resource, *e.g.* the effect of global radiation on plant growth. An enhancement would be to use weather forecasts in control strategies. This idea was implemented in the EET¹-project "weather in control" [LB06]. The main purpose of the project "weather in control" was (i) to compute *optimal control* strategies that anticipate to changing environmental conditions (ii) to analyze the effects of weather forecast and parameter uncertainty on the predicted process states and calculated optimal controls. In this concept, the weather is then no longer a disturbance but an external input driving the system.

For greenhouses optimal control studies related to crop growth can be found in literature *e.g.* [GAS84, vSCB00]. Since these have long term objectives (months) in terms of crop yield the outside temperature used in the model comes from climatic data. However, if the greenhouse climate itself is subject to an optimal control or model predictive control strategy prediction horizons of no more than one hour are used [TVWVS96, COC05, PCKP05]. For these short term horizons the lazy man weather prediction, *i.e.* the observed weather will not change during the next hour, proved to be a better solution than using weather forecasts [TVWVS96]. Ultimately, a two time-scale decomposition can be used to gain from both: a maximum crop yield and minimizing resources (*i.e.* greenhouse climate optimization) [vSvWT02]. In the area of livestock buildings optimal control strategies are very scarce and in these cases no use is made of dynamic models of the indoor climate [SL93, TGB⁺95]. Dynamic models in livestock buildings are used, however, for the estimation of the variables to be controlled such as the temperature in the animal occupied zone, that cannot be measured directly [VWAVB⁺05] and in robust control design [DAPS06]. Model prediction based on weather forecasts are not found. A third application area on which ambient weather conditions act is in post-harvest processes. The interest in control concepts for drying in general has increased only in the last decade [Duf06]. This accounts especially for optimal control concepts. In many cases ambient air is (partially) used to achieve the control objectives (*e.g.* [EJM⁺97, RPM06, SHD06]). Again, no use is made of weather forecasts. In other post harvest processes like storage, studies of control with ambient air are very limited. In [KPL03] it was

¹EET: subsidy program from the Dutch Ministry of Finance, the Ministry of Education, Culture and Science and the Ministry of Housing, Regional Development and Environment (project no. EETK01120)

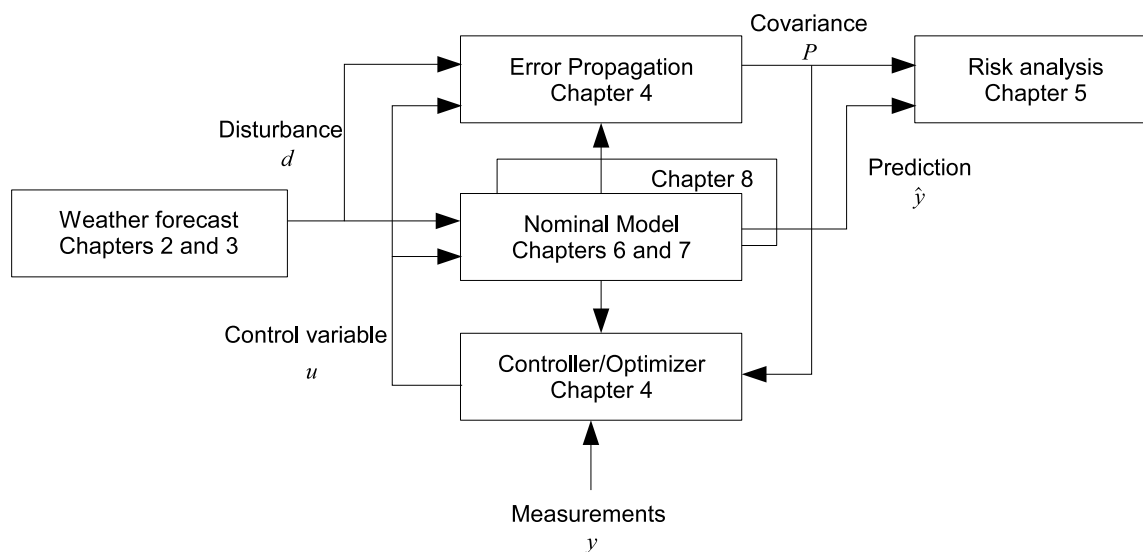


Figure 1.1: schematic overview

shown that given a perfect weather forecast total costs can be reduced up to four days ahead. Disturbances, however, are not accounted for in this study.

In figure 1.1 a schematic overview is given of the use of *weather forecasts* in dynamical models and the analysis of the uncertainty in weather forecasts. The main application area in this thesis relates to storage of potatoes (see appendix B for details of the model). In part I the uncertainty from weather forecasts is analyzed (chapters 2 and 3). Furthermore, the propagation of this uncertainty and a risk evaluation of the predicted output is subject of research (chapter 4). Finally, the predicted uncertainty can be part of an optimal control algorithm (chapter 5). Both model prediction and error propagation require a good nominal model. In part II a model calibration technique for rational models is proposed to obtain a good predictor. First, the technique is applied to the storage model (chapters 6 and 7). Second, the approach is extended to rational biokinetic models (chapter 8).

1.2 Error Propagation

1.2.1 Weather forecasting

Weather forecasts are predictions of the future state of the atmosphere. Because of the chaotic nature of the weather exact initial conditions of the numerical weather prediction models are required in order to provide an accurate forecast [LOR63]. Since these initial conditions are only approximations, the weather forecast becomes uncertain in the longer range. In order to obtain a measure of the uncertainty ensemble prediction [LEI74, PBBP97] was developed. An ensemble forecast consists of several members, where each member has an equal probability of occurrence. The

members are the outcome of a numerical weather prediction model with perturbed initial states.

With the ensemble prediction an uncertainty measure of the weather forecast becomes directly available [Zhu05b] and is also known as skill prediction of the weather forecast. In addition, a distinction can be made between forecasts with lower or higher than average expected uncertainty [TZM01, Zie01].

Three different types of goodness of weather forecasting are distinguished [Mur93]: consistency, quality and value. Consistency is defined by the correspondence between forecasts and the forecasters' judgements, quality by the correspondence between forecasts and observations and value as incremental benefits of forecasts to users. Since consistency depends on the forecasters' judgement which is not known to the user this type of goodness is outside the scope of this thesis. Users in application areas such as agriculture or electricity production are operating on a local scale and are interested in the quality of a weather forecast. Furthermore, if decisions are made upon weather forecasts it can also be related to value [Atz95, KPL03, Wil98, Tei05]. In this thesis the quality plays the central role but the close relationship with value is obvious.

Because local weather forecasts are essential for local systems it is important to have a good local weather forecast (see part I). In meteorology postprocessing techniques like model output statistics, perfect prognosis and reanalysis [Mar06] are in use to correct model outputs. The commercially available (corrected) weather forecasts, however, are only given for specific predefined locations. Methods based on the Kalman filter that reduce the bias of weather forecasts are described in the literature to account for large differences in mountainous areas [Hom95] or in the presence of large water surfaces such as seas or lakes [GA02]. Near the ground the type of vegetation and soil [Gei61] determines the microclimate. However, the vegetation does not only effect the climate near the ground but also at larger distances from the ground [WIL91, CHE93]. Furthermore, urbanization has a substantial effect on the local climate [Mak06].

Local users of weather forecasts are able to purchase forecasts up to ten days ahead by commercial weather agencies. In general, these forecasts are not corrected for the local situation. Hence, there is a need to update regional forecasts for local users. For the short term (up to 36 hours) a single deterministic forecast is available in the Netherlands. For the medium range forecast (up to 10 days) ensembles are used. As mentioned before, the medium range ensembles can be treated as probabilistic forecasts to estimate the probability distribution. Given this probability distribution economic decision rules can be optimized [SRH01, Zhu02].

1.3 Estimation methods

Calibration of a mathematical model is an important procedure to optimize the predictive quality of the model. In this thesis (part II) several parameter estimation techniques, *i.e.* nonlinear least squares, ordinary least squares and total least squares are used to calibrate a model of a storage facility of agricultural produce. The methods that will be used are briefly described in the following subsections. An example of a very simple deterministic dynamic model is given by:

$$y(k+1) = ay(k) + bu(k) \quad (1.1)$$

a linear time-invariant difference equation model. This model will be used to illustrate the differences between the described techniques. It is assumed that the model structure is known beforehand and that the unknowns are the parameters a and b .

1.3.1 Nonlinear least squares

For models that are nonlinear in the parameters iterative nonlinear least squares methods are usually used to estimate the parameters. A disadvantage of these methods is that generally they might end up in local minima. The general procedure of nonlinear least squares is that a quadratic loss function is defined

$$V = \sum_{k=1}^m (e(k))^2 \quad (1.2)$$

with $e(k)$ the prediction error $y(k) - \hat{y}(k)$, *i.e.* the measurement $y(k)$ minus the k -step ahead prediction $\hat{y}(k)$, and m the total number of measurements. By changing the parameters the value of the loss function also changes. A minimum is found if and only if the following criteria hold:

$$\frac{\partial V}{\partial p} = 0 \quad (1.3)$$

$$\frac{\partial^2 V}{\partial p^2} > 0 \quad (1.4)$$

with p the parameter vector. Because this minimum generally cannot be calculated analytically it is to be found numerically. It is therefore hard to ensure the global minimum is found if multiple local minima are present.

To illustrate the existence of local minima in a nonlinear least squares approach the linear deterministic model output (1.1) with b fixed at zero is estimated by:

$$\hat{y}(k+1) = \hat{a}\hat{y}(k) \quad (1.5)$$

It can be easily seen that even for a two steps ahead prediction with y_0 given the model becomes nonlinear in its parameters:

$$\begin{aligned}\hat{y}(1) &= \hat{a}y(0) \\ \hat{y}(2) &= \hat{a}^2y(0)\end{aligned}$$

In the deterministic case shown a minimum is found when $V = 0$. If noise is present (*i.e.* $y(k) = \hat{y}(k) + e(k)$), however, which is generally the case the minimum of the loss function will not be zero. Because the minimum value of V in such cases is not known the conditions (1.3) and (1.4) need to be evaluated. For this particular example (up to two-step ahead prediction) the partial derivative is given by:

$$\frac{\partial V}{\partial \hat{a}} = -2y(1)y(0) + 2\hat{a}y(0)^2 - 4\hat{a}y(0)y(2) + 4\hat{a}^3y(0)^2 \quad (1.6)$$

There are three solutions for the derivative equals zero, *i.e.* a minimum, maximum or inflexion point is found. The solutions are given by:

$$\hat{a} = \begin{bmatrix} a \\ -\frac{1}{2}a + \frac{1}{2}\sqrt{(a^2 - 2)} \\ -\frac{1}{2}a - \frac{1}{2}\sqrt{(a^2 - 2)} \end{bmatrix}$$

If $a > \sqrt{2}$ there are three real valued solutions from which two are minima and one is a maximum. A nonlinear search routine could therefore end up in the wrong minimum.

For a three-step ahead prediction polynomials of order five are already obtained in the derivatives. These polynomials can generally not be solved analytically [Abe26]. Numerical solvers are then required to find the roots. Moreover, if more parameters need to be estimated, nonlinearity increases and more (real valued) minima can be present. Indeed, in general numerical solvers cannot guarantee to find all minima and hence they cannot guarantee to find the global minimum. Frequently used methods like the simplex method or gradient methods [see e.g. DS96] are widely available in commercial computer programs. The nonlinear least squares procedure for estimating parameters is used in this thesis as a reference for alternative estimation methods.

1.3.2 Ordinary least squares

For systems linear in their parameters the ordinary least squares procedure is widely used. Consider again the deterministic system of (1.1) with $y(0)$ given and $a = 0$. If the observations are subject to noise the model can be written as:

$$y(k+1) = bu(k) + e(k+1) \quad (1.7)$$

If e is zero mean white noise and $u(k)$ is known, b can be estimated with ordinary least squares. Because the system is linear in its parameters and the squared 2-norm, *i.e.* $\|e\|_2^2 = e^T e$ with $e = [e(1), e(2), \dots, e(m)]^T$, is minimized a global solution is found by:

$$\hat{b} = (u^T u)^{-1} u^T y \quad (1.8)$$

with $u = [u(0), u(1), \dots, u(m-1)]^T$ and $y = [y(1), y(2), \dots, y(m)]^T$. The basic assumption in this procedure is that the regression vectors are exactly known. The ordinary least squares procedure ensures the global minimum if the regression vectors are not linearly dependant. Detailed information about ordinary least squares can be found in *e.g.* [Lju87, Nor86].

1.3.3 Total least squares

As mentioned before the underlying assumption of the ordinary least squares estimator is that the regressors are known without errors. Frequently, however, errors are not only present in the output vector but also in the data matrix containing the regression vectors. In the ordinary least squares example it was assumed that the error was related to the output $y(k+1)$ and that the input $u(k)$ was known exactly. It is very well possible, however, that also $u(k)$ is subject to errors, *e.g.* $u(k)$ is measured itself. The system is then given by:

$$y(k) + e(k) = b(u(k-1) + w(k-1)) \quad (1.9)$$

This model type is known as errors-in-variables. If the errors e and w are independent and identically distributed the parameter estimation problem is solved by total least squares (TLS). TLS is a fitting technique that compensates for errors in the data and was first introduced by [GVL80]. A graphical representation of TLS estimation compared with OLS estimation is given in Figure 1.2. In stead of minimizing $\|e\|_2$, TLS minimizes $\|[e|w]\|_F$ the frobenius norm of $[e|w]$ which is defined by:

$$\|[e|w]\|_F = \|\Xi\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^2 \xi_{ij}^2} \quad (1.10)$$

The solution of the basic TLS problem is heavily based on the singular value decomposition. For the specific case described above, *i.e.* a single parameter needs to be estimated, the solution is given by:

$$[u|y] = USV^T \quad (1.11)$$

$$\hat{b} = v_{1,2} v_{2,2} \quad (1.12)$$

where $V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}$. An extensive introduction to TLS with algorithms can be found in [VHV91].

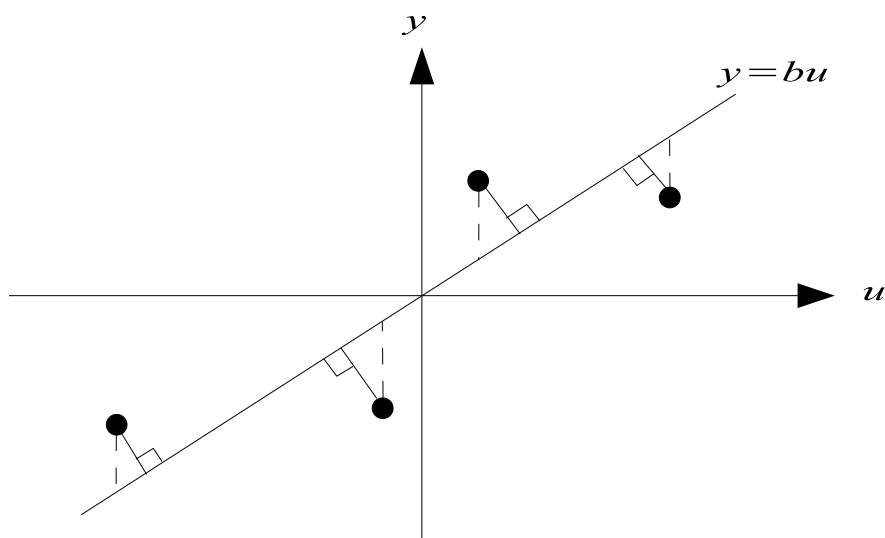


Figure 1.2: Ordinary least squares (---) versus total least squares (—) [GVL80].

Since the introduction of TLS many extensions have been made. The limitation of independently and identically distributed errors e and w was overcome by the introduction of generalized TLS [VHV89]. The dependencies between the errors can be defined in the covariance matrix C and the solution is based on the generalized singular value decomposition [Van76]. If multiple inputs are present, this method also allows the case where some of the inputs are free of errors. If the data matrix, *i.e.* $[y|u]$, is near rank deficient regularization tools *e.g.* [Han94] are needed. Specific methods that are able to regularize TLS are truncated TLS [Fie97] and Tikhonov based regularization [Gol99, BBT06].

For dynamic systems like (1.1) in which all variables are subjected to errors, the error matrix $[E|w]$ becomes structured. This is worked out in the following case:

$$y(k+1) + e(k+1) = a(y(k) + e(k)) + b(u(k) + w(k)) \quad (1.13)$$

For m measurements this leads to the norm of the following structured matrix to be minimized:

$$\left\| \begin{bmatrix} e(1) & e(0) & u(0) \\ e(2) & e(1) & w(1) \\ \vdots & \vdots & \vdots \\ e(m) & e(m-1) & w(m-1) \end{bmatrix} \right\|_F$$

Because in this case the errors $[e|w]$ are no longer independent and the errors are structured the TLS solution is no longer optimal [VHV91]. Several solutions such as constrained TLS [AMH91], structured TLS [Dem93] and structured total least norm [RPG96, VPR96] have been proposed and successfully applied for this specific case. A comparison between these methods can be found in [LVHDM97]. Even in the case

of nonlinearly structured matrices methods are available to cope with these structures [RPG98, LVHDM02]. The methods dealing with structured matrices, however, are all based on iterative algorithms which may suffer from local minima.

1.4 Rational models

A special class of nonlinear models is the class of rational models or models with a polynomial quotient structure. Discrete-time dynamic rational models are in this thesis specified by:

$$x(k+1) = f(Z, p) = \frac{g(Z, p)}{h(Z, p)} \quad (1.14)$$

with $g(\cdot)$ and $h(\cdot)$ polynomials in the elements of Z and p . Furthermore, $Z = (x(k), \dots, x(k-\tau), u(k), \dots, u(k-\tau))$ $k, \tau \in \mathbb{Z}^*$ and $\tau < k$ with τ the time delay. Many models are captured by this class such as biokinetic functions described by the Michaelis-Menten equation. In this thesis the storage model that is used for uncertainty analysis can be confined to this class (see appendix B or chapter 4).

1.5 Research topics

Models of agricultural systems that use weather forecasts to predict their systems behavior need a local forecast for the best performance. However, weather forecasts are only available for specific meteorological weather stations. Furthermore, commercial weather forecasts are only delivered at certain time-intervals. Between two subsequent weather forecasts no updates are available. In addition, many agricultural systems already have installed basic instruments to measure weather data. The first research question therefore is postulated as:

Question 1. *Can data assimilation techniques be used to improve local weather forecasts?*

A model of a *storage facility* of agricultural produce is used in the remainder for analyzing the effect of uncertain weather forecasts. Details of this model can be found in appendix B. For linear systems it holds that given the second statistical moment of the input error the error of the system states can be calculated analytically. If the model state is to be kept within constraints, the uncertainty of the state becomes important. Furthermore, if optimal control is applied the system uncertainty becomes part of the optimal control problem. This results in the following questions:

Question 2. *How can the state uncertainty be integrated in the goal function?*

Question 3. *How are the control variable and state uncertainty affected if the uncertainty of the system states is part of the goal function?*

In case a receding horizon optimal control (RHOC) algorithm is implemented based on a nominal weather forecast the expected costs probably differ from the realized costs based on the actual weather. The same holds for model uncertainty; the realized costs based on the actual system differ from the calculated costs.

Question 4. *What are the effects of weather forecast uncertainty and parameter uncertainty on the calculated costs in a RHOC framework?*

In order to get a good predictive model, model calibration is required. Nonlinear least squares is a general approach to obtain this for nonlinear in the parameter models. Parameter estimation with nonlinear least squares has some drawbacks of which existence of local minima is of most interest in this study. A rational model directly follows from a storage application. Hence, for general rational models that leads to the next problem statements:

Question 5. *Is it possible to linearly reparameterize a rational model such that its parameters can be estimated with linear regression and what effects does it have on the predictive performance of the model?*

Question 6. *Is total least squares a suitable technique to estimate the new parameters obtained with linear reparameterization such that the predictive performance of the predictor increases?*

1.6 Outline of the thesis

This thesis is divided into two parts. In figure 1.1 it can be seen that uncertainty that is inherent to weather forecasting propagates through a nominal model and leads to output uncertainty. The output uncertainty can then be used for risk analysis and used in control. These methods are covered in part I. In order to be able to apply the described methods a good nominal model is needed. In part II emphasis is put on obtaining a good nominal model, *i.e.* a good predictor.

Part 1: Error propagation An effort is made to reduce the uncertainty of the weather forecasts. In Chapter 2 a general framework based on Kalman filtering is presented to adapt weather forecasts to local sites. In chapter 3 this concept has been applied to a greenhouse site with local data available. The next step is to analyze the propagation of the remaining disturbances through the model. In chapter 4 this is done for a storage model of agricultural produce (see appendix B) and an attempt is made to include the uncertain system states into the optimization procedure. The last chapter of this part, chapter 5 concerns the impact of uncertainties in both the model and the disturbances on the calculated costs based on the cost function used in the optimization procedure.

Part 2: Parameter estimation A discrete-time rational storage model (see appendix B) is calibrated by using different parameter estimation techniques. The main goal is to obtain a good predictor. In chapter 6 a linear reparameterization method of discrete-time rational models is proposed in order to uniquely estimate the "new" parameters with ordinary least squares. After backtransformation a new predictor is obtained. A total least squares solution for linearly reparameterized models is presented in chapter 7. The method is further elaborated for rational biokinetic functions in chapter 8.

Conclusions and general discussion Finally, in chapter 9 conclusions from the main research questions are drawn and possibilities for future research are given.

Part I

Uncertainty Analysis

Chapter 2

Adaptive Weather Forecasting using Local Meteorological Information

This chapter has been published as:
Doeswijk, T.G. and K.J. Keesman. Adaptive Weather Forecasting using Local Meteorological Information. *Biosystems engineering*, 91(4):421-431, 2005

2.1 Introduction

Weather forecasts are subject uncertainties. The goodness of a weather forecast [Mur93], however, is not evident as objectives are different among users. Murphy identified three types of goodness: consistency, quality, and value. Consistency, mainly concerns the meteorologist. Quality is related to the match between forecasts and observations. Value, defined by economic benefits, is mainly of interest to users. The types of goodness are related to one another. For instance, if an economic decision is based on a weather forecast, the relation between quality and value is defined by the user and depends on the type of problem.

In general, meteorological parameters such as temperature, rain and global radiation are important for agricultural systems. Anticipating on future conditions is most often needed in these systems. Weather forecasts then become of substantial importance. The uncertainties of weather forecasts have a direct effect on the uncertainty of the system states [Atz95]. If control strategies are used that anticipate to (changing) weather conditions [CBW96, KPL03], the goodness of weather forecasting is related to value. The anticipating control strategies, however, are heavily based on the system model and as such on the quality of the forecast. Therefore, the quality (or accuracy) of the weather forecasts is examined in this chapter.

When very short term weather forecasts are needed, a simple method such as ‘lazy man weather prediction’ [TVWVS96] or a more sophisticated method based on neural networks [CBCdMO02] can be used. However, when forecasts up to a day ahead or more are needed, commercially available weather forecasts are more reliable.

Although work has been undertaken to improve meteorological models [PBBP97] local conditions are not covered by these models. Expert and historical knowledge of a specific location is needed to improve the local weather forecast. In addition, the forecast can be further improved by adaptive techniques using local observations. In previous work, it has been shown that biases of meteorological models can be reduced for short term forecasts (up to 48 hours) [Hom95]. A comparable result has been obtained for maximum and minimum temperature forecasts [GA02]. Both methods are based on prediction of forecast errors. The obtained errors are then added to the available forecasts.

The main purpose of this chapter is to show that local forecasts can be improved by using local meteorological information. This improvement is based on reduction of the standard deviation of the forecast error. The Kalman filter [Gel74] is used to update local weather forecasts, where the covariance matrices can be obtained from actual and historical data. In this chapter, the temperature at research location ‘De Bilt’ in the Netherlands is chosen to demonstrate the procedure. In order to show the wide applicability of the procedure, results of wind speed and global radiation

are included as well.

2.2 Background

2.2.1 Weather data

The weather data normally used for agriculture, are derived from local measurements and short term forecasts. Results are presented for the temperature at 2 m, wind speed and global radiation at location 'De Bilt' in the Netherlands. Data used for analysis range from March 2001 until December 2003.

The local measurements are available with an hourly interval. The weather forecasts become available every six hours at 0100, 0700, 1300 and 1900 coordinated universal time (UTC) and consist of forecasts from 0 to 31 hours ahead with an hourly interval. These external data were extracted from the GFS (global forecast system) model formerly known as AVN (Aviation) model. The model was modified during the period of study several times. Both data are possibly adjusted by a meteorologist.

2.2.2 Diurnal bias corrections

Systematic errors from numerical weather prediction models can be filtered [Hom95, Bha00, GA02]. In this case, the prediction errors are assumed to vary only in a 24 hour context. The estimated prediction errors at forecast times are added to the forecasts independent from the forecast horizon. These corrected forecasts give good results for bias reduction. However, reductions in standard deviations of the forecasts error are, if any, low. Due to homogeneity of the weather in the Netherlands, *i.e.* no large bias is expected, our main focus is to obtain reduction of the standard deviation of the forecast error.

2.2.3 Forecasting system and Kalman filter

Local measurements are used to update the short term forecasts. The updating algorithm uses a linear, time-varying system in state-space form that describes the evolution of the forecasts. Every stochastic l -steps ahead forecast at time instant k is treated as a state variable $x_l(k)$ where $l = 0, \dots, M$ (final forecast horizon). Consequently, $x_0(k)$ represents the actual state, $x_1(k)$ the one-step ahead forecast, etc. The forecasts that become available at time instant k are treated as deterministic input $u(k)$. A discrete-time state-space notation is used to represent the 'forecasting'

system

$$x(k+1) = A(k)x(k) + B(k)u(k) + G(k)w(k) \quad (2.1)$$

$$y(k) = Cx(k) + v(k) \quad (2.2)$$

where $x(k) \in \mathbb{R}^{M+1}$, $u(k) \in \mathbb{R}^{M+1}$ and $y(k) \in \mathbb{R}^p$ with p the number of measurements. Furthermore, it is assumed that the disturbance input $w(k)$ (so called ‘system noise’) and measurement noise $v(k)$ are zero-mean Gaussian random sequences with:

$$E[w(k)] = 0, \quad E[w(k)w^T(k)] = Q \quad (2.3)$$

$$E[v(k)] = 0, \quad E[v(k)v^T(k)] = R \quad (2.4)$$

where $w(k)^T$ denotes the transpose of $w(k)$. The matrices $A(k)$, $B(k)$, C , $G(k)$, Q and R are system dependent and will be defined in the subsequent sections. Furthermore, in what follows the matrices C , Q and R are a constant (time-invariant) matrices whereas A , B and G are time-varying.

In the next sections, the so-called Kalman filter is used to update local weather forecasts. Given a system in state-space form (Eqns 2.1 & 2.2) and a measured output $y(k)$, the Kalman filter estimates the states at time instance k with the smallest possible error covariance matrix P [Kal60]. The following Kalman filter equations for the discrete-time system (Eqns 2.1 & 2.2) are used to estimate the new states (*i.e.* updated l -steps ahead forecasts) when new observations become available:

$$\hat{x}(k+1|k) = A(k)\hat{x}(k|k) + B(k)u(k) \quad (2.5)$$

$$P(k+1|k) = A(k)P(k|k)A(k)^T + G(k)QG(k)^T \quad (2.6)$$

$$K(k+1) = P(k+1|k)C^T [CP(k+1|k)C^T + R]^{-1} \quad (2.7)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)[y(k+1) - C\hat{x}(k+1|k)] \quad (2.8)$$

$$P(k+1|k+1) = P(k+1|k) - K(k+1)CP(k+1|k) \quad (2.9)$$

where $\hat{x}(k+1|k)$ denotes the estimate of state x at time instant $k+1$ given the state at k , and $A(k)^T$ is the transpose of $A(k)$. Furthermore, $K(k+1)$, known as Kalman gain, denotes the weighting matrix related to the prediction error $[y(k+1) - C\hat{x}(k+1|k)]$.

2.2.4 Covariances

In the Kalman filter covariance matrices P , Q and R play an important role. As can be seen from Eqns (2.6), (2.7) and (2.9), given an initial covariance matrix $P(0)$ (typically $P(0) = 10^6 I$ with I the identity matrix) the corrected covariance matrix $P(k+1|k+1)$ can be calculated. The matrix R is related to the measurement noise. It can mostly be determined from the sensor characteristics. The key problem here is how to choose Q , the covariance matrix related to the disturbance inputs or system noise. This system noise covariance matrix of the short-term forecasts can be determined from historical data by comparing forecasts to observations.

2.3 Local adaptive short term forecasting

The general system (2.1)-(2.2) is further elaborated by defining the system variables and matrices. The dimension of both the state (x) and input (u) vectors is 32. The output (y) is the observed meteorological parameter and is a scalar. The system matrix A is chosen such that at every time instance when a new observation becomes available, the states are moved up one place. As a consequence, the state vector x always represents the forecast horizon (from 0 to $31 - k + 1$ hours ahead). Subsequently, the effective system reduces as long as there are no new external forecasts available. Given the assumption that new external forecasts are better than updated old forecasts, the old states are reset and the new initial state is fully determined by the new forecast as soon as new one becomes available at time instance k^* . The system matrices are defined as follows:

$$A(k) = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix} \quad k \neq k^*, k \in \mathbb{N} \quad (2.10)$$

$$A(k) = 0 \quad k = k^*, k \in \mathbb{N} \quad (2.11)$$

$$C = [1 \quad 0 \quad \dots \quad 0] \quad (2.12)$$

The system noise or disturbance input is only relevant at time instant $k = k^*$, *i.e.* when a new external forecast becomes available. The time-varying system matrices B and G are defined as follows:

$$B(k) = G(k) = I \quad \forall k = k^*, k \in \mathbb{N} \quad (2.13)$$

$$B(k) = G(k) = 0 \quad \forall k \neq k^*, k \in \mathbb{N} \quad (2.14)$$

where $A, B, G \in \mathbb{R}^{(M+1) \times (M+1)}$ and $C \in \mathbb{R}^{M+1}$ ($M = 31$). The Kalman filter can now be introduced to use local measurements for improving short term weather forecasts. The stability of the filter is proven in Appendix A.

2.4 Results

From the forecast and observation files the average forecast error (*i.e.* forecasts - observations) of the predicted meteorological parameter was obtained and the covariance matrix Q of the forecast error was calculated. The data used to obtain this matrix covered the period from March 1, 2001 until March 1, 2002. As an example the covariance matrix of the 2 m temperature is given in figure 2.1. On the diagonal

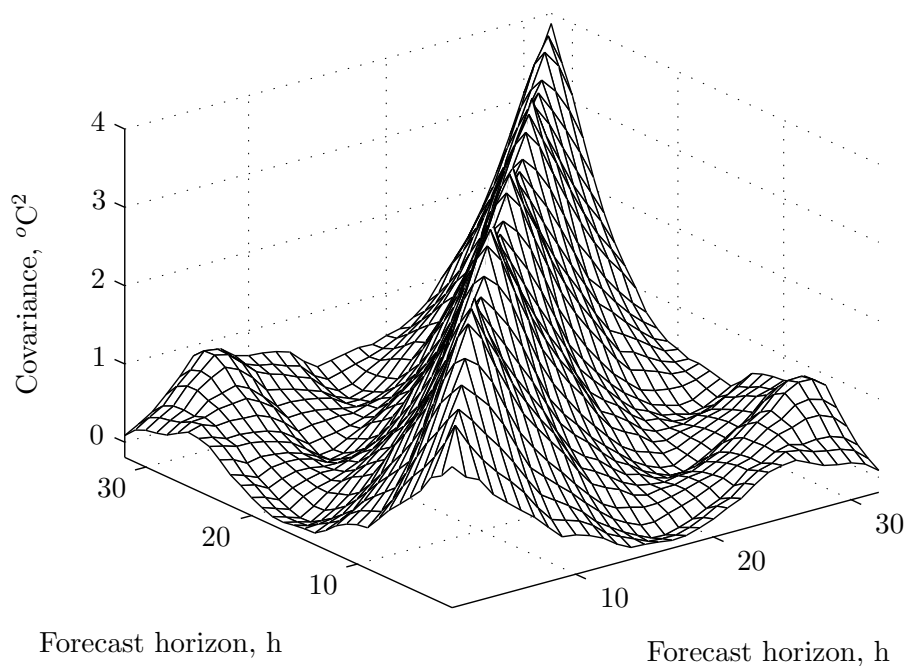


Figure 2.1: Covariance matrix of the short term forecast error

the variance of the forecast error is given along the forecast horizon. Furthermore, Q is symmetrical, *i.e.* $\text{cov}(x, y) = \text{cov}(y, x)$.

For all meteorological parameters the same procedure was used. The Kalman filter was run over the period December 1, 2002 until November 1, 2003. The updated forecasts were compared with the original external forecasts after a new external input entered the system and a measurement was taken.

For the 2 m air temperature a comparison is made between the previously described method and the method described in [Hom95] which is based on diurnal bias corrections.

2.4.1 Air temperature at 2 m

Local adaptive forecasting

The variance of the observational noise R (see Eqn 2.4) was assumed to be $0.1 \text{ } ^\circ\text{C}^2$. Given A, B, C, G (Eqns 2.10-2.14), Q and R , the Kalman filter could then be implemented. The results are presented in figure 2.2. From this figure it can be seen that using a single observation, forecasts up to six hours ahead can be positively adjusted. The periodicity that is observed is due to interpolation; the original model output, given at a 3 hourly interval, is interpolated by the weather agency to an hourly interval.

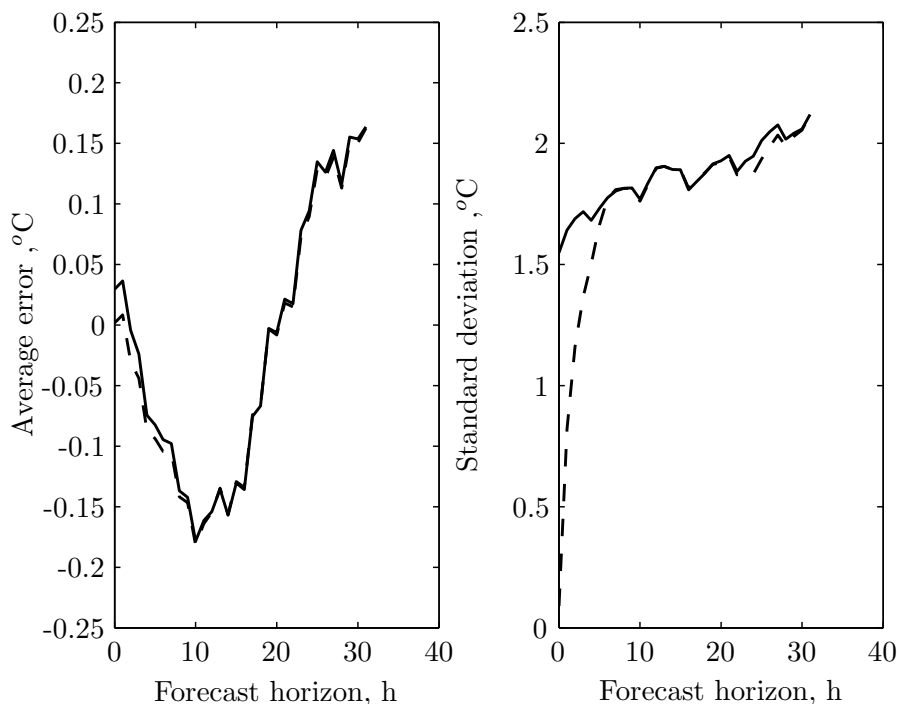


Figure 2.2: Average forecast error and standard deviation of the original forecasts (—) and the local adaptive filtered forecasts (---) for the 2 m temperature

Diurnal bias correction

There are two important design criteria for this Kalman filter. First, the covariance matrix $Q = \sigma_w^2 W$ should be chosen where W is the correlation matrix and σ_w^2 the variance. In this simulation experiment the correlation matrix is chosen similar as described by [Hom95], *i.e.* exponentially decaying until $k + 12$ and then rising again until $k + 24$. No difference in correlation is made between successive hours. The second design criterion is the value σ_w/V where V denotes the standard deviation of the observation. The σ_w/V -ratio appears to be crucial. When the chosen ratio is too large, the standard deviation of the forecast error increases compared to the original forecasts after a few forecast hours. When it is too small, standard deviations are hardly reduced. In this study, the optimum (in terms of minimum average standard deviation of the forecast horizons) is searched by a line search procedure. The results of diurnal bias correction with a σ_w/V ratio of 0.01 are given in figure 2.3. It can be seen that a structural prediction error still remains and that standard deviations can be slightly reduced for the whole forecast range.

Local adaptive forecasting on bias corrected forecasts

While the local adaptive forecasting method works specifically on the short term (+6 hours), the diurnal bias correction gains on the longer forecast horizons. A logical

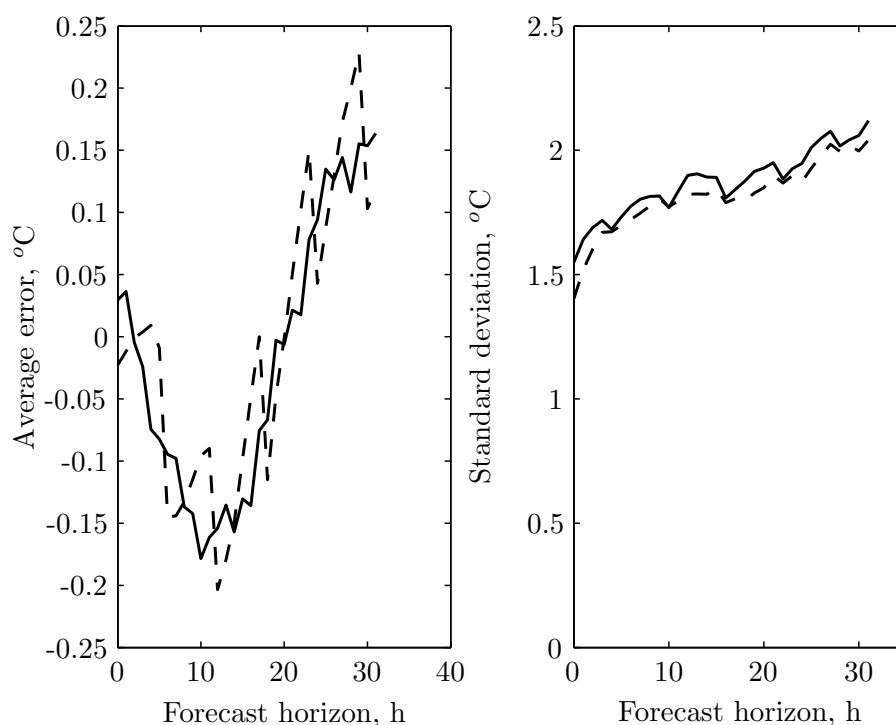


Figure 2.3: Average forecast error and the standard deviation of the original forecasts (—) and of the diurnal bias corrected forecasts (---) of the 2 m temperature

combination would be to first implement the diurnal bias correction and then use the local adaptive forecasting technique on the corrected forecasts. The results are shown in figure 2.4. It can be seen that local adaptive forecasting has the same behaviour when it is used on prefiltered data as it has on the original data. In addition, the effect of decreasing standard deviation lasts longer when prefiltered data are used (6 *vs* 9 hours).

2.4.2 Wind speed

In comparison with the 2 m temperature the system noise covariance matrix of the diurnal bias correction is taken equal and the σ_w/V was found to be optimal around 0.02. The variance of the observation noise of the wind speed in the local adaptive short term system was assumed to be $0.1 \text{ m}^2\text{s}^{-2}$. The results of all three procedures are summarised in figure 2.5. Again, both methods increase forecast performance when used separately. The bias reduction is in this case much more pronounced than for the temperature forecast. When the methods are combined, an extra increase in performance can be seen. The main part, of course, is found at the very short term (+6 hours) but it remains on a small scale for the whole forecast horizon.

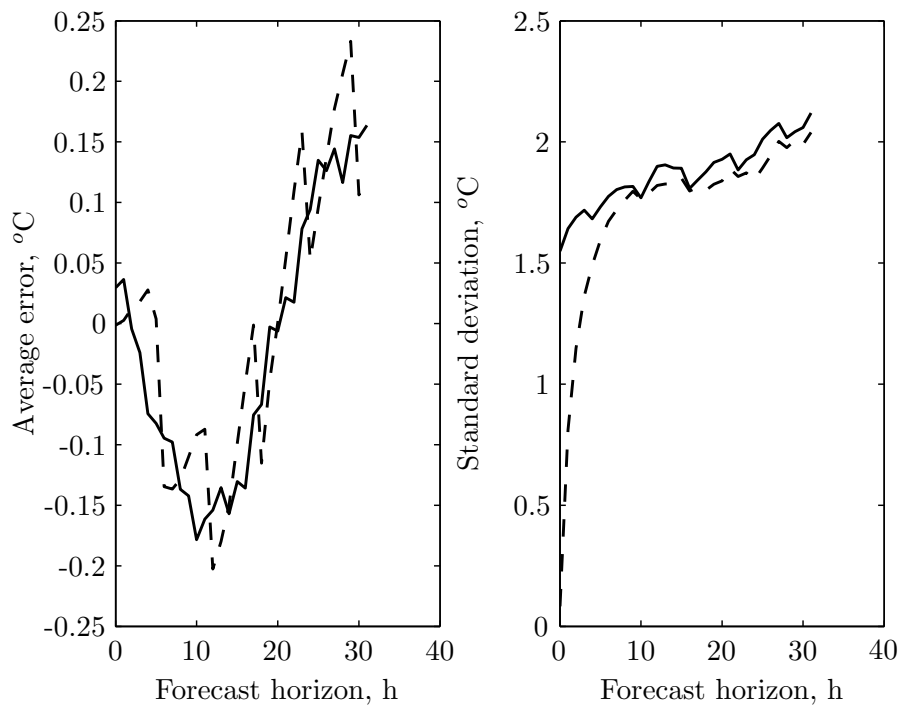


Figure 2.4: Average forecast error and the standard deviation of the original forecasts (—) and of the combined diurnal bias corrected and local adaptive filtered forecasts (---) of the 2 m temperature

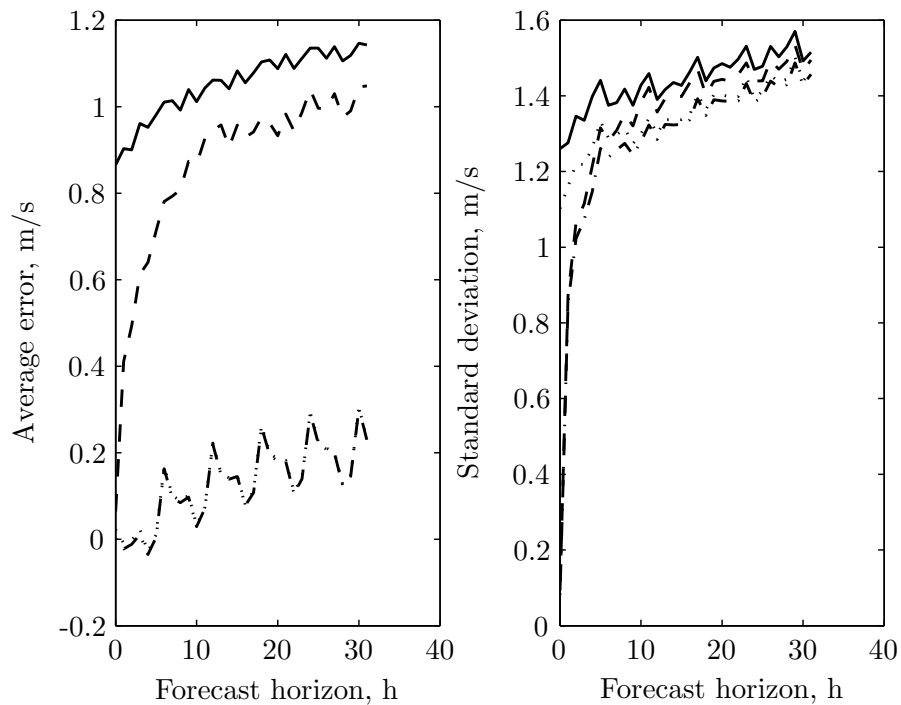


Figure 2.5: Average forecast error and the standard deviation of the original forecasts (—), diurnal bias corrected forecasts (\cdots), local adaptive filtered forecasts (---) and of the combined diurnal bias corrected and local adaptive filtered forecasts (\dashdot) of the wind speed

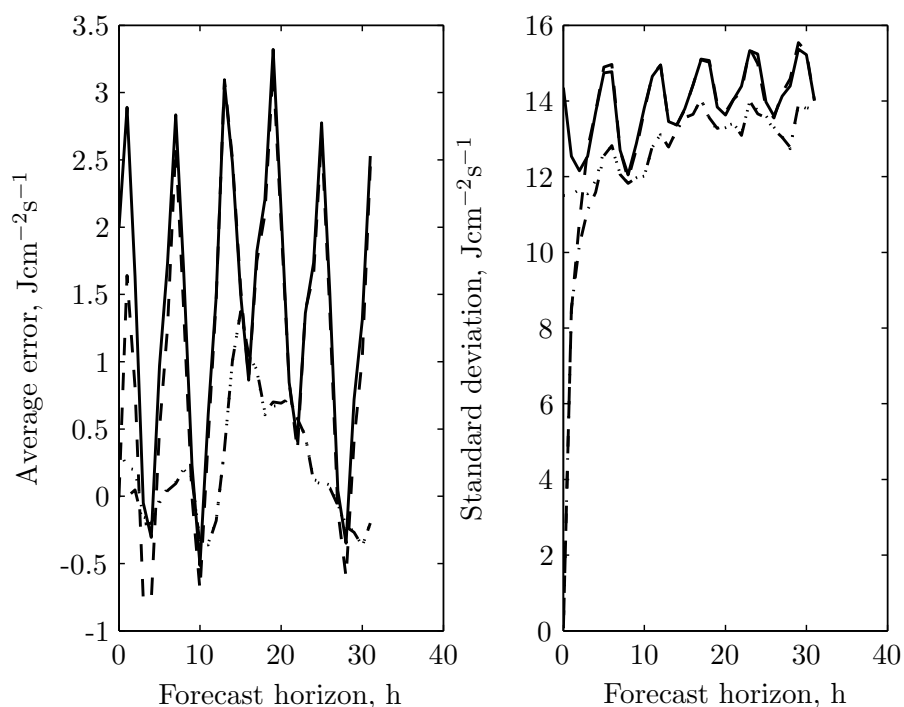


Figure 2.6: Average forecast error and the standard deviation of the original forecasts (—), diurnal bias corrected forecasts (\cdots), local adaptive filtered forecasts (---) and of the combined diurnal bias corrected and local adaptive filtered forecasts (\dashdot) of the global radiation

2.4.3 global radiation

Diurnal bias correction was implemented for temperature and wind speed. However, for global radiation, the covariance matrix must be adjusted. At night, no radiation is available and so no correlation is present. The length of the nights also vary during the year. For simplicity, it is assumed that for the whole year no correlation is needed between 1700 and 0500 UTC. The remaining correlations are kept the same as for the temperature case. The optimum σ_w/V ratio appeared to be around 0.02. For local adaptive forecasting, no special covariance matrices were used. In figure 2.6 the average error and standard deviation for all procedures are given. The twelve hour dependency is clearly seen by the large periodic variations for the average error of the original forecast. The diurnal bias correction works well for the 0-12 hour horizon. From this point the bias increases until 24 hour forecast horizon where it again approaches zero. Furthermore, the same results are seen for global radiation as for temperature and wind speed. The standard deviation is reduced for both procedures. In this case the local adaptive forecasting procedure works only until a 3 hour forecast horizon. However, when both procedures are combined, the standard deviation is reduced for more than 5 hours horizon.

2.5 Discussion

The improvement of weather forecasts with the Kalman filter for the local adaptive forecasting procedure largely depends on the choice of covariance matrices Q and R . The period over which the covariance matrices are determined plays an important role. In this study the covariance matrices are defined over a period of a year. Herewith seasonal effects are neglected. When seasonal effects on variability of the forecast error are suspected Q and R must be determined from the seasons of a previous year. Apart from the seasonal effects, changes in weather forecasting models can effect the variability of the forecast error. Consequently, a ‘window’ for the covariances can be considered. The weather forecasting models were adjusted frequently during the period of research [NCE05]. However, the Kalman filter still showed a good performance. As an alternative to the stochastic filtering approach one may also consider an unknown-but-bounded error approach [Sch73, Kee97].

In Section 2.3, a low variance is assumed for the local observation. This might be true for the measurement device itself but the variance also depends on how and where the device is installed. The measurement represents the weather disturbance input of the real system under study. Weather forecasts are generally made only for a few meteorological stations. As both measurements and forecasts used in this chapter were dedicated to the same specific meteorological location, larger deviations between forecasts and measurements are expected for specific agricultural systems. This is yet another reason why an extra improvement of the forecast is suspected by using the Kalman filter in relation to the real system.

When considerable biases are present, the diurnal bias correction gives good results for bias reduction. When the σ_w/V ratio is chosen correctly, also a reduction in standard deviation is obtained. This reduction can then last for the whole forecast range. However, when the ratio chosen is too large then a loss in performance is seen. When values are too low no change in forecasts is apparent. In this study a line search procedure was performed to find the optimum ratio. However, the optimum value for this ratio highly depends on the season. For temperature, an optimum value approaching 0 was found for the winter while during summer the optimum value increased to 0.03. A time-varying ratio based on past forecast errors gave satisfactory results for maximum and minimum daily temperatures [GA02]. However, the problem remains because the value is always obtained from past data. A time-varying ratio which is based on future uncertainties of the forecasts, *i.e.* uncertainties given by ensemble forecasts, or expected change in weather conditions is worthwhile to investigate.

The correlation matrix must be determined for every meteorological parameter. For global radiation, this matrix must also depend on the time of the year. This

matrix can be found for every meteorological parameter at a specific location by time series analysis.

When both procedures are run in series, the results show that both bias correction as well as standard deviation reduction are obtained, *i.e.* the methods are complementary to each other. It is worthwhile to investigate whether both methods can be integrated into a single system. For instance, the diurnal bias correction supposes a full correlation of the diurnal pattern, *i.e.* a certain error obtained now will also be present tomorrow. A more valid assumption is that this relation is exponentially decaying. As a result the σ_w/V ratio chosen can be larger.

2.6 Conclusions

The improvement of local weather forecasts using local measurements is demonstrated in this chapter for both diurnal bias correction and local adaptive forecasting. If only very short term forecasts (several hours ahead) are needed local adaptive forecasting is the proper method to use. If only bias correction is sufficient, the diurnal bias correction procedure is the best choice. Generally, first a diurnal bias correction followed by the local adaptive forecasting procedure is recommended because this gives the best results in both bias and standard deviation reduction. Slow and fast dynamics of the weather are then incorporated.

Chapter 3

Improving Site-Specific Weather Forecasts for Agricultural Applications

This chapter has been published as:
Doeswijk, T.G. and K.J. Keesman. Improving Local Weather Forecasts for Agricultural Applications. In *Signal and Image Processing, Second IASTED Intern. Multi-Conference on Automation, Control, and Information Technology*, Novosibirsk, Russia, pages 421-431, 2005

3.1 Introduction

In many agricultural systems all kinds of weather variables, such as temperature, radiation and rain, have a dominant effect on the systems' behavior. Weather input variables are not only a disturbance to the system but are also a resource, (*e.g.* global radiation drives plant growth). For maximization of plant production some of the control inputs should closely follow changes in weather conditions, *e.g.* CO₂-dosing in green-houses should anticipate on changes of global radiation. Therefore, when controlling agricultural systems, weather forecasts can be of substantial importance, especially when anticipating control strategies are used. [see *e.g.* CBW96, KPL03]. If forecasts of less than one hour are used, the so-called "lazy man" weather prediction, where the forecast is chosen equal to the most recently measured value, seems to be reasonable [TVWVS96]. If, however, the forecast horizon increases, preferably commercial forecasts should be used.

It is well known that, because of the chaotic behavior of the atmosphere, weather forecasts can be rather uncertain. This uncertainty increases as the forecast horizon increases. Many efforts are taken by meteorologists to improve the quality of the weather forecasts, but forecasts of weather variables remain uncertain. For instance, the 2 meter temperature forecast has a standard deviation of 1.5 °C for the zero-hour ahead forecast in "De Bilt", The Netherlands (see figure 2.2).

Previous research has been done to improve local weather forecasts. For instance, biases can be largely removed from meteorological model outputs [Hom95]. This bias reduction procedure uses a Kalman filter that predicts a diurnal forecast error. This forecast error then has to be added to the original forecast. The procedure is further called "diurnal bias correction" (DBC). In specific cases like mountainous areas and places surrounded by seas [GA02] this DBC proved to reduce the bias drastically. Standard deviations of the forecast error however, are not reduced with this procedure. However, as was shown in chapter 2, reduction of the standard deviation of the forecast error can be obtained as well. The standard deviation reduction procedure uses local measurements in a Kalman filter to update the forecast. This procedure is further referred to as "local adaptive forecasting" (LAF). The best performance was obtained by applying both procedures: first bias reduction, then standard deviation reduction.

The presented studies in [Hom95] and in chapter 2 used meteorological data of specific meteorological stations. Measurements at these meteorological stations need to fulfill specific meteorological requirements [WMO96]. In general, local measurements for agricultural systems do not fulfill the meteorological requirements, but they do represent the actual local situation. Furthermore, the forecasts from weather agencies are specific for some predetermined places, but most often the specific agri-

cultural systems are not located at these places. At last, local circumstances such as soil properties, different altitudes, presence of water resources (lake, sea) etc., have an effect on local weather conditions and thus should be accounted for by the local weather forecast. Therefore it is assumed that weather forecasts for agricultural systems can be improved for both bias and standard deviation.

The purpose of this chapter is to show that local forecasts for agricultural systems can be improved by using local measurements. This improvement is based on reduction of the *bias* as well as the *standard deviation* of the forecast error.

In section 3.2 theoretical background is given and the analyzed data are explained. In section 3.3 the results of three weather variables are presented: temperature, wind speed and global radiation. The results are discussed in section 3.4. Finally, some conclusions are presented in section 3.5.

3.2 Background

3.2.1 Kalman filtering

Both procedures, DBC and LAF, are based on a discrete-time state-space system representation of either forecast errors or forecasts and use the Kalman filter as the main algorithm to update weather forecasts.

The stochastic discrete-time state-space system used in both procedures is given by:

$$x(k+1) = A(k)x(k) + B(k)u(k) + G(k)w(k) \quad (3.1)$$

$$y(k) = C(k)x(k) + v(k) \quad (3.2)$$

where $x(k) \in \mathbb{R}^{M+1}$, $u(k) \in \mathbb{R}^{M+1}$, with M the maximum forecast horizon, and $y(k) \in \mathbb{R}^p$, with p the dimension of the actual output vector. It is assumed that the disturbance input $w(k)$ (so called "system noise") and measurement noise $v(k)$ are zero-mean Gaussian random sequences with:

$$E[w(k)] = 0, \quad E[w(k)w^T(k)] = Q \quad (3.3)$$

$$E[v(k)] = 0, \quad E[v(k)v^T(k)] = R \quad (3.4)$$

The matrices $A(k), B(k), C(k), G(k), Q$ and R are system dependent and will be defined in the subsequent sections.

As mentioned before, the so-called Kalman filter is used to update local weather forecasts. For an elaborate description of the Kalman filter we refer to [Gel74]. Here, the algorithm is briefly outlined. Given a system in state-space form (3.1)-(3.2) with noise properties (3.3)-(3.4) and a measured output $y(k)$, the Kalman filter estimates the states at time instance k with the smallest possible error covariance matrix. The

following Kalman filter equations for the discrete-time system (3.1)-(3.2) are used to estimate the new states (updated forecast errors in DBC or forecasts in LAF) when new observations become available:

$$\hat{x}(k+1|k) = A(k)\hat{x}(k|k) + B(k)u(k) \quad (3.5)$$

$$P(k+1|k) = A(k)P(k|k)A(k)^T + G(k)QG(k)^T \quad (3.6)$$

$$K(k+1) = P(k+1|k)C^T [CP(k+1|k)C^T + R]^{-1} \quad (3.7)$$

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)[y(k+1) - C\hat{x}(k+1|k)] \quad (3.8)$$

$$P(k+1|k+1) = P(k+1|k) - K(k+1)CP(k+1|k) \quad (3.9)$$

where $\hat{x}(k+1|k)$ denotes the estimate of state x at time instant $k+1$ given the state at k , and $A(k)^T$ is the transpose of $A(k)$. Furthermore, $K(k+1)$, known as Kalman gain, denotes the weighting matrix related to the prediction error $[y(k+1) - C\hat{x}(k+1|k)]$.

3.2.2 Diurnal bias correction

It has been shown by [Hom95] that systematic errors from numerical weather prediction models can be largely removed. A brief outline of the algorithm is given.

The basic assumption is that the prediction errors are assumed to vary only in a 24 hour context. The states $x_1 \cdots x_{24}$ represent the forecast errors at times from 0000 UTC until 2300 UTC. No input is present in this system. The output y is defined by the measurement at a specific time and is a scalar. The system matrices are given by: $A = I$, $G = I$ and C is time-varying *e.g.* $C = [1 \ 0 \cdots 0]$ at 0000 UTC, $C = [0 \ 1 \ 0 \cdots 0]$ at 0100 UTC *etc.* The Kalman filter matrices Q and R are time-invariant where Q is a symmetric matrix with ones on the diagonal, premultiplied by the variance σ_w^2 and $R = V^2$ the variance of the measurement error. The σ_w/V -ratio determines the update rate. The optimal σ_w/V -ratio can vary between meteorological stations and can vary between weather variables (see [Hom95] and chapter 2). The initial covariance matrix $P(0)$ is typically chosen as: $P(0) = 10^6 I$, with I the identity matrix. The estimated prediction errors ($\hat{x}(k+1|k+1)$) are added to the external forecasts independent of the forecast horizon.

3.2.3 Local adaptive forecasting

Local measurements are used to update the short term forecasts. The updating algorithm as described in section 2.3 uses a linear, time-varying system in state-space form that describes the evolution of the forecasts. Every stochastic l -steps ahead forecast at time instant k is treated as a state variable $x_l(k)$ where $l = 0, \dots, M$, with M the maximum forecast horizon. Consequently, $x_0(k)$ represents the actual state, $x_1(k)$ the one-step ahead forecast, *etc.* As a consequence, the state vector x

always represents the forecast horizon (from 0 to $M - k$ hours ahead). Subsequently, the effective system dimensions reduce as long as there are no new external forecasts available. The external forecasts that become available at time instant k^* are treated as deterministic input(s) $u(k)$. Given the assumption that new external forecasts are better than updated old forecasts, the old states are reset and the new initial state is fully determined by the new forecast, *i.e.* $x(k^* + 1) = u(k^*)$. The output $y(k)$ is the observed weather variable and is a scalar. The system matrices are then defined by

$$A(k) = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix} \quad k \neq k^*, k \in \mathbb{N} \quad (3.10)$$

$$A(k) = \mathbf{0} \quad k = k^*, k \in \mathbb{N} \quad (3.11)$$

$$C = [1 \ 0 \ \dots \ 0] \quad (3.12)$$

The system noise or disturbance input is only relevant at time instant $k = k^*$, *i.e.* when a new external forecast becomes available. The time-varying system matrices B and G are defined as follows:

$$B(k) = G(k) = I \quad \forall k = k^*, k \in \mathbb{N} \quad (3.13)$$

$$B(k) = G(k) = \mathbf{0} \quad \forall k \neq k^*, k \in \mathbb{N} \quad (3.14)$$

where $A, B, G \in \mathbb{R}^{(M+1) \times (M+1)}$ and $C \in \mathbb{R}^{1 \times (M+1)}$.

The Kalman filter matrices Q and R must still be specified. The measurement noise covariance matrix R may be found from the sensor characteristics. The key problem is how to choose Q , the covariance matrix related to the disturbance inputs or system noise. In figure 2.1 an example of this system noise covariance matrix of the short-term forecasts was presented. This matrix has been determined from historical data by comparing forecasts with observations. The initial covariance matrix $P(0)$ is typically chosen as: $P(0) = 10^6 I$, with I the identity matrix.

3.2.4 Weather data

Three different weather variables, that are most relevant for greenhouse systems, are studied: temperature, wind speed and global radiation. Data is obtained from January 1, 2002 until June 31, 2002 and January 1, 2003 until June 31, 2003. The origin of forecasts and local measurements are specified by:

short term forecasts The commercial weather agency Weathernews Benelux delivered forecasts for location ‘Deelen, The Netherlands’ (see figure 3.1). These

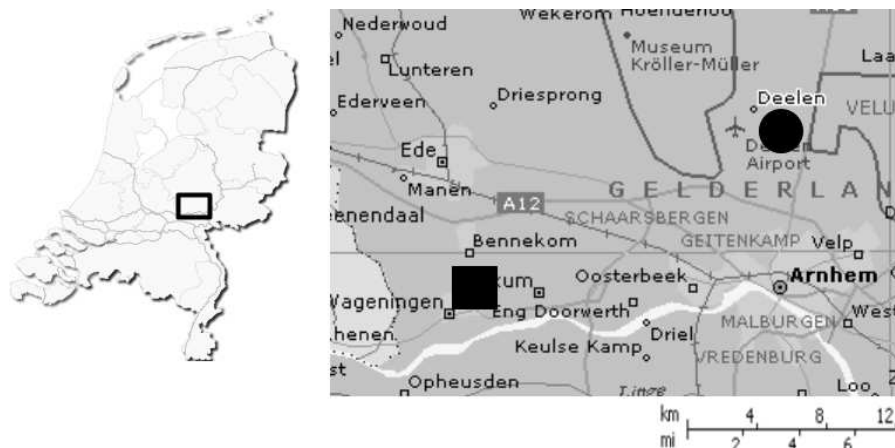


Figure 3.1: Indication of forecast (●) and measurement (■) location within the Netherlands

data become available every six hours and consist of forecasts from 0 to 31 hours ahead with an hourly interval. These external data are extracted from the GFS model (National Weather Service). The data are possibly adjusted by a meteorologist. The data are made available at 0100, 0700, 1300 and 1900 UTC.

local measurements These data are obtained from a greenhouse in ‘Wageningen, The Netherlands’ located at about 20 km from ‘Deelen’ (see figure 3.1). Measurements were stored with a 2 minute interval. The hourly averages were calculated and used for analysis.

3.3 Results

Forecast data are compared with local observations. For the DBC, matrix Q describes the correlation of the forecast error over a 24 hour horizon. In this experiment it is chosen similar as described by [Hom95], *i.e.* exponentially decaying until $t + 12$ and then rising again until $t + 23$ ($e^{-0.0744t}$ with $0 \leq t \leq 12, t \in \mathbb{N}$), for each weather variable. The optimal σ_w/V -ratio is found with a line search procedure. Optimality in this case is defined by: minimum average standard deviation (σ_e) of the forecast error

$$\min_{\sigma_e/V} \frac{1}{M+1} \sum_{i=0}^M \sigma_i(\sigma_w/V) \quad (3.15)$$

The optimal ratio is calculated over the period January 1, 2002 until June 31, 2002. If W is chosen as 1, and σ_w/V is given, R can be calculated.

The covariance matrix of the forecast error Q of the LAF procedure is calculated for each weather variable with data from January 1, 2002 until June 31, 2002.

As both DBC and LAF are complementary the procedures are run consecutively after a new measurement becomes available. The procedure is run over the period January 1, 2003 until June 31, 2003. The updated forecasts (DBC+LAF) are compared with the original forecasts provided by the weather agency.

The results related to a specific weather input variable contain the optimal σ_w/V -ratio for DBC. In addition, the assumed measurement noise covariance matrix R used in LAF is given. Furthermore, the forecast error, *i.e.* forecast - observation, is calculated and the average forecast error and the standard deviation of the forecast error are presented.

3.3.1 temperature

The results are given in figure 3.2 with a σ_w/V -ratio of 0.011 for DBC and the variance of the measurement noise R in LAF of $0.1 \text{ } ^\circ\text{C}^2$. In figure 3.2 it can be seen that the bias is reduced for each forecast horizon. The standard deviation is reduced for each forecast horizon but especially up to 10 hours ahead this reduction is clear.

3.3.2 wind speed

The σ_w/V -ratio used in the DBC was 0.042. The variance of the observation noise of the wind speed in the local adaptive short term system is assumed to be $0.1 \text{ (ms}^{-1}\text{)}^2$. The results are summarized in figure 3.3. The bias is almost completely removed compared to the original forecasts for each forecast horizon. Again, the standard deviation is lowered. The reduction of standard deviation in this case clearly remained until the maximum forecast horizon.

3.3.3 global radiation

For global radiation the same LAF procedure as for temperature and wind has been implemented but with $R = 10 \text{ (Js}^{-1}\text{m}^{-2}\text{)}^2$. However, in DBC the covariance matrix should be adjusted. At night no radiation is available and so no correlation is present. The length of the nights should also vary during the year. For reasons of simplicity the yearly variation is neglected and hence it is assumed that for the whole year no correlation is needed between 1700 and 0500 UTC. The remaining correlations are kept the same as for the temperature case. The optimum σ_w/V -ratio appeared to be around 0.010. In figure 3.4 the average error and standard deviation for the original forecast and the adjusted forecast are presented. Overall, the bias is reduced. However, only in the first few hours this can be seen clearly. Furthermore, the standard deviation is reduced for each forecast horizon, particularly in the first 5 hours. In addition to this, the peaks are largely removed.

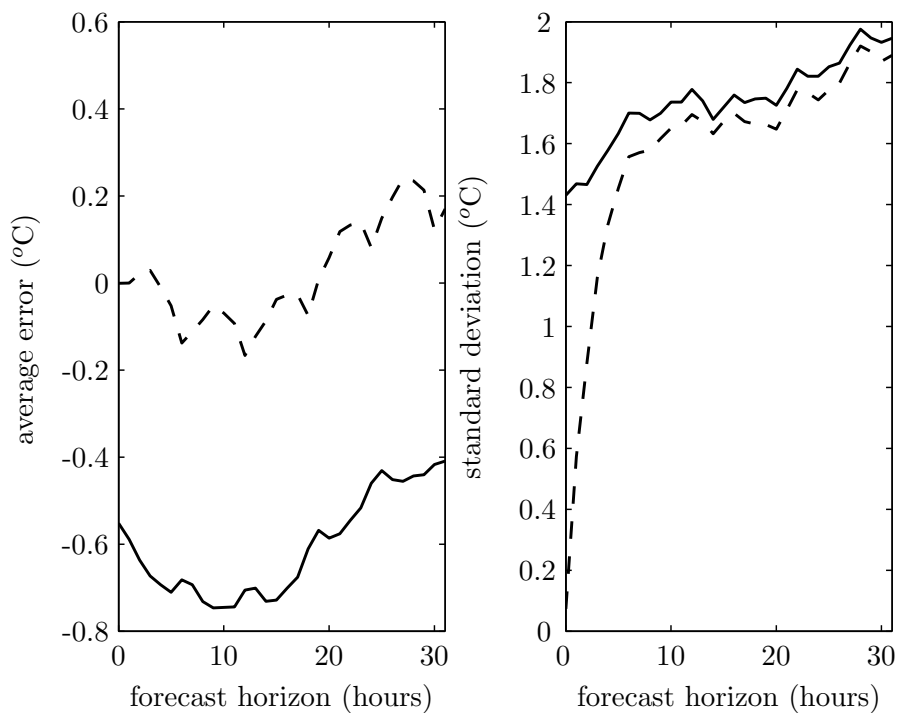


Figure 3.2: Average forecast error and the standard deviation of forecast error of the original forecasts (—) and of the adjusted forecasts with DBC+LAF (---) of the temperature

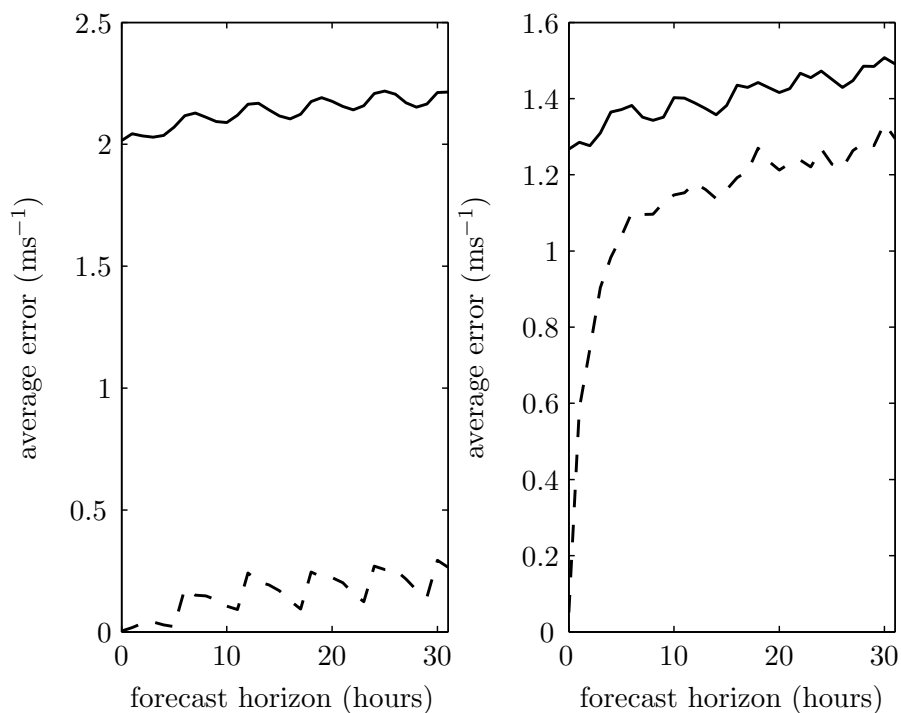


Figure 3.3: Average forecast error and the standard deviation of forecast error of the original forecasts (—) and of the adjusted forecasts with DBC+LAF (---) of the wind speed

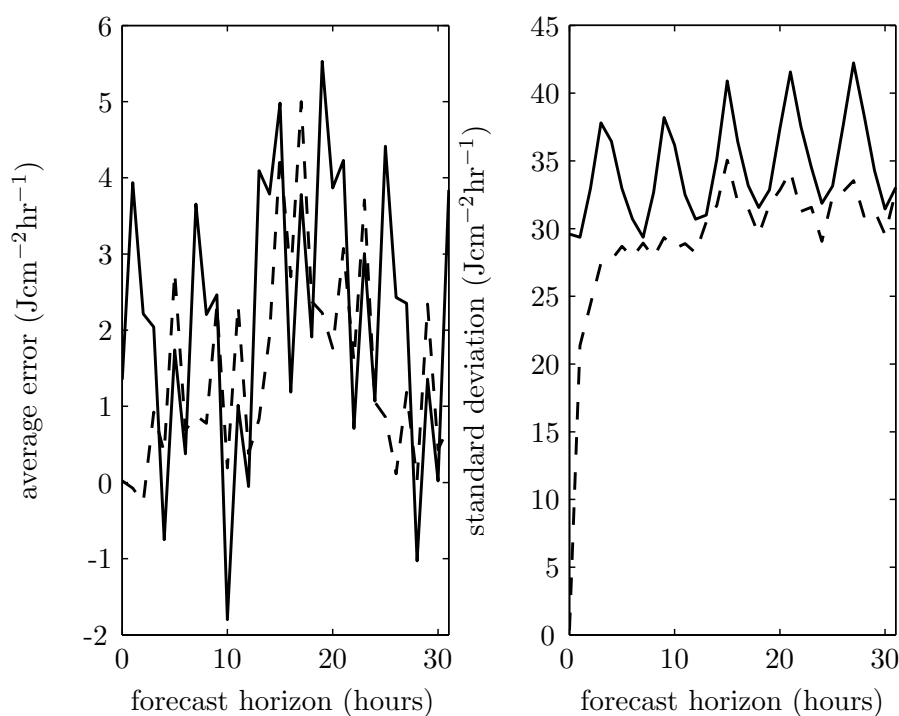


Figure 3.4: Average forecast error and the standard deviation of forecast error of the original forecasts (—) and of the adjusted forecasts with DBC+LAF (---) of the global radiation

3.4 Discussion

Bias of the forecast error is often present even if the forecast is related to a specific meteorological measurement location. Therefore, it can be expected that for local measurements, as in agricultural systems, bias is present and is probably larger than for the meteorological station from which the forecast originated. In figures 3.2 - 3.4 the bias is clearly present.

When considerable biases are present the DBC works quite well for bias reduction as can be seen in figures 3.2 and 3.3. The σ_w/V -ratio appears to be crucial for the performance of DBC. If the σ_w/V -ratio is chosen too large then a loss in performance is observed, *i.e.* the standard deviation increases for larger forecast horizons. For instance, values chosen by [Hom95], *i.e.* $\sigma_w/V = 0.06$, does not satisfy for temperature in our case. Too low values will lead to negligible changes in the forecasts. Furthermore, the value for this ratio depends on the weather type and change in weather conditions. A time-varying ratio was proposed by [GA02]. The problem remains, however, because the value is always obtained from past data.

The line search procedure with optimality criterion (3.15) to minimize the standard deviation of the forecast error will result in an average standard deviation equal to or lower than the average standard deviation of the original forecast. The bias,

however, is then not necessarily minimized. Therefore, other minimization criteria, such as minimum mean square error or minimum bias, must be used according to the defined purpose. It should be noted that the calculated optimal σ_w/V -ratio is kept constant for each year. In practice, the optimal σ_w/V -ratio changes every year. From the results in figures 3.2-3.4, however, it can be seen that the calculated optimal ratio is applicable for the following year.

The correlation in the covariance matrix Q in DBC has to be determined for every weather variable. For global radiation, this matrix should also depend on the time of the year because daylight duration can change largely within a year. In this case the matrix Q must be time-varying.

The LAF procedure largely depends on the choice of covariance matrices Q and R . The matrix Q can be found for every weather variable at a specific location by time series analysis. The period over which the covariance matrices are determined may play an important role. In our study we defined the covariance matrices over a period of six months. This was the same period of the year over which the Kalman filter was run. When seasonal effects on variability of the forecast error are suspected one can choose to determine Q and R from the seasons of a previous year, *e.g.* month by month, quarter by quarter etc. Apart from the seasonal effects, meteorological models may have a more significant effect on the variability of the forecast error. Consequently, one could consider to use a "window" for the covariances.

In this chapter a low variance is assumed for the local measurement. This might be true for the measurement device itself but the variance also depends on how and where the device is installed. The measurement should represent the weather disturbance input of the real system under study. For local measurements devices it is crucial that they are properly calibrated and maintained. Furthermore, the measurement device must be situated on a proper place, *e.g.* a temperature device should not be exposed to direct sunlight. As ambient measurements are frequently used to control greenhouse climates, it is acceptable to update forecasts with local measurements. On the other hand, if it is known that local measurements are unreliable, it is worthwhile to investigate if the measurements can be updated with weather forecasts to generate more reliable measurements. As an example the observations of the meteorological stations of 'Deelen' and 'Wageningen' and the local observations of 'Wageningen' are compared in table 3.1. The meteorological station 'Wageningen' is located at about 4 km of the local observations. It can be seen in this table that all differences are quite similar. Intuitively, the difference between local measurements and 'Wageningen' meteorological station is expected to be smaller than the others because of the small distance. Apparently, local measurements, as stated in the introduction, can behave quite differently than meteorological measurements.

In this study DBC and LAF were executed consecutively when a new measure-

Table 3.1: Mean and standard deviation (σ) of observation differences of meteorological stations in ‘Deelen’(D) and ‘Wageningen’(W) and local observations in ‘Wageningen’(L).

	temperature		wind speed		global radiation	
	mean	σ	mean	σ	mean	σ
L-W	0.23	0.91	-0.88	0.84	-1.90	27.5
L-D	0.44	0.92	-1.77	1.17	2.1	29.5
W-D	0.22	1.17	-0.90	1.04	4.5	23.1

ment became available. It could be worthwhile to examine the possibility of integrating both procedures into a single system. For instance, the DBC supposes a full correlation of the diurnal pattern, *i.e.* a certain error obtained now will also be present tomorrow. A more valid assumption is that this relation is exponentially decaying. As a result the σ_w/V ratio can be chosen larger. As an alternative to the stochastic filtering approach one may also consider an unknown-but-bounded error approach [see *e.g.* Sch73, Kee97].

3.5 Conclusions

It has been demonstrated in this chapter that both bias and standard deviations of forecast errors are reduced for three different weather variables: temperature, wind speed and global radiation. The tuning parameters of DBC and LAF, however, must be chosen carefully. Using historical data for tuning the parameters gives adequate results.

Chapter 4

Uncertainty Analysis of a Storage Facility under Optimal Control

This chapter is submitted for publication as:
Doeswijk, T.G., K.J. Keesman and G. van Straten. Uncertainty analysis of a storage facility under optimal control

4.1 Introduction

Predictive models that are subject to disturbances generate uncertain outputs. In addition, if these disturbances are related to the inputs, like ventilation with outside air, the uncertainty of the system states increases with increasing control inputs. If optimal control strategies are used for these systems the optimal controls are usually only valid around the nominal trajectory of the disturbances. In this paper, it is shown for a model of a storage facility of agricultural produce that it is possible to solve an optimal control problem in which the analytically calculated uncertainty is incorporated into the objective function.

The relative humidity of air can act as driving force of evaporation from water containing products. Water loss in stored agricultural products results in economic and quality losses [CG01] and should therefore be minimized. In drying processes the opposite goal is aimed for and water needs to be removed from the product until a specific water activity is reached [*e.g.* GM03]. In both processes, however, psychrometrics [Ber79] play an important role.

Moisture content in air and relative humidity are linearly related at a constant temperature. However, with increasing temperature the maximum moisture content increases nonlinearly, *i.e.* the relative humidity is nonlinearly related with temperature. The actual water content is usually calculated via partial and saturated vapor pressures. There are numerous nonlinear expressions describing the saturated vapor pressure, see *e.g.* [Son90, PGM97, TCR02]. the overall model is simplified if these relationships can be linearized over a temperature range of interest.

Model linearization of dynamical systems is usually done with a first-order Taylor-series expansion in specific points on the input/state trajectories. The validity of this linearization is limited around this point. As refrigerating and cold-storage of products generally have a limited temperature range, the error made by linearization can be calculated for the whole temperature range. The error caused by linearization can then be compared with the errors resulting from the uncertain inputs or disturbances of the system. In cold-storage or refrigerators these disturbances are for instance the time-varying outside temperature.

As uncertainty plays an important role in almost every model prediction, the propagation of errors through the model becomes important. For linear or linearized systems, it holds that if the input is normally distributed then the output is also normally distributed [Sch73]. Let us illustrate this with a simple example

$$y = ax + b, \quad x = \mathcal{N}(\mu, \sigma) \tag{4.1}$$

where, $a, b \in \mathbb{R}$. Then using the rules of expected value and variance

$$E(y) = E(ax + b) = aE(x) + b \quad (4.2)$$

$$\text{var}(y) = \text{var}(ax + b) = a^2\text{var}(x) + b \quad (4.3)$$

with $E(x) = \mu$ and $\text{var}(x) = \sigma^2$. Hence, $y = \mathcal{N}(a\mu + b, a\sigma)$, which is exact if the system is linear. In case of a linearized system the expressions are approximations. Once the uncertainty is known, the risks of predictive control actions can be analyzed. For example, the confidence intervals can be calculated [KG91]. For nonlinear systems, no general rules for error propagation are available. Numerical approaches, such as Monte Carlo analysis, are then needed to quantify the uncertainty of the output. For bilinear systems the error propagation can only be calculated analytically if the control inputs are known in advance. The possible future control inputs can be obtained by, for instance, solving an optimal control [Ste94] problem.

In this paper, a model of stored agricultural products is used as an example. Psychrometrics are part of this model and the main cause of nonlinearity of the system. The objective of this paper is to analyze the uncertainty of a storage model driven by weather forecasts by means of error propagation rules. In order to be able to use the generic rules of error propagation the model is simplified by linearization, discretization and model reduction. Given the control inputs, the predicted uncertainty is calculated and the risk analyzed. Finally, in stead of an additive analysis of the uncertainty, it is shown that the uncertainty can be directly incorporated in optimal control strategies.

This chapter is structured as follows. In section 4.2 the storage model is given and subsequently model reduction, linearization and discretization is performed. Furthermore, a noise model for the weather forecast is derived. General rules for error propagation are given in section 4.3. Optimal control problem formulations including uncertainty of the storage model are presented in section 4.4. Finally, in section 4.5, some concluding remarks are given.

4.2 Modeling

4.2.1 Storage model

The model used by [KPL03] (see appendix B), in terms of states (x), inputs u , disturbance d and parameters p , is given as

$$\dot{x} = \begin{pmatrix} p_6x_2 - p_6x_1 + p_7 - p_8e^{p_4x_1} + p_9x_3(p_5 + x_3)^{-1} \\ p_{10}x_1 + (-p_{10} - p_{11} - p_{12} - p_{13}u)x_2 + (p_{11} + p_{12} + p_{13}u)d_1 \\ (-p_{12} - p_{13}u)x_3 + (p_{12} + p_{13}u)d_2 - p_{14} + p_{15}e^{p_4x_1} - p_{16}x_3(p_5 + x_3)^{-1} \end{pmatrix} \quad (4.4)$$

where $x = [T_p, T_a, X_a]^T$ *i.e.* temperature of produce, temperature of air and absolute humidity of air and $d = [T_e, X_e]^T$, *i.e.* outside temperature and absolute humidity. Furthermore, u is the control input $\alpha\phi$ with α fraction of valve opening and ϕ the fraction of the maximum possible ventilation rate. Finally, p is a vector of functions of design and physical parameters.

The driving force for exchange is the difference between actual vapor pressure in the surrounding air and the saturated vapor pressure at the surface of the product. The saturated vapor pressure is given by

$$p_s = 100 \left(-1.7001 + 7.7835e^{\frac{T}{17.0798}} \right) \quad (4.5)$$

which explains the presence of the exponential term in (4.4). The partial pressure can be calculated from the nonlinear relationship [Ber79]

$$p_a = \frac{X_a P_{tot}}{0.6228 + X_a} \quad (4.6)$$

with P_{tot} the total pressure. This explains the fractional in (4.4).

4.2.2 Model reduction

The model (4.4) was reduced by [KPL03] on the basis of the singular perturbation theory. For this analysis we therefore refer to this paper.

In cold storage, *i.e.* $T_e < 10^\circ C$, the maximum water content $X_s(T = 10^\circ C) = 0.0077$. Because $X_a < X_s(T = 10^\circ C) \ll 0.628$ eqn. (4.6) is reduced to

$$p_a = \frac{P_{tot}}{0.6228} X_a \quad (4.7)$$

The error that is introduced is smaller than the assumption that P_{tot} is constant. By applying (4.7) for the partial vapor pressure the physical model is reduced to:

$$\begin{aligned} \dot{T}_p = & \left(\frac{p_6 p_{10}}{(p_{10} + p_{11} + p_{12} + p_{13} \alpha \phi)} - p_6 \right) T_p + \left(\frac{p_9 p_{15}}{p_5 (p_{12} + p_{13} \alpha \phi) + p_{16}} - p_8 \right) e^{p_4 T_p} \\ & + \frac{p_6 (p_{11} + p_{12} + p_{13} \alpha \phi)}{(p_{10} + p_{11} + p_{12} + p_{13} \alpha \phi)} T_e + \frac{p_9 (p_{12} + p_{13} \alpha \phi)}{p_5 (p_{12} + p_{13} \alpha \phi) + p_{16}} X_e \\ & + \left(p_7 - \frac{p_9 p_{14}}{p_5 (p_{12} + p_{13} \alpha \phi) + p_{16}} \right) \end{aligned} \quad (4.8)$$

with α , ϕ the control inputs equivalent to u of (4.4) and quasi steady-states T_a , X_a given by

$$\bar{T}_a = \frac{p_{10} T_p + (p_{11} + p_{12} + p_{13} \alpha \phi) T_e}{(p_{10} + p_{11} + p_{12} + p_{13} \alpha \phi)} \quad (4.9)$$

$$\bar{X}_a = \frac{(p_{12} + p_{13} \alpha \phi) X_e - p_{14} + p_{15} e^{p_4 T_p}}{p_{12} + p_{13} \alpha \phi + \frac{p_{16}}{p_5}} \quad (4.10)$$

4.2.3 Interval linearization

For a specific interval of T_p the non-linear term of (4.8), *i.e.* $e^{(p_4 T_p)}$, can be approximated by the linear expression, $p_{20} T_p + p_{21}$. This can be done by minimizing the quadratic error over the temperature range.

$$\min_{p_{20}, p_{21}} J = \min_{p_{20}, p_{21}} \int_{T_p, min}^{T_p, max} (e^{(p_4 T_p)} - (p_{20} T_p + p_{21}))^2 dT_p \quad (4.11)$$

Now let us assume that for cold storage the temperature range of the produce varies between $3^\circ C$ and $9^\circ C$. The parameters p_{20} and p_{21} are chosen such that the integral is minimized. In this particular case where $p_4 = 17.0798^{-1}$ this results in $p_{20} = 0.0834$ and $p_{21} = 0.9275$.

Given the approximation $e^{p_4 T_p} \approx p_{20} T_p + p_{21}$ the exponential term in (4.8) is replaced by the linear one. This results in

$$\begin{aligned} \dot{T}_p = & \left(\frac{p_6 p_{10}}{(p_{10} + p_{11} + p_{12} + p_{13} \alpha \phi)} + \frac{p_{20} p_9 p_{15}}{p_5 (p_{12} + p_{13} \alpha \phi) + p_{16}} - p_6 - p_{20} p_8 \right) T_p \\ & + \frac{p_6 (p_{11} + p_{12} + p_{13} \alpha \phi)}{(p_{10} + p_{11} + p_{12} + p_{13} \alpha \phi)} T_e + \frac{p_9 (p_{12} + p_{13} \alpha \phi)}{p_5 (p_{12} + p_{13} \alpha \phi) + p_{16}} X_e \\ & + \left(p_7 + \frac{p_9 (p_{21} p_{15} - p_{14})}{p_5 (p_{12} + p_{13} \alpha \phi) + p_{16}} - p_{21} p_8 \right) \end{aligned} \quad (4.12)$$

Now, if the control inputs α and ϕ are given and vary in time, a linear (time-varying) model is obtained. The error of (4.12) compared to the original model is very small ($< 0.01^\circ C$) for a constant control $\alpha \phi = 1$ (see figure 4.1). For smaller values of $\alpha \phi$ the error will be larger because the steady state temperature increases and may exceed the $9^\circ C$. Because this is outside the linearization interval where the error is larger, the error will increase. However, even in case $\alpha \phi = 0$ the error is less than $0.25^\circ C$ with $T_p = 24^\circ C$. Note that the linearization of (4.11) can be done for every kind of nonlinear realization of p_s given in literature. Clearly, the error size depends on the size of the linearization interval.

4.2.4 Discretization

The weather forecasts used for prediction of the indoor climate are given on an hourly basis. Furthermore, in practice, the control variables are manipulated on a specific time-interval. Hence, the variables u , T_e and X_e of (4.12) are discrete-time variables. There are two ways to line-up the model. First, the discrete-time variables can be transformed into continuous-time variables by for instance interpolation or zero-order hold methods. Second, the continuous-time model can be discretized with an interval equal to the time interval related to the discrete-time variables. In this case, the latter option is chosen.

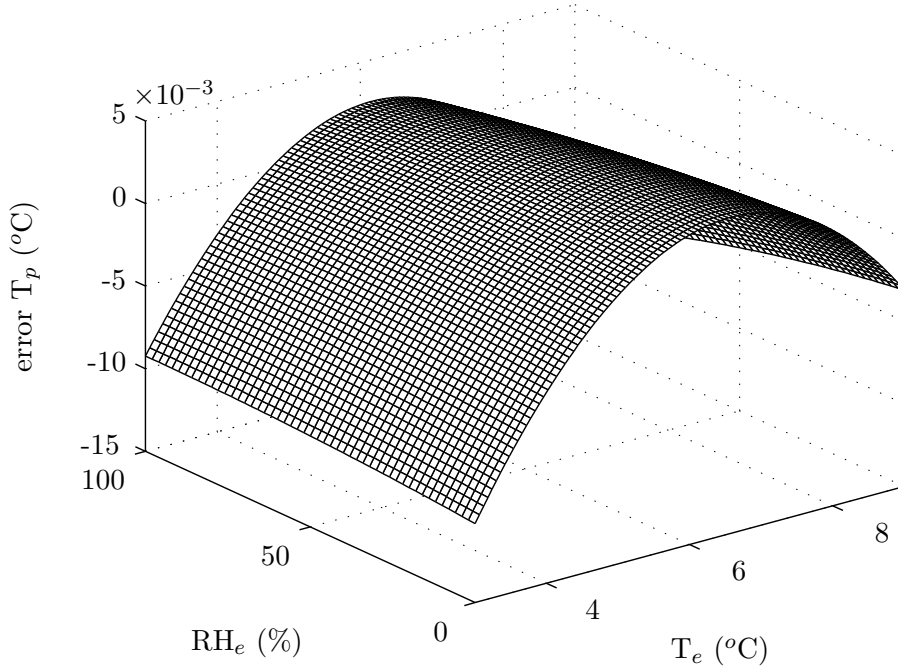


Figure 4.1: steady state error of T_p for constant inputs T_e and X_e introduced by linearization

A continuous-time model can be approximated by a discrete-time model with *e.g.* Euler discretization. Given a linear (time-varying) model

$$\dot{x} = F(t)x(t) + G(t)w(t) \quad (4.13)$$

The discretization procedure is given by [Sch73]

$$x(k\Delta + \Delta) = A(k\Delta)x(k\Delta) + \Delta G(k\Delta)w(k\Delta) \quad (4.14)$$

with,

$$A(k\Delta) = I + \Delta F(k\Delta) \quad (4.15)$$

If Δ is sufficiently small then equation (4.14) approximates (4.13). For the storage model (4.12) discretization leads to

$$\begin{aligned} T_p(k+1) = & \left(1 + p_{22} \left(\frac{p_6 p_{10}}{(p_{10} + p_{11} + p_{12} + p_{13} \alpha(k) \phi(k))} + \frac{p_{20} p_9 p_{15}}{p_5 (p_{12} + p_{13} \alpha(k) \phi(k)) + p_{16}} - p_6 - p_{20} p_8 \right) \right) T_p(k) \\ & + p_{22} \frac{p_6 (p_{11} + p_{12} + p_{13} \alpha(k) \phi(k))}{(p_{10} + p_{11} + p_{12} + p_{13} \alpha(k) \phi(k))} T_e(k) + p_{22} \frac{p_9 (p_{12} + p_{13} \alpha(k) \phi(k))}{p_5 (p_{12} + p_{13} \alpha(k) \phi(k)) + p_{16}} X_e(k) \\ & + p_{22} \left(p_7 + \frac{p_9 (p_{21} p_{15} - p_{14})}{p_5 (p_{12} + p_{13} \alpha(k) \phi(k)) + p_{16}} - p_{21} p_8 \right) \end{aligned} \quad (4.16)$$

with, $p_{22} = \Delta$

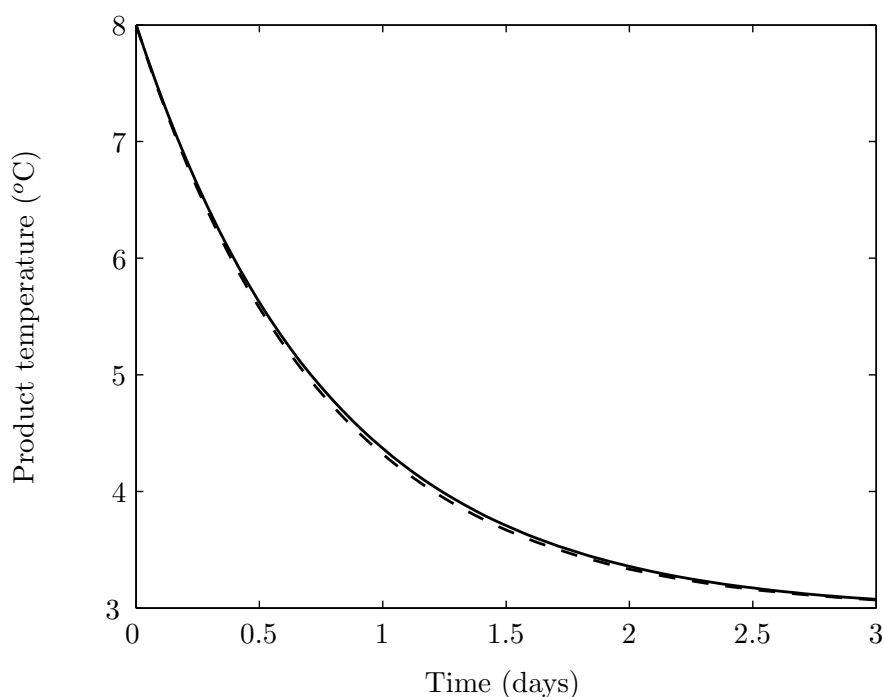


Figure 4.2: Responses of the original model and the approximated model with $\alpha\phi = 1$, $T_e = 3^\circ C$ and $X_e = 3 \cdot 10^{-3} kg/kg$

4.2.5 Approximation validity

In this section the procedure of model reduction, linearization and discretization is validated. The approximated model (4.16) is compared with the original model (4.4) by generating a model response with a constant input. The two extreme cases are considered, *i.e.* $u = 0$ and $u = 1$. The two external inputs are considered to be constant, *i.e.* $T_e = 3^\circ C$ and $X_e = 3 \cdot 10^{-3} kg/kg$. The responses are presented in figures 4.2 and 4.3. Note that the time scales of both figures differ. The responses of both models are very much alike. The maximal differences of the predicted temperatures are 0.052 for $u = 1$ and -0.018 for $u = 0$. These differences are much smaller than the uncertainties in the measurements in practice. Consequently, the simplified model (4.16) can then be used for analysis.

Note 4.1. The model is not validated with experimental data. However, similar models [LdKCvdV06] were fitted on experimental data and produced good results.

4.2.6 Noise modeling

In what follows, the uncertain environmental variables (T_e and X_e) are divided into a deterministic and a zero mean noise part. Hence, the model (4.16) becomes of the form:

$$x(k+1) = a(k)x(k) + b_1(k)d_1(k) + b_2(k)d_2(k) + b_1(k)c_1(k) + b_2(k)c_2(k) + \gamma(k) \quad (4.17)$$

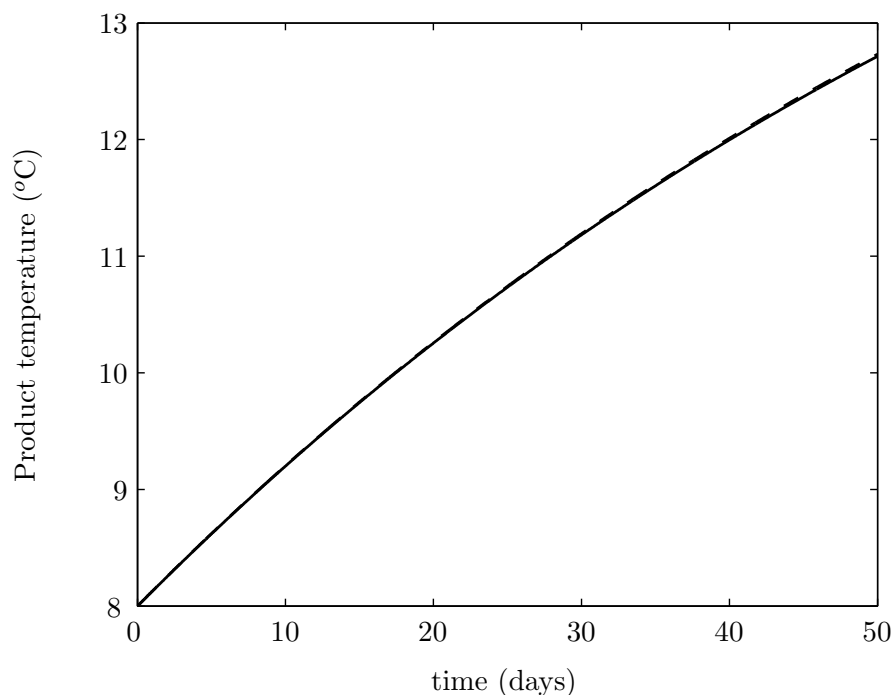


Figure 4.3: Responses of the original model and the approximated model with $\alpha\phi = 0$, $T_e = 3^\circ C$ and $X_e = 3 \cdot 10^{-3} kg/kg$

where $x = T_p$, $d_1 = T_e$, $d_2 = X_e$, c_1, c_2 are discrete noise terms related to T_e and X_e and γ the constant term in (4.16). The noise sequences can be calculated from the forecast errors that are known from historical data. The mean and variances of the noise sequences $\{c_1\}$ and $\{c_2\}$ are calculated for hourly intervals. However, the input uncertainty related to T_e and X_e cannot be seen as white noise sequences because these variables are outputs of meteorological models. In chapter 2 the covariance matrix for a 0 to 31 hours ahead forecast of the temperature is given. From this matrix the correlation between the zero and k hour ahead forecast along the whole forecast range, *i.e.* $k = 0, \dots, 31$ has been calculated. In figure 4.4 the correlation of the forecast error of the temperature can be clearly seen. Therefore the noise sequences $\{c_1\}_{k=1}^N$ and $\{c_2\}_{k=1}^N$ uncertainties should be modeled as colored noise which can for instance be achieved by the first-order fitter

$$\begin{aligned} c_1(k+1) &= a_2 c_1(k) + w_1(k) \\ c_2(k+1) &= a_3 c_2(k) + w_2(k) \end{aligned} \quad (4.18)$$

If $\{w\}$ is a zero mean white noise sequence then $\{c\}$ is also zero mean. The constants, a_2 and a_3 are calculated using the autocorrelation function that satisfies the first-order difference equation $\rho(k) = \phi_1 \rho(k-1)$ [BJ76, p.57]. This result in: $a_2 = 0.839$ and $a_3 = 0.856$. The auto-correlation function with its approximation is shown in figure 4.4.

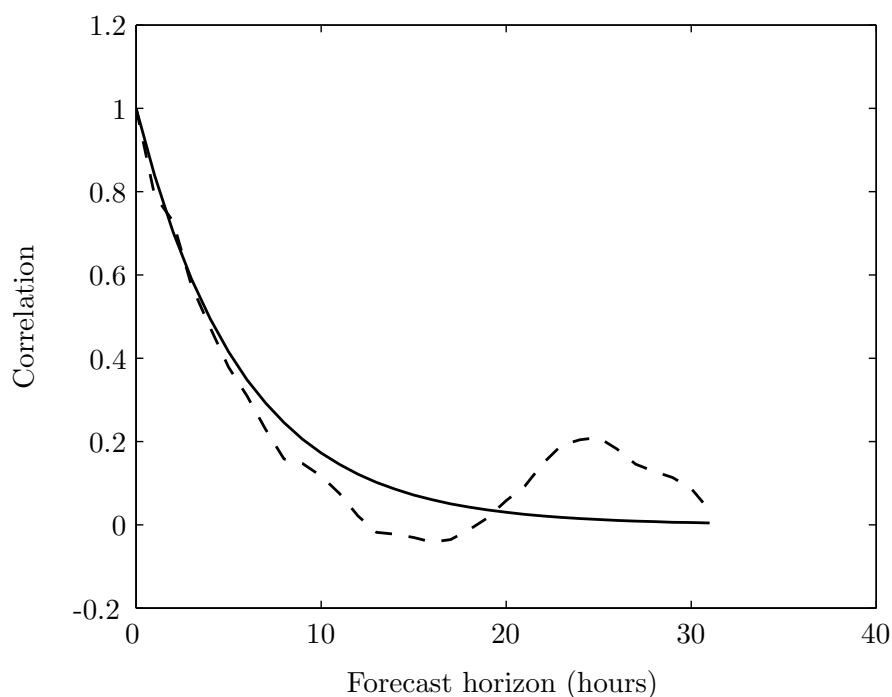


Figure 4.4: Auto correlation function of forecast error of T_e and its first-order approximation.

Notice that the expectation of T_p is not changed after substituting the discrete-time the state equations of the noise, (4.18) in (4.17), as the noise is assumed to be zero-mean white noise. However, the error propagation through the model is affected. This is further analyzed in section 4.3.

4.3 Error propagation

Given a general linear time-invariant system:

$$x(k+1) = Ax(k) + Gw(k) \quad (4.19)$$

with $w(k)$ zero mean white noise, the second statistical moment is given by [Sch73]:

$$P(k+1) = AP(k)A^T + GQG^T \quad (4.20)$$

where P the covariance matrix of the states and Q the covariance matrix of the disturbances, *i.e.* $E(w^T w)$. This is analogous to the static system example given in eqns. (4.1)-(4.3).

The model of the second statistical moment cannot be applied without some modifications to the storage model (4.17)-(4.18). First, the variance of the white noise sequences ($\{w\}_{k=1}^N$) of (4.18) is not known. If it is assumed that c_1 and c_2 are independent, however, the white noise sequence follows from:

$w(k) = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} (k) = c(k+1) - \Gamma c(k)$ where $\Gamma = \begin{bmatrix} a_2 & 0 \\ 0 & a_3 \end{bmatrix}$ and $c(k) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} (k)$. The variance is then defined by $\text{var}(w(k)) = \text{var}(c(k+1) - \Gamma c(k))$. Now, let us assume that $\text{var}(c(k+1)) \approx \text{var}(c(k))$. Generalizing the statistical rules: $\text{var}(x - \alpha y) = \text{var}(x) + \alpha^2 \text{var}(y) - 2\alpha \text{cov}(x, y)$ and $\text{cov}(x, y) = \rho_{xy} \sigma_x \sigma_y$, with ρ_{xy} the correlation coefficient, the variance of the noise vector $w(k)$ is then given by

$$\text{var}(w(k)) = (I - \Gamma^T \Gamma) \text{var}(c(k)) \quad (4.21)$$

The variance of $c(k)$ can be determined from weather data [e.g. DK05]. Given $c(k)$, the mean and variance of $w(k)$ is calculated.

Second, the storage model is time-varying. For time-varying matrices A , G and Q the model of the second statistical moment (4.20) can be adapted to [Åst70]:

$$P(k+1) = A(k)P(k)A(k)^T + G(k)Q(k)G(k)^T \quad (4.22)$$

Hence, the uncertainty throughout the storage model (4.17)-(4.18) can be calculated provided that the controls, *i.e.* $A(k)$ and $G(k)$, and the weather forecast uncertainty $Q(k)$ are known.

Note 4.2. In this case only the error propagation caused by uncertain inputs is calculated. The model itself is assumed to be perfect.

Given the uncertainty, the 95% confidence limits of the states are known. In the next section an example will be presented in which the 95% confidence limits of the storage temperature are given.

4.4 Optimal control

In optimal control theory a cost function is minimized by adjusting the controls [Ste94]. In case of the storage model, the optimal control trajectory can be calculated if the forthcoming weather is exactly known in advance. The best approximation, however, is to use weather forecasts. Unfortunately, the weather forecast is generally rather uncertain. In this case a cost function around the nominal trajectory (d_0) can be used to calculate the optimal controls. In general terms, find u from

$$\min_u J = \Phi(x(H)) + \sum_{k=0}^H f(x(k), u(k), d_0(k), w(k)) \quad (4.23)$$

Given the linearized model, the variance of the of the disturbances, the control inputs and the error propagation rules, the 95% confidence interval of the state variable can be calculated. By varying the control inputs the matrices $A(k)$ and $G(k)$ of (4.22) the covariance matrix $P(k+1)$ also changes. As uncertainty is undesirable in many cases, reduction of the state variance is needed. Minimum variance control

[Åst70] is frequently used in these cases. In practical optimal control problems, however, in addition to the variance of the predictor the cost function generally will contain other terms. For instance, in the storage case minimum variance control would lead to no ventilation because the uncertainty is directly related to the weather forecast. More uncertainty is brought into the system during ventilation, *i.e.* $G(k)$ increases if $\alpha(k)\phi(k)$ increases. Furthermore, other elements like ventilation costs and maintaining a temperature near a reference still are important in the cost function. Therefore, an approach where the uncertainty is incorporated into the cost function seems more natural. Let us illustrate this with a simple example. A quadratic cost function could be formulated as follows:

$$J = \mathbb{E} \sum_{k=0}^H ((x(k) - x_{ref})^2 + u(k)^2) \quad (4.24)$$

As the uncertainty of the state $x(k)$ is subject to the uncertainty $P(k)$ the cost function can be expanded by $P(k)$, so that:

$$J = \mathbb{E} \sum_{k=0}^H ((x(k) - x_{ref})^2 + u(k)^2 + P(k)) \quad (4.25)$$

Note that $P(k)$ is not squared because it is already a squared variable (σ^2). A more natural way to handle the uncertainty is not to use x but *e.g.* $x \pm 2\sigma$, the 95% confidence limit of x . The uncertainty is now directly taken into account and no further weighting of the uncertainty is necessary. By choosing $|x - x_{ref}| + 2\sqrt{P}$ in stead of the expected value of $x - x_{ref}$, with $|x - x_{ref}| = \sqrt{(x - x_{ref})^2}$ the quadratic cost function changes to:

$$J = \sum_{k=0}^H \left((x(k) - x_{ref})^2 + 4P(k) + 4\sqrt{P(k)(x(k) - x_{ref})^2 + u(k)^2} \right) \quad (4.26)$$

In the next example the cost functions (4.24)-(4.26) will be evaluated for the storage model (4.17)-(4.18).

Example 4.1. The weather forecasts used in this example consist of an ensemble of forecasts [PBBP97]. This ensemble of forecasts is then used to calculate the mean and the variance at each forecast horizon. The ensemble forecast of the temperature of 1 april 2005 is given in figure 4.5. The nominal weather forecast (*i.e.* mean value of the ensemble) is used as input whereas the variance from the ensemble forecast is used to calculate the variance of the white noise sequence (*i.e.* $Q(k)$) at every time instance k .

Now that $T_e(k)$, $X_e(k)$ and $Q(k)$ are known, a parametric optimization procedure [Ste94] is used to calculate the optimal controls for each of the cost functions (4.24)-(4.26), with $x_{ref} = 7^\circ C$. In figures. 4.6-4.8 the predicted product temperature profiles are given with their 95% confidence limits and the controls. From these figures it

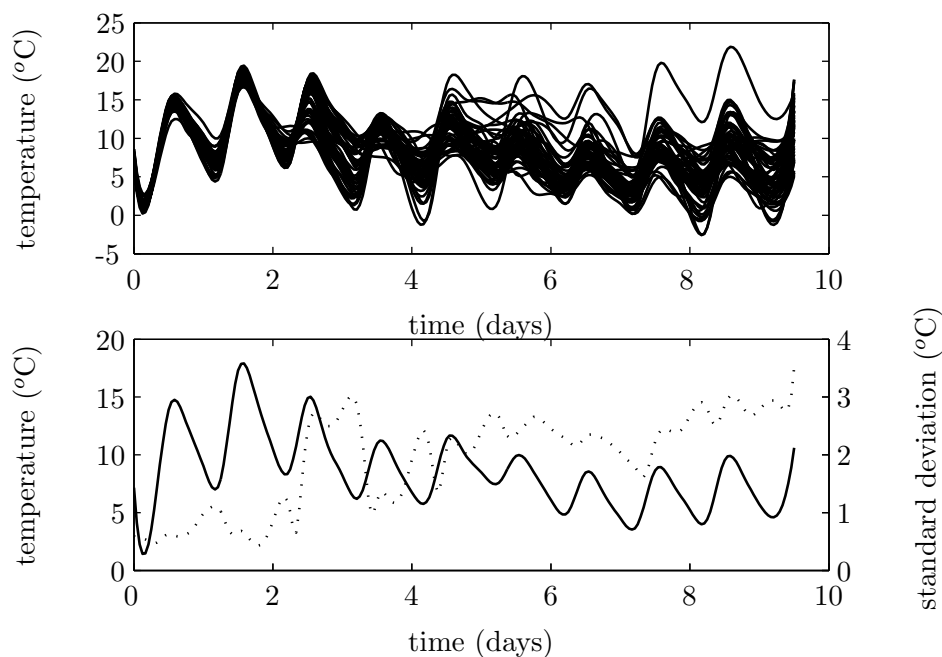


Figure 4.5: Ensemble prediction of the temperature of 1 april 2005 (top) and the mean value (—) and the standard deviation (\cdots) of the ensemble (bottom)

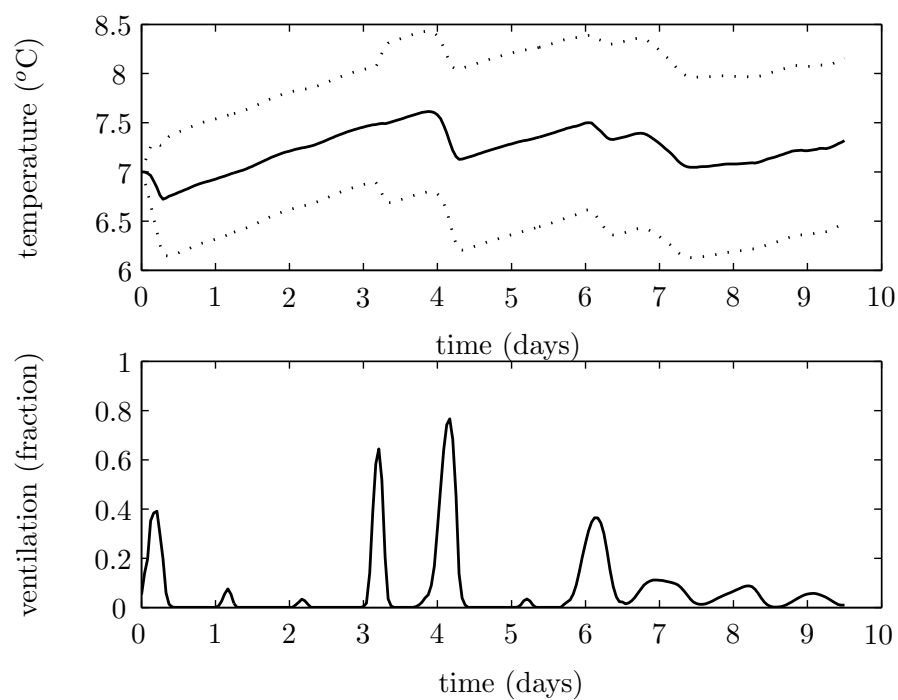


Figure 4.6: Optimal state trajectory (—) with the 95% confidence limit (\cdots) (top) with the accompanying controls (bottom) using cost function (4.24).

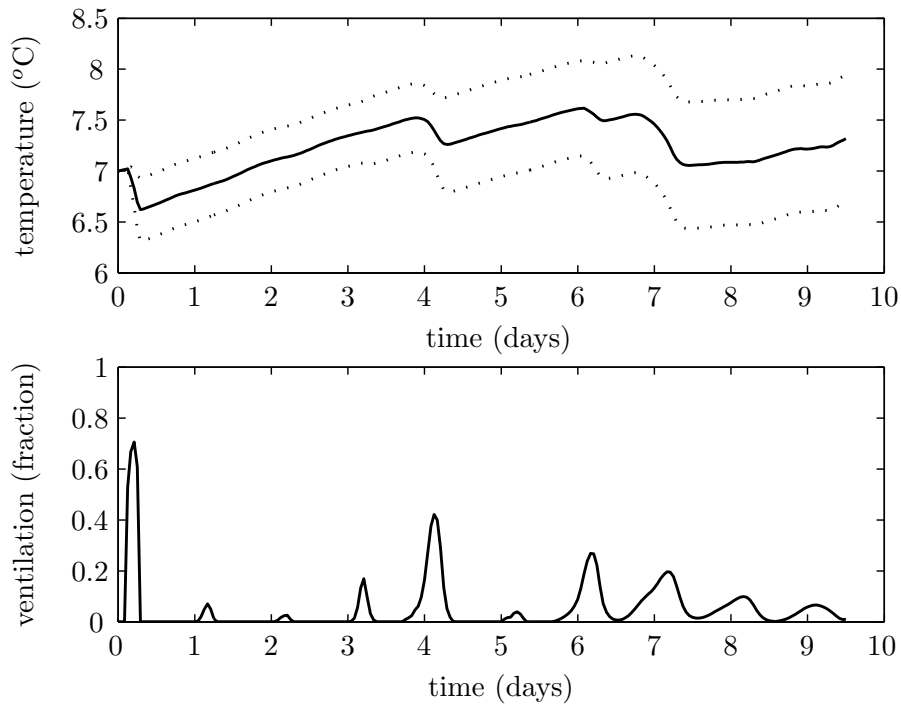


Figure 4.7: Optimal state trajectory (—) with the 95% confidence limit (···) (top) with the accompanying controls (bottom) using cost function (4.25).

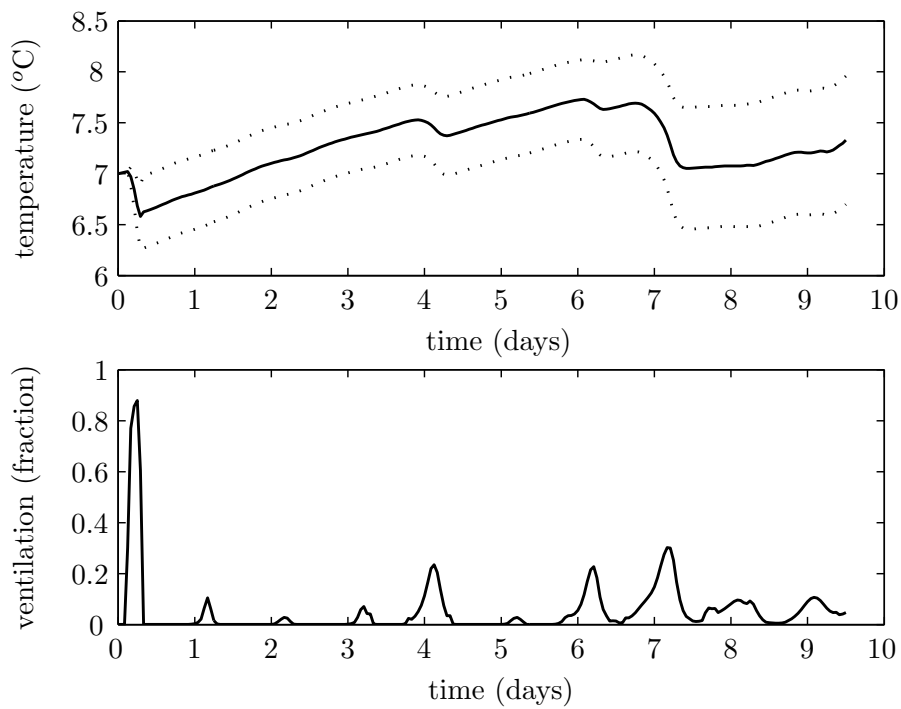


Figure 4.8: Optimal state trajectory (—) with the 95% confidence limit (···) (top) with the accompanying controls (bottom) using cost function (4.26).

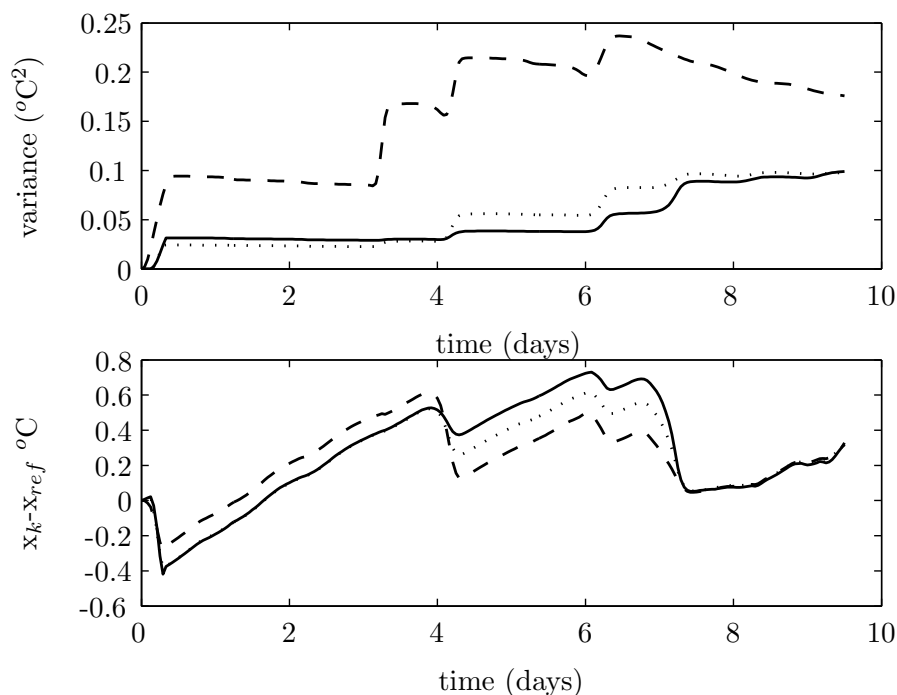


Figure 4.9: Variance (top) and residuals ($x_k - x_{ref}$, bottom) of the calculated state trajectories using (4.24) (---), (4.25) (···) and (4.26) (—).

can be seen that by incorporating uncertainty in the cost function the control inputs shift to points of lower uncertainty in the weather forecast. This finally results in a lower state covariance (see figure 4.9). Furthermore, it can be seen that the state covariance increases less over time if the confidence limit is taken into account during optimization. It can also be seen that the residuals, *i.e.* $x - x_{ref}$, in general become larger. This is more clear by calculating the sum of squared errors (see table 4.4). This is the trade-off by minimizing the uncertainty.

Note 4.3. In the example no real risk avoidance is incorporated. This could be done as well by using cost function (4.24) and a hard state inequality constraint *e.g.* $|x - x_{ref}| + 2\sqrt{(P)} \leq \delta$. It is then possible, however, that there is no feasible solution.

Table 4.1: sum of squared errors of the expected state trajectory using different cost functions

applied cost function	(4.24)	(4.25)	(4.26)
sse	21.4	26.4	35.7

4.5 Concluding remarks

When the operating point of a process is kept within bounds the model can be linearized on this interval by using an optimality criterium instead of using standard

linearization techniques around a point. For a storage model, this linearization procedure proved to be successful. If the controls are known or have been calculated in advance, the storage model becomes linear and the error propagation calculation can be done analytically. Finally, the model uncertainty is used in an optimal control framework such that the predicted 95% confidence limits are kept as close as possible to the reference trajectory. If implemented as state constraints the confidence limits can also be used for risk avoidance.

Chapter 5

Impact of Weather Forecast Uncertainty in Optimal Climate Control of Storehouses

This chapter has been published as:
Doeswijk, T.G., K.J. Keesman and G. van Straten. Impact of weather forecast uncertainty in optimal climate control of storehouses. In *Control Applications in Post - Harvest and Processing Technology*, Bornimer Agrartechnische Berichte, 55, pages 46-55, Potsdam, Germany, 2006

5.1 Introduction

Indoor climate in greenhouses, office buildings and storage facilities for agricultural produce are generally related to outdoor weather conditions [CLW⁺01, LdKCvdV06, Tap00]. The indoor climate is affected by for instance heat transfer through walls, radiation, respiration and evapotranspiration of organisms, ventilation with outdoor air and by active heating or cooling. If anticipating control strategies such as receding horizon optimal control are to be used for controlling the indoor climate using outdoor climate, weather forecasts become important. Uncertainties in weather forecasts then lead to uncertain predicted indoor climates. Moreover, the cost function relating to any kind of optimal control strategy is subject to errors in the weather forecasts. The sensitivity of the model outputs and the related cost function therefore needs to be investigated. For both air conditioned buildings [HKFK04] and a potato storage facility [KPL03, LvdVdKC06] it was shown that a good weather forecast reduces the cost function almost as much as a perfect weather forecast. In these studies short term weather forecasts (1 to 2 days ahead) were used. In this study which focuses on effect of the uncertainty in weather forecasts, medium range weather forecasts (up to 10 days ahead) are used.

The medium range weather forecast [PBBP97] consists of an ensemble of 50 different weather forecasts. All 50 ensemble members have an equal probability of occurring. Hence, the uncertainty of the weather forecasts is known a priori. This knowledge can be used to evaluate a calculated optimal control solution by calculating the costs of each of the ensemble members. Next to this, the optimal control solution can be evaluated a posteriori with the observed weather.

In general, a model is an approximation and not an exact representation of the system investigated. Hence, not only uncertainties in the inputs (*i.e.* weather forecasts) but also uncertainties of the model effect the outcome of an optimal control algorithm.

The objective of this paper is to evaluate the effect of uncertainty in the weather forecast on the costs of the calculated optimal control problem of a potato storage facility. Furthermore, the effect of model uncertainty on the costs is investigated. Finally, a time criterion for which the optimal control problem is to be recalculated (*i.e.* feedback is applied) based on model and input uncertainties is proposed.

This chapter is structured as follows: first a brief description of the storage model and the cost function are given. Next, the type of weather forecasts used is explained. Then, the open loop and closed loop evaluations using weather forecasts and observations are given in subsequent sections. Finally, results are discussed and conclusions are drawn.

5.2 Preliminaries

5.2.1 Storage model

The model used to simulate a potato storage facility is based on the work of [LdKCvdV06]. The original model was simplified to make it suitable for use in a receding horizon optimal control algorithm [LvdVdKC06]. A brief description of this discrete-time model is

$$x_T(k+1) = x_T(k) + p_1 + (d_{T_{wb,ext}}(k) - x_T(k)) (p_3 + p_2 p_5 u_{mix}(k) (p_7 + (1 - p_7) u_{vent}(k))) \quad (5.1)$$

$$x_{CO_2}(k+1) = x_{CO_2}(k) + p_1 p_4 + (d_{CO_2,ext}(k) - x_{CO_2}(k)) (p_2 u_{mix}(k) (p_7 + (1 - p_7) u_{vent}(k)) + p_6) \quad (5.2)$$

where x_T represents the product temperature, x_{CO_2} the CO_2 concentration in the bulk, $d_{T_{wb,ext}}$ the ambient wet bulb temperature, $d_{CO_2,ext}$ the ambient CO_2 concentration, u_{vent} the fraction of maximum possible internal ventilation, u_{mix} the fraction of ambient air in the air flow, and p a vector of parameters containing physical and design parameters. Hence, the control input is defined by: $u = [u_{mix}, u_{vent}]^T$. The sample time used for this model is one hour.

The temperature of the bulk can be measured. The CO_2 -concentration, however, is difficult to measure in practice. The storehouse therefore is controlled on the product temperature. The CO_2 -concentration is simulated to avoid too high concentrations that lead to damage of the product (see section 5.2.3). This is possible because the CO_2 -concentration reduces very fast during ventilation, *i.e.* the CO_2 -concentration reaches equilibrium fast.

5.2.2 Weather forecasts

The weather forecasts used in this paper are the so-called medium range weather forecasts provided by Weathernews Benelux. This means that the forecast up to 10 days ahead is given. Due to the chaotic nature of the atmosphere, the weather forecast is sensitive to the initial conditions. Therefore, an ensemble prediction system [PBBP97] is introduced to improve the forecast quality and to get more insight into the uncertainty of the forecast. The weather forecast used in this paper consists of 50 ensemble members. Each of the members have an equal chance of occurring. The ensemble mean is used as the nominal weather forecast.

A new weather forecast becomes available every 24 hours. Short term weather forecasts (up to 36 hours ahead) can be improved by using local observations as was shown in chapters 2 and 3. A similar method is used here to correct the medium range weather forecast.

5.2.3 Cost function

To give a little more insight into the control objectives of the potato storehouse, the elements of the cost function are mentioned below. How the weighting factors are chosen is beyond the scope of this paper. The following control objectives are to be fulfilled:

- The temperature of the bulk must be kept as close as possible to a pre-specified reference temperature ($\min_u ||x_T(k) - T_{ref}||$).
- The temperature must always be kept above a specified minimum temperature (inequality constraint $x_T(k) > T_{min}$).
- The temperature may not decrease faster than a specified limit within 24 hours (inequality constraint). ($x_T(k - 24) - x_T(k) < T_{\Delta}$).
- The weight loss of the product due to evaporation must be minimized ($\min_u \sum_{k=0}^H f_1(x_T(k), d_{T_{wb}}(k), u(k))$) with H the control horizon.
- The CO_2 -concentration must always be kept below a specified maximum (inequality constraint $x_{CO_2} < CO_{2,max}$).
- The energy costs related to ventilation must be minimized ($\min_u \sum_{k=0}^H f_2(u_{vent}(k))$).

5.2.4 Receding horizon optimal control

Given the model (5.1)-(5.2), a weather forecast (containing $d_{T_{wb,ext}}(k)$) and assuming $d_{CO_2,ext}(k) = 0.0314$ to be constant, a control trajectory can be calculated that minimizes a cost function according to:

$$\min_{u(k)} J = \min_{u(k)} E \left[\phi[x(k+H)] + \sum_{k=0}^H L[x(k), u(k), d(k)] \right] \quad (5.3)$$

where the expected value is taken because the weather forecast is a stochastic variable [Ste94]. In the investigated case there are no final costs *i.e.* $\phi[x(k+H)] = 0$. If this open-loop control problem is solved repeatedly every l hours ($l < H$) given the updated (or measured) states and weather conditions the control loop is closed. This control strategy is called receding horizon optimal control (RHOC).

The RHOC solution with the nominal weather forecast is taken as the reference point for the uncertainty evaluation. The RHOC was started at 13 April 2005 and ended at 2 June 2005 using medium-range weather forecasts for location 'De Bilt, The Netherlands'. The forecast horizon of the weather forecast also determines the horizon of the RHOC ($H = 217$ hours). During this period every six hours a new optimal control trajectory was calculated (*i.e.* $l = 6$) with a new external weather

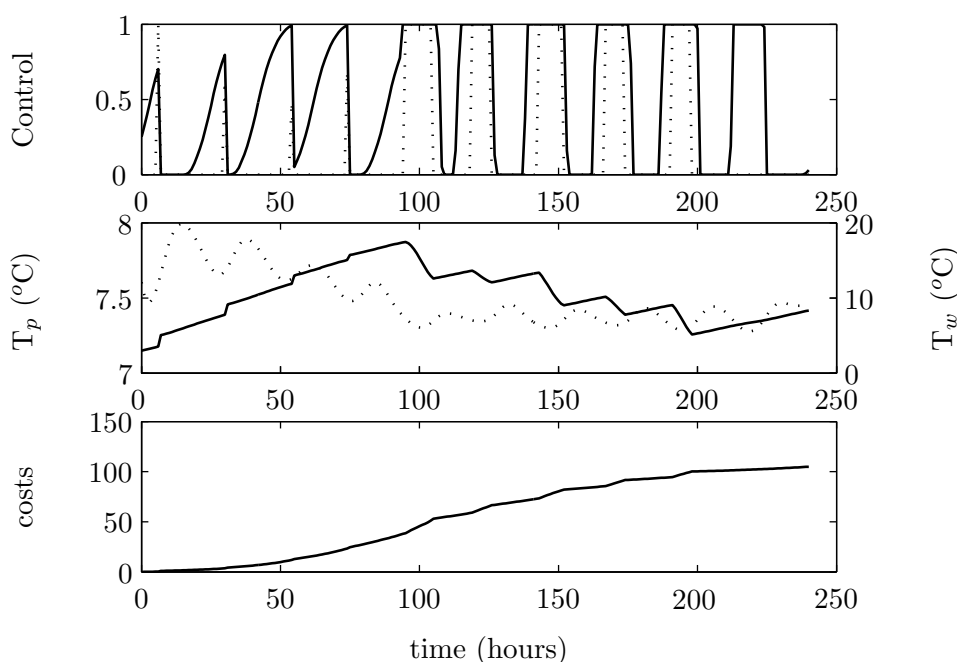


Figure 5.1: Optimal control output of 1 May 2005. The subfigures show respectively: calculated optimal controls ($- u_{\text{mix}}$, $\cdots u_{\text{vent}}$; predicted product temperature ($-$) and external wet bulb temperature (\cdots); predicted cumulative costs

forecast. The initial conditions of each optimal control run was taken equal to the measurements. In addition, the newly calculated controls were implemented. A total of 203 optimal control trajectories are calculated (4 times a day, almost 51 days). An example of a calculated control trajectory with the accompanying predicted state and cost evolution is presented in figure 5.1. The reference temperature (T_{ref}) in this case was 7°C . With the increasing temperature also the costs increase. Therefore, long periods of ventilation occurs if the external wet bulb temperature is lower than the product temperature.

5.3 Weather uncertainty

There are several possibilities to evaluate the uncertainties in the cost function. First the change in cost function is investigated when observations of the weather are used instead of the forecast. Second, the effect of ensemble members on the cost function is investigated. It should be noted that no new RHOC calculations are performed based on different weather forecasts or parameter values. The calculated costs based on the nominal weather forecast are compared with other weather forecasts or parameter sets with the same calculated control trajectories of u_{mix} and u_{vent} .

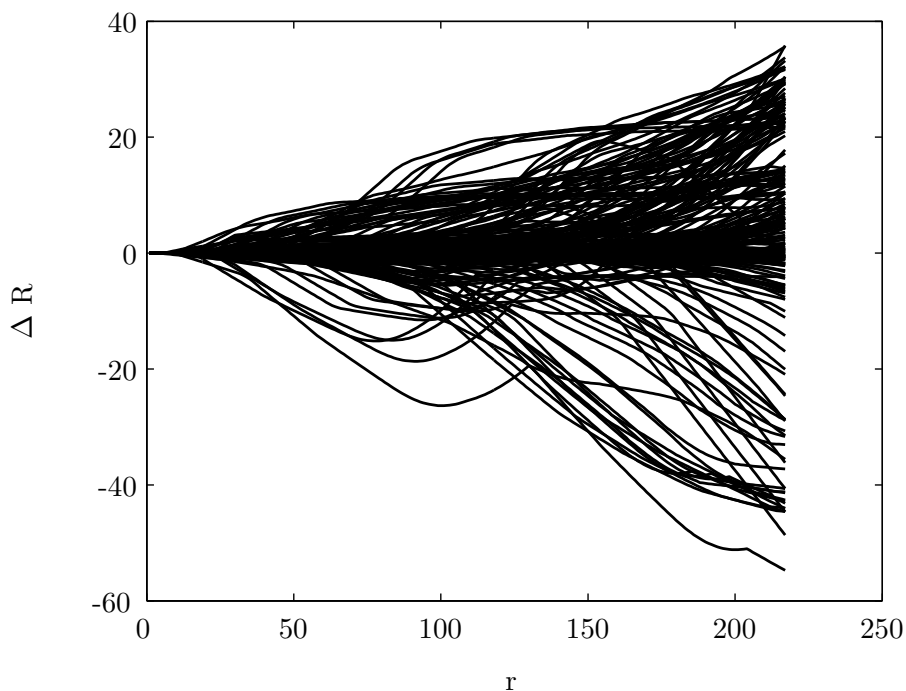


Figure 5.2: Difference in calculated absolute costs and realized costs of 203 optimal control runs

5.3.1 Open loop evaluation

If the model is run again but now with the observed weather instead of the forecasted weather, a change in costs is observed. In figure 5.2 the difference between calculated and realized running costs are given with the running costs defined by:

$$R = \sum_{k=0}^r L(x(k), u(k), d(k)) \quad (5.4)$$

If $r = H$ then $R = J$. All 203 cost function trajectories are recalculated with the observed weather and the given control trajectories. It can be seen that the total cost difference can be both positive and negative. This means that, given the control trajectory, the realized weather can reduce the total costs more than was expected from the forecasted weather. It does not mean that the calculated control trajectory is also related to the minimum if the observed weather was used in the RHOC.

Because the total absolute costs J change for every optimal control run, because of changing initial states and changing weather forecasts, the relative change in the cost function, *i.e.* $\Delta R^{rel} = \frac{R_{obs} - R_{fct}}{J_{obs}}$, is given in figure 5.3. The absolute costs can change dramatically (not shown here) over the researched period because the outside temperature increases. Especially, the cost criterion $\|x_T(k) - T_{ref}\|$ then increases. Relatively to the costs at $t(k) + H$, however, the change in costs do not seem to change that dramatically.

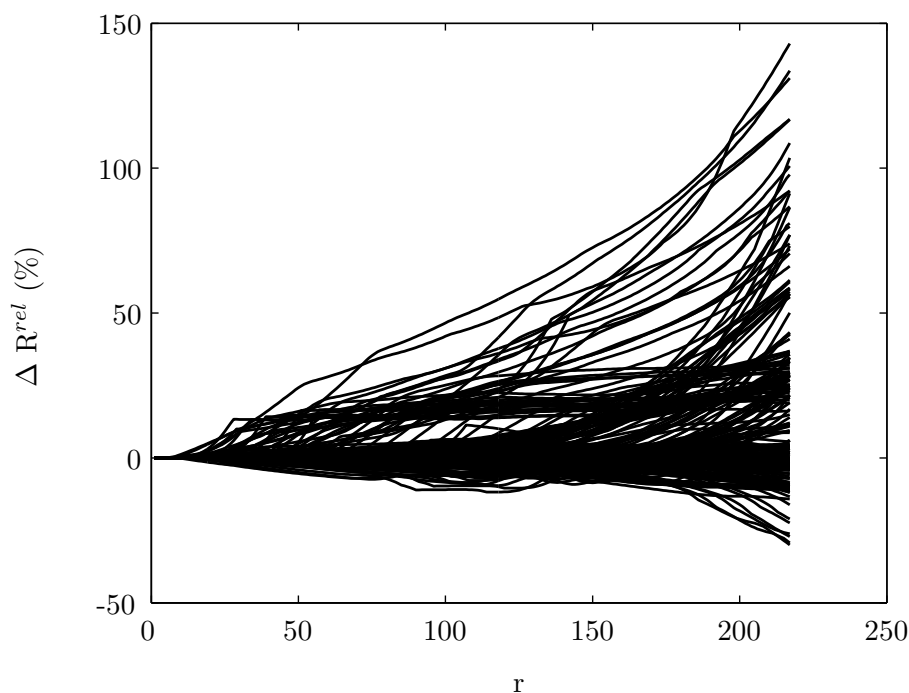


Figure 5.3: Difference in calculated relative costs and realized costs of 203 optimal control runs

Using the observed weather to calculate the costs is a useful a posteriori tool. To evaluate the uncertainty of the costs a priori, however, other information about the uncertainty of the weather forecast is needed. The uncertainty of the weather forecast is embedded within the ensemble (see section 5.2.2). By calculating the costs related to each of the ensemble members the worst case scenario can be evaluated, *i.e.* the ensemble member for which the costs are highest of all. In figure 5.4 the worst-case cost differences are presented. It can be seen here that the worst-case scenarios always lead to increased costs and these costs are considerably larger than the realized costs (figure 5.2).

5.3.2 Closed loop evaluation

In the specific RHOC implementation every six hours the newly calculated controls are implemented. This means that the uncertainty in the cost function after six hours needs to be evaluated. In figure 5.5 the the differences between calculated costs using nominal weather and actual costs based on observed weather are presented for each control run at $k = 6$. The increased and decreased costs can be clearly seen from this figure. If all the costs shown in this figure are summed the additional costs are known for the given period. These costs as well as the relative contribution to the total costs at the final time are given in table 5.1.

From the a priori known weather forecast uncertainty from the ensemble the

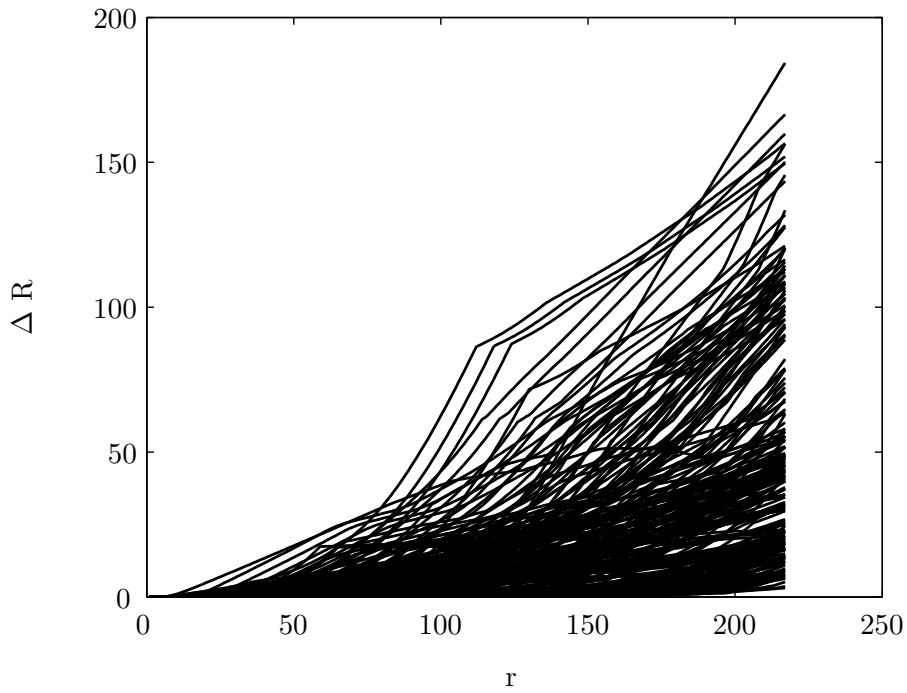


Figure 5.4: Difference in calculated absolute costs and maximum costs based on the ensemble of the 203 optimal control runs

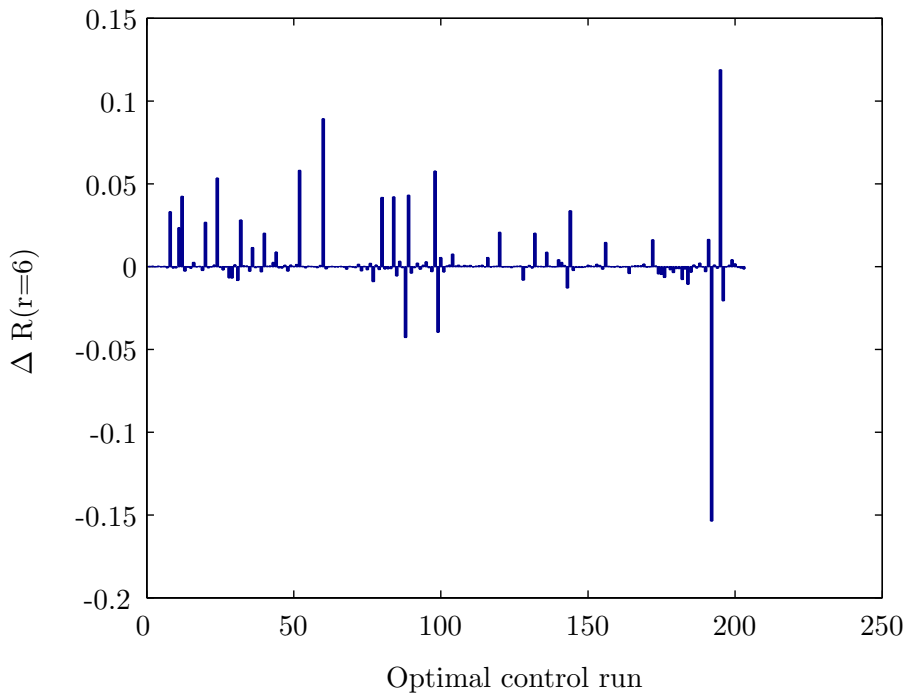


Figure 5.5: Difference in calculated costs and realized costs of each of the 203 optimal control runs after six hours

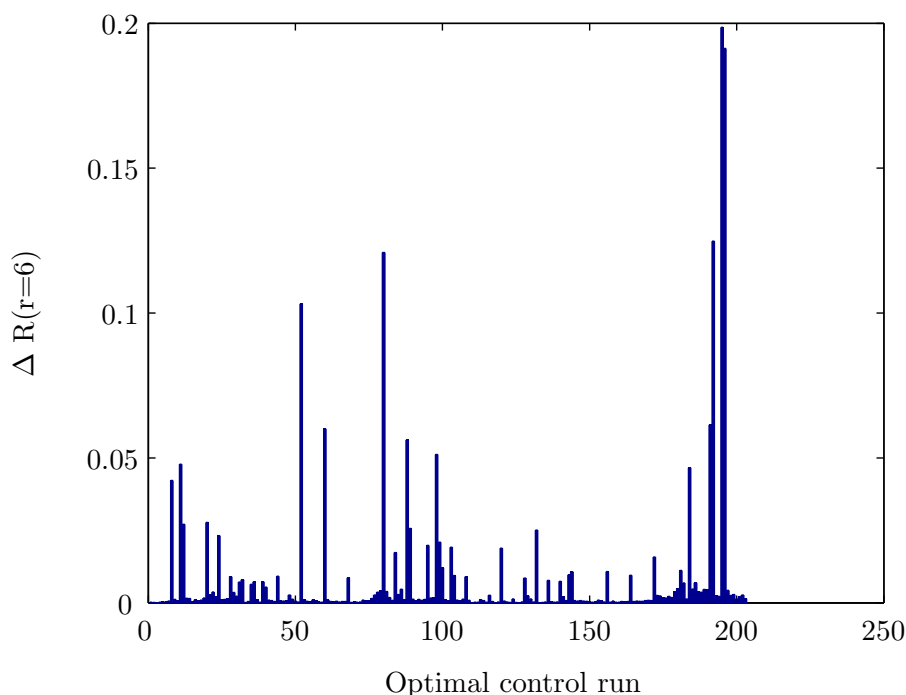


Figure 5.6: Difference in calculated costs and maximum costs based on the ensemble of each of the 203 optimal control runs after six hours

worst-case scenario can be calculated. The additional worst-case costs at the moment that a new optimal control trajectory was calculated, *i.e.* six hours, are presented in figure 5.6 for each of the calculated RHOC runs. Again, the figure shows that the costs always increase. The total additional costs (see table 5.1) in this case, however are still very low.

5.4 Model uncertainty

The model was calibrated with available data from a storehouse. A perfect fit, however, is not a realistic assumption because of modeling deficiencies, changing internal conditions and so on. From the calibration it results that the model lacks information to be a precise predictor. In general, the model fits quite well but at some instances misfits occur.

To evaluate the effect of an imperfect model, model parameters are varied and this modified model is assumed to be the actual system. The model is changed by changing the parameters p_1 , p_2 and p_3 of eqn. (5.1). The changes may vary between 50% and 150% of the nominal value. The selected parameter variability is chosen quite arbitrarily but are within range with the estimated parameters of several different storehouse compartments. Fifty parameter sets are obtained with Latin hypercube sampling, assuming that the parameters are uncorrelated. The

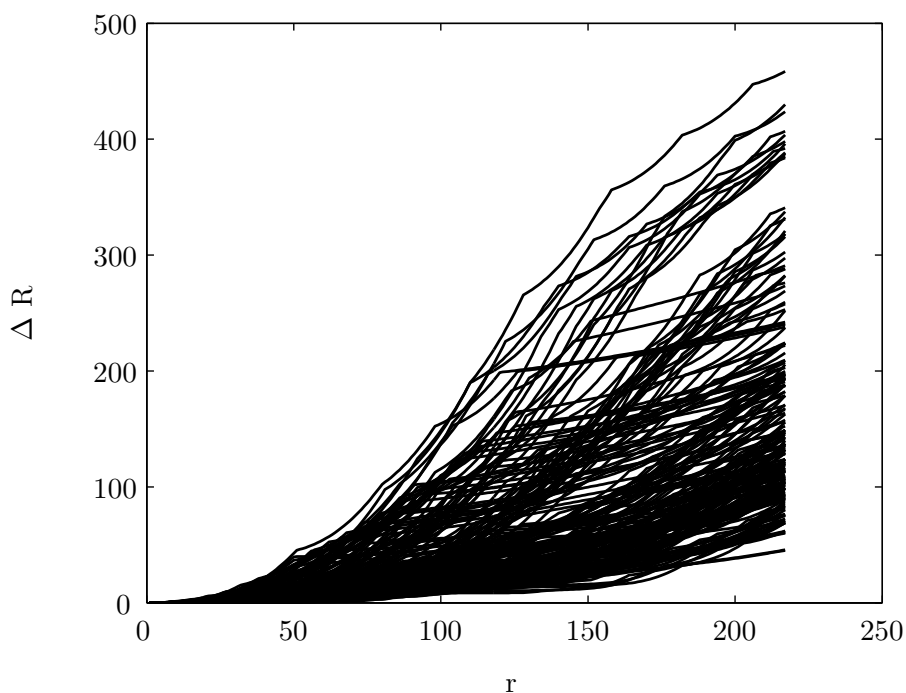


Figure 5.7: Difference in calculated nominal costs and upper bound of costs based on the parameter ensemble of each of the 203 optimal control runs

Table 5.1: Additional costs of the realized costs and the worst-case scenarios for model and weather forecast uncertainty.

	$\Sigma(\Delta J)$	$\Sigma(\Delta J)/J_{tot}$
observations	0.47	0.10%
weather ensemble	1.72	0.37%
parameter ensemble	15.4	3.29%

worst-case scenarios for all 203 optimal control runs are given in figure 5.7. For the given parameter sets the costs increase considerably. To get more insight in the closed-loop uncertainty, the maximum cost differences at the 6 hour forecast for each of the subsequent control runs are given in figure 5.8. At this specific horizon the cost increase is within reasonable limits as can be seen in table 5.1.

5.5 Discussion

If in open loop the calculated costs based on the weather forecast and the *actual* costs given observed weather and realized controls are compared, it is evident that the costs can both increase and decrease. Not only the total costs are of importance but also the relative costs. During a long warm period the optimal calculated costs will increase significantly. The absolute increase or decrease due to uncertain weather forecasts can then be quite large whereas the relative change may be reasonable. The

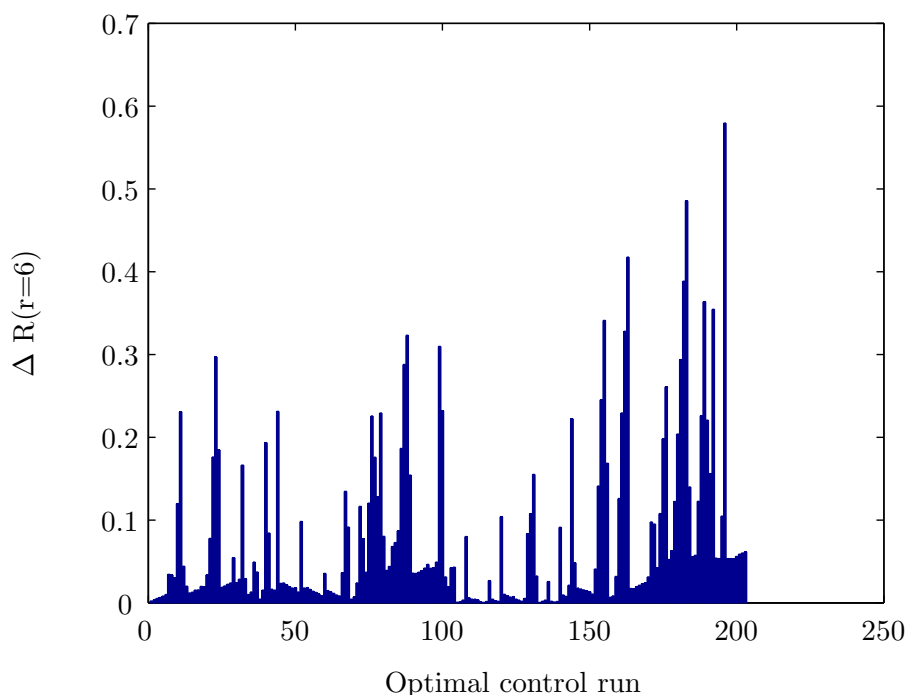


Figure 5.8: Difference in calculated nominal costs and upper bound of costs based on the parameter ensemble of each of the 203 optimal control runs after six hours

opposite may also occur, *i.e.* large relative change versus a low absolute change.

Prior knowledge about the uncertainty of the forthcoming weather is very useful to study uncertainties in the near-future costs. From figures 5.2 and 5.4 it can be calculated that for this specific case the cost-increase is in general much larger for the worst-case scenario than for the observed weather. The same can be seen in the closed loop control in figures 5.5 and 5.6. From these figures it can be seen that the ensemble, indeed, is a useful tool to evaluate the uncertainty of the cost function a priori.

From the chosen parameter variability it can be seen that the worst-case scenario generally has a larger effect on the cost function than the worst-case weather uncertainty. The actual uncertainty based on the true parameters, however, could not be evaluated as these were not available. It is expected that, in analogy to the weather forecast uncertainty, the actual cost increase is lower and in some optimal control runs even decreases.

It has been shown in this case study that applying feedback every six hours reduces the uncertainty of the calculated costs tremendously. Given an uncertainty interval the maximum feedback time can be calculated as well. In table 5.2 the maximum deviation in the cost function from the weather forecast uncertainty is given for different time-intervals between two subsequent optimal control runs (l). It can be seen that the difference in costs up to 24 hours is relatively small. Hence, in case of a

Table 5.2: Additional costs of the realized costs and the worst-case scenarios for weather forecast uncertainty.

l	$\Sigma(\Delta J)/J_{tot}$
6	0.37%
12	1.15%
18	2.05%
24	3.17%
48	8.0%
72	14.2%

communication failure between optimization algorithm and control computer of less than *e.g.* one day no manual intervenience is required.

5.6 Concluding Remarks

In post-harvest storage of agricultural produce, optimal control strategies can be used to anticipate on future weather conditions. In the simulation case-study with real weather forecasts and observed weather it appeared that there are only slight cost increases due to uncertainty in weather forecasts if the optimal control problem is calculated every 6 hours in a RHOC context. Furthermore, even the a priori known worst-case scenario from the weather forecast ensemble leads to a satisfactory result with a 6 hour interval RHOC. The interval between two subsequent optimal control runs can be increased given a user defined uncertainty limit. In the shown case an increase up 24 hours leads to a maximum increase of less then 5%.

If the model is known to be inaccurate it is expected that the uncertainty of the total costs are more subject to model uncertainty than to weather forecast uncertainty. For the investigated case, the worst-case scenario still showed an increase of the total cost of less than 5% which seems to be acceptable. The effect of the actual model uncertainty, however, can only be calculated a posteriori.

Part II

Parameter Estimation

Chapter 6

Parameter Estimation and Prediction of a Rational Storage Model: an ordinary least squares approach

This chapter has been published as:
Doeswijk, T.G. and K.J. Keesman. Parameter Estimation and Prediction of a Nonlinear Storage Model: an algebraic approach. In *2005 International Conference on Control and Automation*, Budapest, Hungary, 2005

6.1 Introduction

In general, the model structure is an approximate representation of *e.g.* a specific physical process. Hence, there is a need to fit the model to the experimental data by parameter estimation. More specifically, as many physical models of real processes are nonlinear in the parameters, the model is fitted by nonlinear parameter estimation. In many applications, such as in optimal control studies, predictive quality is very important. Hence, the model with the estimated parameters should be validated on a different data set.

In literature various parameter estimation methods have been proposed. Among them the ordinary least-squares estimator, using matrix calculus, (see *e.g.* [Lju87] and [Nor86]) is most frequently used. Various extensions of this method have been proposed in the past. For instance, for solving nonlinear least-squares problems, different numerical optimization-based procedures are available, see [DS96]. The key problems encountered in nonlinear least-squares estimation, especially in non-convex optimization problems, are the existence of local minima and the limited amount of parameters that can be reasonably estimated. Another extension is recursive least-squares estimation, where the parameters are updated every time instant new data becomes available. This procedure can be extended to the nonlinear parameter case, see *e.g.* [Gel74], as well.

The objective of this paper is to show how to estimate parameters in a nonlinear discrete-time model structure using ordinary least-squares techniques. Next to this, the predictive quality of the obtained reparameterized model is shown. Therefore, a method, together with the corresponding conditions on the model structure, is presented to reparameterize models that are nonlinear in their parameters as models linear in their parameters. The parameters of the reparameterized model can then be estimated by ordinary least-squares. Finally, the original model structure is retained but with the new parameters. The method is tested on data and a model of a storage facility. Basically, this paper has been inspired by the work of Ljung and Glad [LG94] and it is a natural extension of the paper by Lukasse et al. [LKvS96].

In section 6.2 the general derivation from a model with a polynomial quotient structure to a model linear in its parameters is presented and demonstrated by some examples. Section 6.3 shows three least-squares estimation techniques to estimate the parameters of a storage model with real data. Some validation results on the storage model are shown in section 6.4. In section 6.5 the estimation results as well as the applicability of the method are discussed. Finally, some conclusions are drawn in section 6.6.

6.2 Modelling

Physical modelling can generate equations that are not only nonlinear in their states and inputs but also in their parameters. If these nonlinear functions can be rearranged and reparameterized such that a function arises that is linear in its new parameters, these parameters can be directly and uniquely estimated using an ordinary least-squares procedure. After rearranging, the reparameterized model can be put in a predictor form. For a discrete-time model this is formally given in the following theorem:

Theorem 6.1. *Given the discrete-time nonlinear model*

$$x_{k+1} = f(Z, p) \quad (6.1)$$

where $Z = (x(k), \dots, x(k - \tau), u(k), \dots, u(k - \tau))$ $k, \tau \in \mathbb{Z}^*$ and $\tau < k$ with τ the time delay. If $f(\cdot)$ is a finite polynomial quotient in the elements of Z and p , the predictor

$$\hat{x}(k + 1) = \tilde{f}(Z, \hat{\theta}) \quad (6.2)$$

is equivalent to (6.1). The unique ordinary least-squares estimate $\hat{\theta}$ is given by

$$\hat{\theta} = (F(\cdot)^T F(\cdot))^{-1} F(\cdot)^T F_0(\cdot) \quad (6.3)$$

provided $(F(\cdot)^T F(\cdot))^{-1}$ exists and where $F(\cdot) = [F_1(x(k + 1), Z) \dots F_n(x(k + 1), Z)]$ and $F_0(\cdot) = F_0(x(k + 1), Z)$.

Proof. If $f(\cdot)$ is a finite polynomial quotient in Z and p , equation (6.1) can be written as

$$x(k + 1) = \frac{g(Z, p)}{h(Z, p)} \quad (6.4)$$

with $g(\cdot)$ and $h(\cdot)$ finite polynomials in elements of Z and p . Then, multiplying both sides with $h(\cdot)$ and rearranging terms by elementary algebraic operations, as addition, subtraction, multiplication and division, results in

$$\begin{aligned} F_0(x(k + 1), Z) &= F_1(x(k + 1), Z)\theta_1 + \dots + F_n(x(k + 1), Z)\theta_n \\ &= [F_1(\cdot) \quad F_2(\cdot) \quad \dots \quad F_n(\cdot)] \theta \end{aligned} \quad (6.5)$$

with, for $i = 1, \dots, n$: $\theta_i = \varphi_i(p)$, a polynomial quotient in p . A unique least-squares estimate $\hat{\theta}$, given by (6.3), exists and is equal to θ , if $F(\cdot)$ has rank n and thus $(F(\cdot)^T F(\cdot))^{-1}$ exists. Since $F_0(x(k + 1), Z), \dots, F_n(x(k + 1), Z)$ are again finite polynomial quotients, equation (6.5) can be rearranged such that $\hat{x}(k + 1) = \tilde{f}(Z, \hat{\theta}) = \tilde{f}(Z, \theta)$ with $\tilde{f}(\cdot)$ a finite polynomial quotient. After re-substitution using $\theta_i = \varphi_i(p)$ and rearranging terms: $\hat{x}(k + 1) = \tilde{f}(Z, \theta) = f(Z, p)$. \square

Remark 6.1. Note that the model (6.1) is noise free. Then the estimate $\hat{\theta}$ is optimal, i.e. $\hat{\theta} = \theta$, which does not always hold for p using nonlinear least squares as will be demonstrated in example 6.1. However, if noise is present the estimate is only unbiased if the noise sequence $\{e\}$ in $F_0(\cdot) = F(\cdot)\theta + e$, is uncorrelated with $F(\cdot)$ and has zero mean [Nor86].

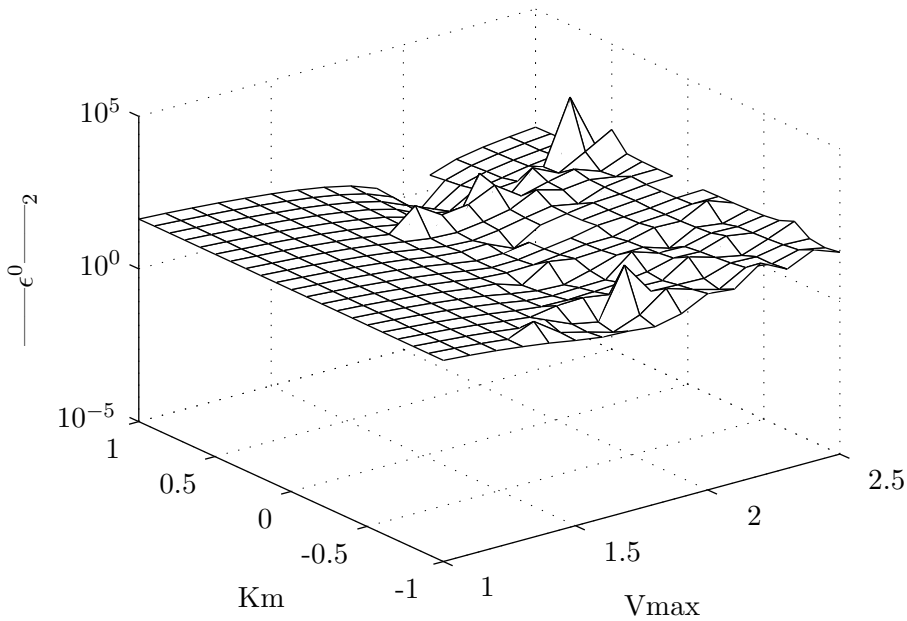


Figure 6.1: Norm of the prediction error with $K_m = 1$, $V_{max} = 2$.

Remark 6.2. The reparameterized model (6.4) can also be put in an output error form by adding the output equation

$$y(k) = x(k) + e(k) \quad (6.6)$$

The ordinary least squares estimate is then no longer optimal and the estimate may be biased. However, the estimate could be optimal in the sense that the predictor has a minimal mean square error over a validation data-set.

Example 6.1. Let us consider substrate consumption with Michaelis-Menten kinetics in a batch reactor. The substrate concentration is described by the following discrete-time model

$$S(k+1) = S(k) - V_{max} \frac{S(k)}{K_m + S(k)} \quad (6.7)$$

By rearranging (6.7) we get

$$S(k)(S(k+1) - S(k)) = [-(S(k+1) - S(k)) \quad -S(k)] \begin{bmatrix} K_m \\ V_{max} \end{bmatrix} \quad (6.8)$$

which is of the form $F_0(\cdot) = F(\cdot)\theta$ that can be solved by ordinary least-squares. Note that in this case $\theta = p$. The validity of the procedure can further be verified by generating model outputs with $S(0) = 30$, $K_m = 1$, and $V_{max} = 2$. A least-squares estimation using (6.8) gives $\hat{\theta} = [1 \quad 2]^T$

In contrast to linear estimation, the outcome of a nonlinear parameter estimation procedure is not guaranteed as can be seen in figure 6.1. If no physical insight in terms of bounds is used and the estimation procedure is treated as a fitting problem one could easily end up in a local minimum. But also with knowledge about system or parameter bounds care must be taken and the solution due to possible singularities in (6.7) is not always obvious. \square

Corollary 6.2. *If, after rearranging (6.1), in (6.5): $F_0(\cdot) = 0$, then a solution can be obtained from*

$$\hat{\theta} = \ker [F_1(\cdot) \cdots F_n(\cdot)] \quad (6.9)$$

In general, the non-empty solutions are not unique. The number of normalized solutions (nullity(F)) can be obtained from the well-known rank-nullity theorem [Kai80]

$$\text{rank}(F) + \text{nullity}(F) = n \quad (6.10)$$

□

Example 6.2. Given the discrete-time system

$$y(k+1) = y(k) - \frac{a}{bu(k) + c} \quad (6.11)$$

it can be rearranged as:

$$0 = [1 \quad u(k)(y(k+1) - y(k)) \quad y(k+1) - y(k)][a \quad b \quad c]^T \quad (6.12)$$

and the solution is found by:

$$[\hat{a} \quad \hat{b} \quad \hat{c}]^T = \ker([1 \quad u(k)(y(k+1) - y(k)) \quad y(k+1) - y(k)]) \quad (6.13)$$

Now, with $y_0 = 10$, $u = [0 \quad 1 \quad 2 \quad 3]^T$ and $a = 1, b = 2, c = 3$ the output $y = [10 \quad 9.67 \quad 9.47 \quad 9.33]^T$ has been generated. Evaluating (6.13) gives the normalized solution: $[\hat{a} \quad \hat{b} \quad \hat{c}]^T = [0.27 \quad 0.53 \quad 0.80]^T \simeq 0.27[a \quad b \quad c]^T$, indicating that the system is unidentifiable. □

Let us now apply the previous theory to a real-world application, *i.e.* storage of agricultural produce in a storage facility.

A discrete-time nonlinear model describing the temperature of the produce in a storage facility (see appendix B) is given by:

$$\begin{aligned} T_p(k+1) = & \left(\tilde{p}_1 + \frac{\tilde{p}_2}{\tilde{p}_3 + \tilde{p}_4 u(k)} + \frac{\tilde{p}_5}{\tilde{p}_6 + \tilde{p}_7 u(k)} \right) T_p(k) + \frac{\tilde{p}_8 + \tilde{p}_9 u(k)}{\tilde{p}_3 + \tilde{p}_4 u(k)} T_e(k) \\ & + \frac{\tilde{p}_{10} + \tilde{p}_{11} u(k)}{\tilde{p}_6 + \tilde{p}_7 u(k)} X_e(k) + \left(\tilde{p}_{12} + \frac{\tilde{p}_{13}}{\tilde{p}_6 + \tilde{p}_7 u(k)} \right) \end{aligned} \quad (6.14)$$

where T_p is the product temperature, T_e and X_e the external temperature and absolute humidity respectively, u the controlled input, *i.e.* product of ventilation and valve opening, and $\tilde{p}_1 \dots \tilde{p}_{13}$ are functions of physical and design parameters. Equation (6.14) can now be written as

$$\begin{aligned} T_p(k+1) = & \begin{bmatrix} u(k)T_p(k+1) & u^2(k)T_p(k+1) & T_p(k) & u(k)T_p(k) & u(k)^2T_p(k) & T_e(k) \\ u(k)T_e(k) & u^2(k)T_e(k) & X_e(k) & u(k)X_e(k) & u^2(k)X_e(k) & u(k) & u^2(k) & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \cdots \theta_{14} \end{bmatrix}^T \end{aligned} \quad (6.15)$$

where $\theta_i = \varphi_i(\tilde{p}_1, \dots, \tilde{p}_{13})$, a polynomial quotient. Note that after reparameterization, $F_0(\cdot) \neq 0$ in (6.5).

Once the least-squares estimate $\hat{\theta}$ has become available the product temperature can be predicted by

$$\hat{T}_p(k+1) = \frac{\tilde{f}(T_p(k), T_e(k), X_e(k), u(k), \hat{\theta})}{\tilde{g}(u(k), \hat{\theta})} \quad (6.16)$$

with

$$\begin{aligned} \tilde{f}(\cdot) &= T_p(k)(\hat{\theta}_3 + \hat{\theta}_4 u(k) + \hat{\theta}_5 u(k)^2) + T_e(k)(\hat{\theta}_6 + \hat{\theta}_7 u(k) + \hat{\theta}_8 u(k)^2) \\ &\quad + X_e(k)(\hat{\theta}_9 + \hat{\theta}_{10} u(k) + \hat{\theta}_{11} u(k)^2) + \hat{\theta}_{12} u(k) + \hat{\theta}_{13} u(k)^2 + \hat{\theta}_{14} \\ \tilde{g}(\cdot) &= 1 - \hat{\theta}_1 u(k) + \hat{\theta}_2 u(k)^2 \end{aligned}$$

6.3 Estimation

In this section, different least-squares techniques are used to estimate parameters of the storage model (6.14).

6.3.1 Nonlinear least-squares

Depending on the type of problem an appropriate algorithm must be chosen to solve a nonlinear least squares problem [DS96, Nor86]. All methods have in common that the solution is iteratively found. Existence of local minima can be a serious issue (*e.g.* figure 6.1) if initial estimates are inaccurate. Hence, in practice, in particular in non-convex problems, most often the number of parameters to be estimated should be limited. It is therefore necessary to find the most sensitive parameters. A (local) sensitivity analysis can be used to gain insight in the parameter sensitivity. Now, with knowledge of the sensitivities of the parameters, a set of parameters to be estimated is chosen. When physical and design parameters are to be estimated, usually these parameters may not vary unlimited. The least-squares problem then becomes constrained. However, this may lead to a non-optimal estimation in least-squares sense. Let us now demonstrate this to the storage model (6.14).

The parameter vector p of the discrete-time storage model (6.14) consists of several design and physical parameters. The results of a sensitivity analysis are presented in figure 6.2. The parameters c_{pp} , V_p and ρ_p have the same sensitivity. They always appear product-wise and could be replaced by a combined parameter. Hence, only one of these parameters has to be estimated. Given the sensitivity in figure 6.2 and some indication of the uncertainty of the physical and design parameters the following parameters have been selected for estimation: $\hat{\theta} = [P_{resp}, \phi_{max}, c_{pp}, \alpha_{ea}, h]^T$, respectively the respiration heat of the stored product, maximum ventilation flow,

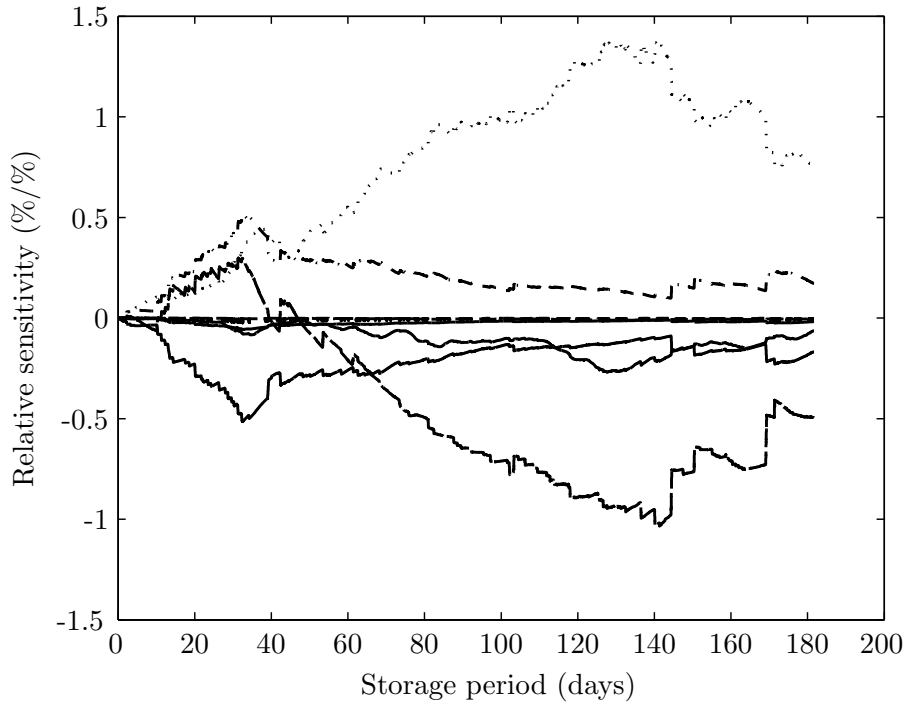


Figure 6.2: Sensitivity of some design and physical parameters of a storage model.

Table 6.1: Nominal and estimated parameter values.

	θ^0	$\hat{\theta}$
P_{resp}	5.42	5.17
ϕ_{max}	0.0972	0.0744
c_{pp}	3600	4277
α_{ea}	0.347	0.347
h	3.5	3

the heat capacity of the stored product, heat transfer from environment to air in bulk and the height of the bulk. Subsequently, a constrained nonlinear least-squares estimation is performed. The nominal values (θ^0) and the estimates ($\hat{\theta}$) are given in table 6.1. Only the parameter estimate of h is at its lower bound; the other parameter estimates are within the associated bounds.

6.3.2 Truncated least-squares

Given a model that is linear in its parameters, *i.e.*

$$y = C\theta + e \tag{6.17}$$

with $y \in \mathbb{R}^N$, $C \in \mathbb{R}^{N \times n}$ and $\theta \in \mathbb{R}^n$. An ordinary least-squares estimate is given by

$$\hat{\theta} = (C^T C)^{-1} C^T y \tag{6.18}$$

However, when the problem is ill-conditioned the estimates are very sensitive to the data. To overcome this problem, the truncated least-squares method can be used. Given the singular value decomposition of C , *i.e.*

$$C = U\Sigma V^T \quad (6.19)$$

with $U \in \mathbb{R}^{N \times N}$, $\Sigma \in \mathbb{R}^{N \times n}$ and $V \in \mathbb{R}^{n \times n}$, one can determine from the singular values (diagonal elements of Σ) whether the problem is ill-conditioned. Premultiplying (6.17) with U^T [Nor86, p.77] gives:

$$\begin{aligned} y^* &= U^T y = U^T C \theta + U^T e \\ &= U^T U \Sigma V^T \theta + U^T e \\ &= \Sigma \theta^* + e^* \end{aligned} \quad (6.20)$$

where $\theta^* = V^T \theta$ and $e^* = U^T e$. The parameters θ_i^* that correspond to singular values that are very small compared to the largest singular value, defined by the numerical rank determinant \mathcal{R} , are set to 0. Now, the modified linear regression model (6.20) can be solved. The sum of squares $(y^* - \Sigma \theta^*)^T (y^* - \Sigma \theta^*)$ is minimized when $\hat{\theta}_i^* = y_i^* / \sigma_i$ where σ are the singular values and $i = 1, \dots, r$ with $r \leq n$ the numerical rank. The parameter estimates can now be obtained from $\hat{\theta} = V \hat{\theta}^*$.

Let us now analyze the storage model (6.15) and add the equation error e_k . Observations from a 50 days period with a sampling interval of 900 seconds are used to estimate the parameter vector θ . Using the singular value decomposition (6.19) we obtain the matrices U, V and Σ related to the regression matrix of (6.15). By choosing an appropriate numerical rank determinant *e.g.* $\mathcal{R} = \bar{\sigma}/1000$, with $\bar{\sigma}$ the largest singular value, the following results are obtained:

$$S = \begin{bmatrix} 617.5027 \\ 225.9039 \\ 206.0795 \\ 27.7489 \\ 15.7058 \\ 8.1518 \\ 2.0180 \\ 1.6773 \\ 0.2375 \\ 0.1451 \\ 0.0270 \\ 0.0246 \\ 0.0019 \\ 0.0003 \end{bmatrix} \quad \hat{\theta}^* = \begin{bmatrix} -0.7604 \\ -0.3290 \\ -0.5414 \\ -0.0237 \\ -0.0196 \\ 0.1382 \\ 0.0052 \\ -0.0041 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\hat{\theta} = \begin{bmatrix} -0.0038 & -0.0007 & 0.9998 & -0.0038 & 0.0007 & 0.0001 & -0.0003 & 0.0043 \\ -6.0 \cdot 10^{-5} & -2.9 \cdot 10^{-6} & 1.6 \cdot 10^{-6} & -0.0009 & -4.6 \cdot 10^{-5} & 0.0024 \end{bmatrix}^T$$

From the right singular vectors (not shown here) it appears that the first singular value is predominantly related to θ_3 and θ_6 and the lowest singular value is related to θ_{10} and θ_{11} .

If physical interpretation of the estimates is desired, it is possible to retain the original physical parameters from the estimated parameters using:

$$\hat{\theta}_i = \varphi_i(\hat{p}) \quad (6.21)$$

where $\hat{\theta} \in \mathbb{R}^n$ is the estimated parameter vector and $\hat{p} \in \mathbb{R}^q$ the physical and/or design parameter vector to be estimated. If $q = n$ direct inversion could be used, while in other cases one should use *e.g.* a minimum length or least-squares approximation. However, and this should be stressed again, for prediction (6.16) can be directly applied.

6.3.3 Recursive estimation

A Kalman filter approach can be used to estimate parameters recursively [Lju87]. The linear regression model of (6.17) should then be written in a form like

$$\theta(k+1) = \theta(k) + w(k) \quad (6.22)$$

$$y(k) = C(k)\theta(k) + v(k) \quad (6.23)$$

Under the assumption that the parameters are constant the covariance matrix $E(w(k)w(k)^T) = 0$. In addition to this, an assumption about the observation noise properties must be made and the initial estimate $\theta(0)$ and initial covariance matrix $P(0)$ must be specified. The parameters of (6.15) are now estimated recursively.

Given a sampling time much smaller than the system time-constant of (6.15) the initial parameter estimate of $\theta_3 = 1$ and all other initial parameter estimates are zero, *i.e.* the model reduces to $T_p(k+1) = T_p(k)$. The initial covariance matrix $P(0) = 0.1 * I$, *i.e.* all parameters are considered to be independent and identically distributed with the knowledge that the actual parameter values are around the initial estimates. The covariance matrix, $R = E(v(k)v(k)^T)$, represents the variance of the observations, *i.e.* the product temperature, and determines the parameter update rate of the filter. In this example it is taken as $1^\circ C^2$. With given y and C from section 6.3.2 the parameter trajectories of θ_1, θ_2 and $\theta_4 \dots \theta_{14}$ are given in figure 6.3. Not shown is θ_3 which remains close to one. From figure 6.3 it can be seen that it takes until day 8 before the parameters start to change. This is due to the fact that this is the first ventilation period of the data-set. Furthermore, it can be seen that

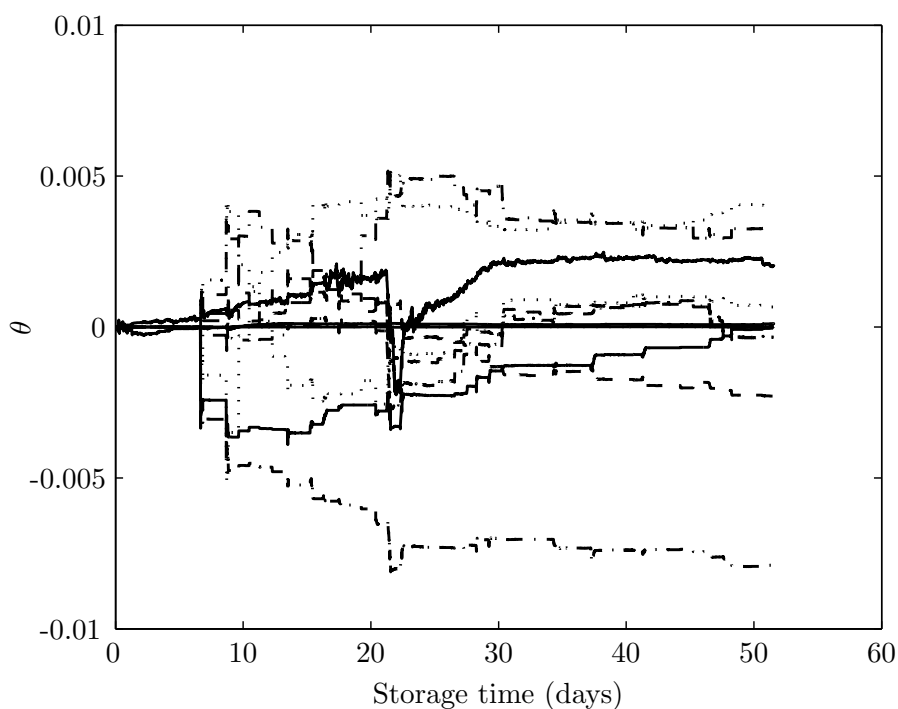


Figure 6.3: Parameter evolution of the recursive parameter estimates θ_1, θ_2 and $\theta_4 \dots \theta_{14}$

the parameters converge in time. The recursive parameters estimates at the final time are given by:

$$\hat{\theta} = \begin{bmatrix} -4.5 \cdot 10^{-5} & 0.0040 & 0.9998 & -0.0079 & -0.0023 & 0.0001 & 0.0007 & 0.0033 \\ 3.9 \cdot 10^{-5} & 4.0 \cdot 10^{-6} & 4.4 \cdot 10^{-6} & -0.0003 & -0.0003 & 0.0021 \end{bmatrix}^T$$

This final parameter vector at day 52 is used in the next section to test the validity of the recursively estimated model.

6.4 Validation

In this section the estimated parameters from section 6.3 are validated by evaluating the model predictions obtained from an open loop simulation. First, the same period over which the calibration was performed is considered and then a validation period is chosen with about the same length.

First, the mean-square error (MSE) and a graphical presentation of the results of the calibration period are given in table 6.2 and figure 6.4. It can be seen that the MSE of the recursive estimate and the truncated least-squares estimate are within the same range. The physical model performs slightly worse. Notice furthermore, that the sudden changes of the recursive estimates of figure 6.3 correspond to the fast dynamics, *i.e.* ventilation, of figure 6.4.

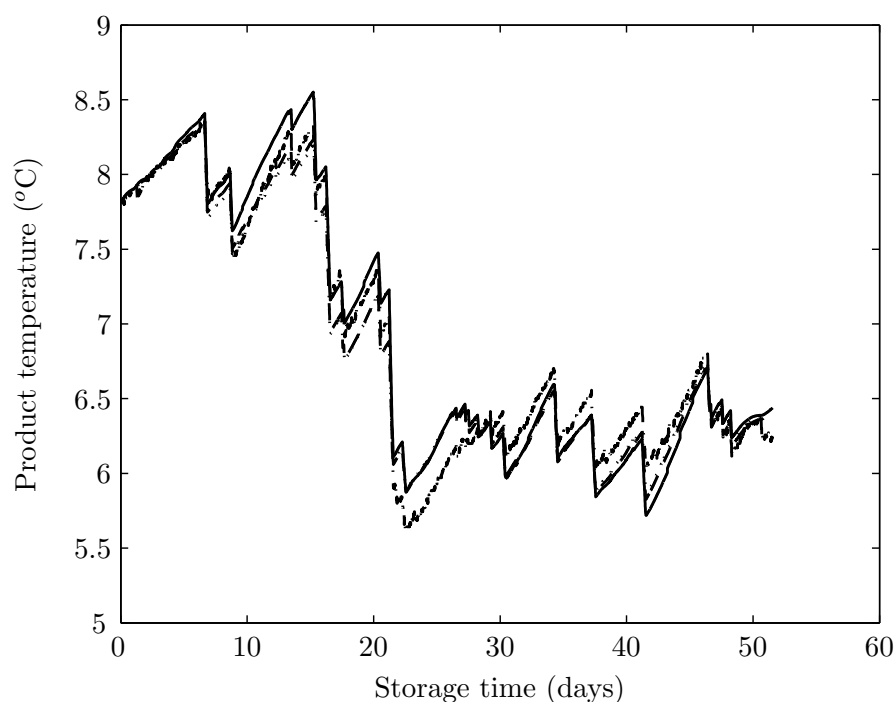


Figure 6.4: Measured and predicted product temperatures in the calibration period (\cdots physical model, $\cdot - \cdot$ truncated least squares, $- - -$ recursive estimate, $—$ measurements)

A validation step is performed over a different period in the same storage season. Hence, for the same storage facility and product, the same parameter vector as is obtained from calibration could be used. The results of the validation are given in table 6.2 and figure 6.5. The model predictions with the truncated least-squares estimate and the recursive estimate perform within the same range. The physical model predicts significantly worse.

6.5 Discussion

Reparameterizing and rearranging a discrete-time model with polynomial quotient structure into a model linear in its parameters is helpful in getting optimal parameter estimates in least-squares sense if the model is noise free. The physical interpretation of the variables is lost but the main structure is conserved. In contrast, black-box modelling on the basis of neural networks or nonlinear regressions can also be used.

Table 6.2: mean squared errors of calibration and validation period

	physical model	truncated	recursive
calibration	0.027	0.020	0.021
validation	0.113	0.032	0.040

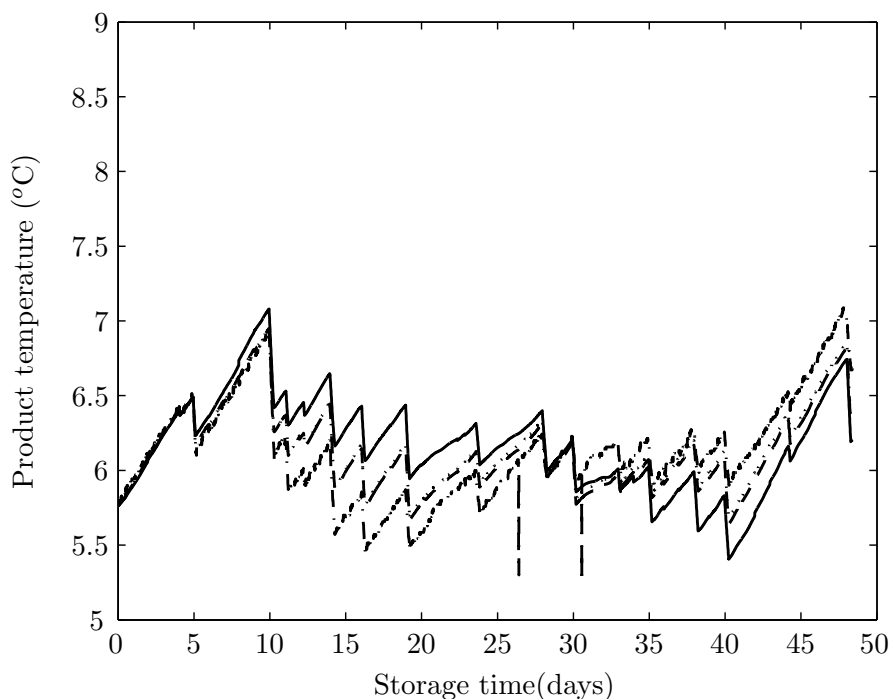


Figure 6.5: Measured and predicted product temperatures in the validation period (\cdots physical model, $\cdot - \cdot$ truncated least squares, $- - -$ recursive estimate, $—$ measurements)

However, in those cases the system order must be estimated from the data. Usually, no direct information about the physical system is then taken into account.

Some difficulties may arise in least-squares estimation of the reparameterized system. Rearranging and reparameterizing may lead to correlated columns and interdependent parameters, which lead to (near) rank deficient problems. This problem can be solved by using truncated least-squares (see section 6.3.2). But, the main problem is that due to the reparameterization there is not only (structural) unknown uncertainty in the output vector but also in the data matrix (see equation (6.5)). This problem can be tackled by a total least-squares approach [GVL80] and will be subject of another paper.

In the storage model example the inputs, as external temperature (T_e) and humidity (X_e), are measured variables, which directly leads to an errors-in-variables problem. Consequently, a nonlinear estimation procedure should take this into account, as well. However, the key problem in this example was the numerical rank deficiency of the data matrix. This problem not only occurred due to the introduction of correlated columns, but also due to the experimental setup. First, the fan is mainly controlled as on/off and it is switched on for less than 10% of the time. Therefore, the columns of C that are multiplied by $u(k)$ and $u^2(k)$ will contain many zeros and as such will be highly correlated. In addition to this, the control frequency

was quite high, *i.e.* 15 minutes compared to the time constant of the system which is approximately 19 hours with maximum ventilation and 20 days with no ventilation, where the optimal sample time is about 1/10 of the time constant [Lju87], and thus leading to high correlations.

Truncated least-squares came out quite well in the example. However, an appropriate value for the numerical rank determinant must be chosen. Here it is chosen by tuning. The nonlinear least-squares estimation performs worse for both calibration as well as for validation. From table 6.1 it can be seen that the parameter h is bounded at its lower bound, which indicates that the estimate is non-optimal. Furthermore, the solution can be in a local minimum. The recursive estimate performs well for both calibration and validation period.

6.6 Concluding remarks

It has been demonstrated that a discrete-time model with polynomial quotient structure in input, output, and parameters can be rearranged and reparameterized such that a model arises that is linear in its parameters. Consequently, the parameters can be uniquely estimated by ordinary least-squares methods. Furthermore, it has been shown that using the estimated parameters in the back-transformed model (via rearranging terms) leads to a predictor with fairly good open loop predictive performance in the sense of mean-square error. Further research will focus on a total least-squares approach to account for the errors-in-variables problem.

Chapter 7

Parameter Estimation and Prediction of a Rational Storage Model: a total least squares approach

This chapter has been published as:
Doeswijk, T.G. and K.J. Keesman. Parameter estimation and prediction of nonlinear biological systems: some examples. In *14th IFAC Symposium on System Identification*, Newcastle, Australia, 2006

7.1 Introduction

In general, in a model calibration procedure, the parameters of the model are estimated such that the model predictions fit well on measured data. Several techniques are available to obtain suitable estimates. For models that are *linear* in their parameters, the ordinary and weighted least squares techniques [Nor86, Lju87] are probably the most frequently used. However, when errors are not only present in the current output but also in the previous outputs and inputs, a modification of the ordinary least squares technique is required. The overdetermined set of linear equations is then given by:

$$A\theta \approx b \quad (7.1)$$

where $A \in \mathbb{R}^{m \times n}$ with $m > n$ the data matrix, $\theta \in \mathbb{R}^n$ the linear parameters and $b \in \mathbb{R}^m$ the output. The total least squares (TLS) approach [GVL80] provides a solution if errors are not only present in the output b but also in the data matrix A . The TLS technique is currently widely used in several application areas, *e.g.* [VHL02] and references in [VHV91], and examples are present where TLS shows its superiority over ordinary least squares.

Usually, for *nonlinear* in the parameter models, parameter estimates are found iteratively using optimization algorithms. The existence of local minima and high computational effort may impede the application of nonlinear estimation methods.

Given the nonlinear discrete-time model

$$x(k+1) = f(Z, p) \quad (7.2)$$

where $Z = (x(k), \dots, x(k-\tau), u(k), \dots, u(k-\tau))$, $k, \tau \in \mathbb{Z}^+$ and $\tau < k$ with τ the time delay, and $f(\cdot)$ a *finite polynomial quotient* in the elements of Z and $p \in \mathbb{R}^q$. The model can be rearranged and reparameterized such that a model arises that is linear in its (new) parameters θ , a polynomial quotient in p . These parameters can then be estimated with a non-iterative least squares estimator. If noise is added to (7.2), then by rearranging the model, errors become part of the data matrix. Hence, in general the parameter estimation problem becomes an errors-in-variables (EIV) problem for which TLS provides a natural solution. Finally, the linear in the parameters model can be rewritten in predictor form, *i.e.*

$$\hat{x}(k+1) = \tilde{f}(Z, \hat{\theta}) \quad (7.3)$$

where $\hat{\theta}$ contains the least squares estimate in the linearly reparameterized model. Note that the original parameter vector p is *not* necessarily re-estimated from $\hat{\theta}$. In what follows, the main focus is on the predictive quality of (7.2) and (7.3).

The validity of reparameterization and linear estimation was illustrated in chapter 6 with an example of Michaelis-Menten kinetics. In the *noise free* case the linearly

reparameterizing method led to the exact solution where the nonlinear least squares approach could end up in local minima. Furthermore, the linear regressive reparameterization approach was applied to a storage facility containing a biological product. Real data were used to evaluate the predictive quality of (7.3) with θ estimated with a truncated least squares estimator. The results were compared with the original model (7.2) where the parameters p were estimated by a traditional nonlinear estimation approach. The linearly reparameterized storage model, however, got an EIV structure because errors appeared in the data matrix. In chapter 6 this was neglected initially but it has been noticed that a TLS approach would be more appropriate for parameter estimation than the truncated least squares method. To our knowledge this reparameterization, analysis and application of a TLS approach to a *nonlinear in the parameter model* has never been explicitly reported.

The objective of this paper is to evaluate the predictive quality of (7.3) with θ estimated with OLS and TLS and compare this with the predictive quality of the *nonlinear* original model (7.2) with p estimated with a nonlinear least squares technique. First, the Michaelis-Menten model is used with simulated noisy data. Second, a storage model with real data is used.

7.2 Background

7.2.1 Algebraic non-linear parameter estimation

If $f(\cdot)$ is a polynomial quotient in Z and p , reparameterizing the original model (7.2) leads to

$$F_0(x(k+1), Z) = [F_1(\cdot) \quad F_2(\cdot) \quad \dots \quad F_n(\cdot)] \theta \quad (7.4)$$

with, $F_i(\cdot) = F_i(x(k+1), Z)$, $i = 1, \dots, n$. The new parameters are given by $\theta_i = \varphi_i(p)$, a polynomial quotient in p . The model (7.4) is a linear regressor (see also (7.1)) and the parameters θ can be estimated by least squares.

As $F_i(\cdot)$ can be a function of $x(k+1)$, even in case of an equation error structure, the data matrix may contain errors. Furthermore, the columns $F_i(\cdot)$ can become linear dependant which leads to (near) rank deficiency. Finally, nonlinearities can occur in uncertain regression variables which in turn can lead to biased estimates. Regularization and bias compensation therefore might be needed.

After estimating the parameters θ , the model can be rewritten in predictor form (7.3). Remark that no effort is done to estimate the original parameters p from $\hat{\theta}$ (see for details chapter 6).

7.2.2 Michaelis-Menten kinetics

A discrete-time model describing the substrate concentration in a batch bioreactor with Michaelis-Menten kinetics is given by:

$$S(k+1) = S(k) - V_{\max} \frac{S(k)}{K_m + S(k)} \quad (7.5)$$

with S the substrate concentration, V_{\max} the maximum substrate conversion rate and K_m the Michaelis-Menten constant. Estimating K_m and V_{\max} may lead to local minima (see figure 6.1). Rearranging (7.5) leads to:

$$S(k)(S(k+1) - S(k)) = [-(S(k+1) - S(k)) \quad -S(k)] \begin{bmatrix} K_m \\ V_{\max} \end{bmatrix} \quad (7.6)$$

which is linear in the parameters. Remark that in this case no reparameterization has taken place, *i.e.* $\theta = p$. As the substrate concentration is the measured (noisy) variable in this system, it can be clearly seen that the data matrix A , in this case $[-(S(k+1) - S(k)) \quad -S(k)]$ for $k = 1, \dots, m$, contains measurement errors and hence, (7.6) has become EIV.

7.2.3 Storage model

A model that describes temperature dynamics in storage facilities for biological products such as fruits and vegetables, is given by (see appendix B:

$$T_p(k+1) = \left(\tilde{p}_1 + \frac{\tilde{p}_2}{\tilde{p}_3 + \tilde{p}_4 u(k)} + \frac{\tilde{p}_5}{\tilde{p}_6 + \tilde{p}_7 u(k)} \right) T_p(k) + \frac{\tilde{p}_8 + \tilde{p}_9 u_{k-1}}{\tilde{p}_3 + \tilde{p}_4 u(k)} T_e(k) \\ + \frac{\tilde{p}_{10} + \tilde{p}_{11} u(k)}{\tilde{p}_6 + \tilde{p}_7 u(k)} X_e(k) + \left(\tilde{p}_{12} + \frac{\tilde{p}_{13}}{\tilde{p}_6 + \tilde{p}_7 u(k)} \right) \quad (7.7)$$

where $T_p(k+1)$ is the measured output variable. The variables T_p denotes the temperature of the produce ($^{\circ}C$), T_e the external temperature ($^{\circ}C$) and X_e the external absolute humidity (kg/kg). Finally, the input u denotes the product of fresh inlet and ventilation and is bounded by: $0 \leq u \leq 1$. The nonlinearities in (7.7) are related to heat and mass transfer.

Rearranging (7.7) into a linear regression format leads to:

$$T_p(k+1) = \begin{bmatrix} u(k)T_p(k+1) & u(k)^2 T_p(k+1) & T_p(k) & u(k)T_p(k) & u(k)^2 T_p(k) & T_e(k) & u(k)T_e(k) \\ u(k)^2 T_e(k) & X_e(k) & u(k)X_e(k) & u(k)^2 X_e(k) & u(k) & u(k)^2 & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \cdots \theta_{14} \end{bmatrix}^T \quad (7.8)$$

where $\theta_i = \varphi_i(p)$.

As can be seen from (7.8) the model clearly is EIV as $T_p(k+1)$ is no longer only the output variable but now also appears in the data matrix. Furthermore, $T_p(k)$,

$T_e(k)$ and $X_e(k)$ are also measured uncertain variables. After estimating θ the model (7.8) can be rewritten in predictor form

$$\hat{T}_p(k+1) = \frac{\hat{\theta}_3 T_p(k)}{1 - \hat{\theta}_1 u(k) - \hat{\theta}_2 u(k)^2} + \dots + \frac{\hat{\theta}_{13} u(k)^2}{1 - \hat{\theta}_1 u(k) - \hat{\theta}_2 u(k)^2} + \hat{\theta}_{14} \quad (7.9)$$

In what follows, the quality of the predictor (7.9) will be evaluated in terms of the mean square error (MSE) of the prediction errors.

7.3 Estimation methods

To evaluate the performance of predictors (7.2) and (7.3) with estimates \hat{p} and $\hat{\theta}$, several estimation methods are compared.

A nonlinear parameter estimation procedure for estimating the parameters of (7.7) is used in this paper as a reference for the alternative estimation methods. The vector p of (7.7) consists of several physical and design parameters. A selection of these physical and design parameters has been estimated using a nonlinear least squares procedure.

For systems linear in their parameters the ordinary least squares procedure is widely used. For ill-conditioned systems, however, the estimates are very sensitive to the data and hence, the errors therein. The predictive quality of the original system with the estimated parameters is then very poor. A regularization method to overcome these limitations is the truncated least squares method which uses the numerical rank of the data matrix to stabilize the solution [Nor86, p.77].

An underlying assumption of the ordinary least squares estimator is that all errors are subjected to the output vector. However, frequently errors are not only present in the output vector but also in the data matrix. A fitting technique that compensates for errors in the data is TLS. [GVL80] outlined a TLS-solution of the EIV problem which is heavily based on the singular value decomposition (SVD) of (7.1), *i.e.*:

$$U\Sigma V^T = [A, b] \quad (7.10)$$

with $U \in \mathbb{R}^{m \times (n+1)}$ and $V \in \mathbb{R}^{(n+1) \times (n+1)}$ orthogonal and $\Sigma \in \mathbb{R}^{(n+1) \times (n+1)}$ diagonal, containing the singular values. Their method assumes that the errors are independent and identically distributed. Then, if the absolute size of the errors are all roughly equal [VHV91] good results can be obtained.

Several extensions of the TLS algorithm have been proposed. In this paper the generalized TLS (GTLS) [VHV89] algorithm is used. GTLS overcomes the limitation of independent and identically distributed errors. GTLS considers different sized and correlated errors as well as error-free variables where the covariance matrix of the

Table 7.1: Nominal and estimated parameter values for different numerical ranks.

θ^0	$\hat{\theta}(\mathcal{R} = 5)$	$\hat{\theta}(\mathcal{R} = 4)$	$\hat{\theta}(\mathcal{R} = 3)$	$\hat{\theta}(\mathcal{R} = 2)$
0.9500	0.9643	0.9600	0.9527	0.9827
0.1000	2.1947	-1.7327	0.0903	0.0355
0.2000	-8.2878	-0.2826	0.1877	0.0706
0.3000	5.1155	1.1424	0.2879	0.1061

errors must be known or estimated. If there are no error-free data and the covariance matrix equals the identity matrix the GTLS solution is equal to the classical TLS solution. The key problem is how to choose the covariance matrix if the error distributions are unknown.

In the storage model example, the matrix of interest for GTLS estimation, *i.e.* $[A, b]$, of the linearly reparameterized model (7.8) has a large condition number ($> 10^6$). The large condition number implies that some columns are highly correlated which leads to (near) rank deficiency. This, in turn, leads to an estimator that is very sensitive to the data. Hence, a regularization method was chosen. Both truncated least squares and the nongeneric GTLS use a numerical rank to overcome the problem of ill-conditioning. By choosing an appropriate numerical rank a compromise is found between the stabilization of the solution and the accuracy of the GTLS estimator.

Let us, as an illustrative example, consider the following dynamic system:

$$\begin{aligned}x(k+1) &= \theta_1 x(k) + \theta_2 u_1 + \theta_3 u_2 + \theta_4 u_3 \\y(k) &= x(k)\end{aligned}\tag{7.11}$$

where u_1 , u_2 and u_3 are all constant inputs related by $u_3 = 1.5u_2 = 3u_1$. Given the set of nominal parameters, the constant input signals and the initial condition, the system is simulated for 1000 steps. All columns of the data matrix $A = [y(k), u_1, u_2, u_3] \in \mathbb{R}^{m \times n}$ and output vector $b = y(k+1) \in \mathbb{R}^{m \times 1}$ are then corrupted with additive independent and identically distributed zero mean gaussian white noise. The matrix $[A, b]$ is now ill-conditioned due to the near linear dependency. However, the parameters can still be identified with nongeneric GTLS by evaluating the numerical rank \mathcal{R} of $[A, b]$. From table 7.1 it can be seen that for $\mathcal{R} = 3$ the estimates approach the nominal parameters. Notice from figure 7.1 that the predictive value of the estimates is best in terms of MSE for numerical rank $\mathcal{R} = 3$ and the predictive performance dramatically decreases for higher (and lower) numerical rank.

The GTLS is elaborated for the bioreactor model (7.6) and for the storage model (7.8). In section 7.4 the predictive quality of the predictors (7.5) and (7.9) with GTLS estimates are evaluated.

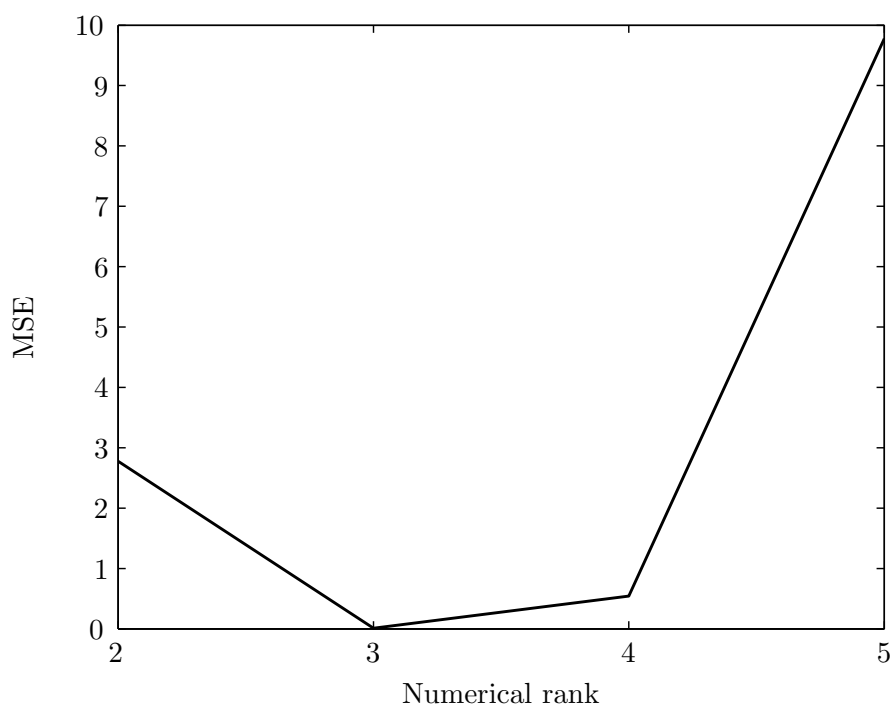


Figure 7.1: MSE of simulated output with nominal parameters minus simulated output with estimated parameters for various numerical ranks

7.4 Results

7.4.1 Michaelis-Menten kinetics

It is known that the parameters of the Michaelis-Menten equation (7.5) in a batch reactor are theoretically identifiable if the initial conditions are known [God83]. If noise is present, however, the parameters, specifically K_m , are hard to identify [Hol82]. Furthermore, a nonlinear estimation method can suffer from local minima. In the following, a comparison between ordinary least squares and GTLS using (7.5) and simulation data is made. The aim of this simulation experiment is to distinguish between the predictive quality of (7.5) with the least squares estimates versus the GTLS estimates.

First, the substrate concentration was simulated for $k = 1, \dots, m$, with $m = 30$ using (7.5) and where $K_m = 10$, $S_0 = 30$, and $V_{\max} = 2$. In addition, a Gaussian white noise sequence is generated with zero mean and unit variance. Next, the noise sequence is multiplied with a factor varying from 0.01 to 0.2 to obtain a range of standard deviations of the measurement noise. The noise corrupted substrate concentrations were then used to generate the data vectors: $-S(k)$ and $\frac{S(k)}{S(k+1)-S(k)}$ for $k = 1, \dots, m$. Given these data K_m and V_{\max} could be estimated. This procedure was repeated 100 times. The mean of the estimated parameters \hat{K}_m and \hat{V}_{\max} for each standard deviation were then used for open loop prediction of the substrate

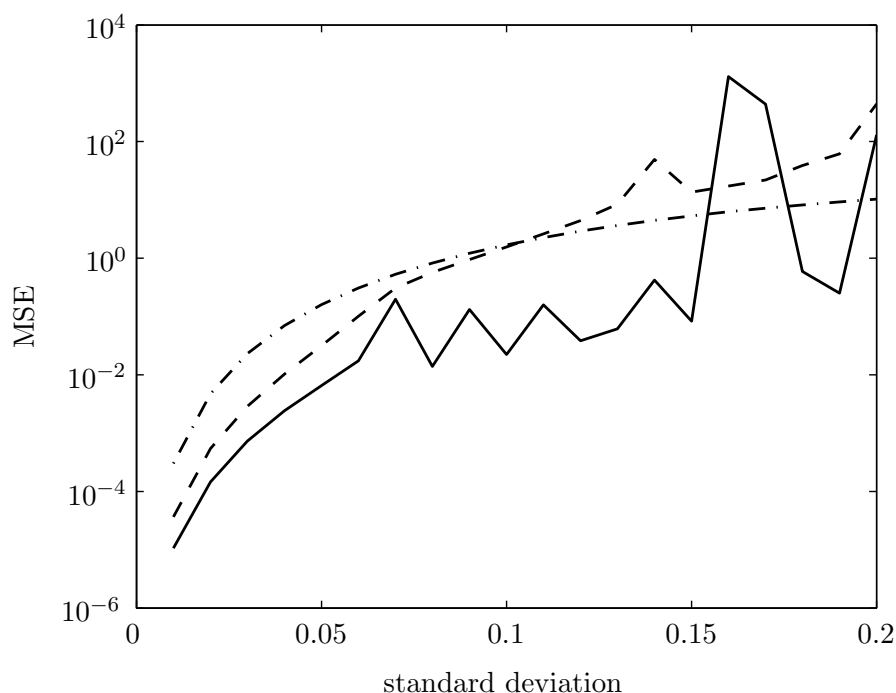


Figure 7.2: MSE of predictors (7.5) with parameters estimated by LS of (7.6) ($\cdot - \cdot$) and of (7.12) ($- - -$), and by GTLS of (7.12) ($—$).

concentration with (7.5) and $S_0 = 30$.

In order to reduce the number of data errors, the modified Michaelis-Menten equation (7.6) was further rearranged such that only one column of the data-matrix contains errors, *i.e.*

$$-S(k) = K_m + \frac{S(k)}{S(k+1) - S(k)} V_{\max} \quad (7.12)$$

Next, for applying GTLS, the covariance matrix had to be chosen. Given the error $e(k) = b(k) - S(k)$ the variance of the regressor of V_{\max} was assumed to be much larger than the variance of the output error because the error of the denominator $e(k+1) - e(k)$ can become close to zero. The covariance matrix was chosen as:

$$\text{cov}\left(\left[\frac{e(k)}{e(k+1) - e(k)}, -e(k)\right], -e(k)\right) = \begin{bmatrix} 10^6 & 0 \\ 0 & 1 \end{bmatrix}$$

Subsequently, the parameters were estimated with the least squares estimators related to (7.6) and (7.12) and with the GTLS estimator related to (7.12). The MSE of the prediction error ($\frac{1}{m} \sum_{k=1}^m (S_k - \hat{S}_k)^2$) is given in figure 7.2.

GTLS clearly outperforms the ordinary least squares. At higher noise levels, however, the estimates may induce singularities. At this point the ordinary least squares estimates already have a poor predictive performance.

7.4.2 Storage model

Let us consider again eqn (7.8). If the size of the errors of each column of the matrix $[A, b]$ correspond to the absolute values of the measurements then the errors are not identically distributed. The absolute humidity X_e is an order of magnitude of 3 smaller than T_e and T_p . Furthermore, the last column of the data matrix A is error free. Hence, better estimates are expected for GTLS compared to truncated least squares.

For applying GTLS a proper covariance matrix must be chosen. As the exact sizes of the errors were not known these were approximated by the variance of the columns. The correlation terms are neglected for convenience. They are relatively small and therefore will have a minor effect on the results.

If the system is rewritten as a predictor (7.9) there appears to be a constraint on the estimated parameters $\hat{\theta}_1$ and $\hat{\theta}_2$. The denominator may not be equal to zero for the whole range of u . Hence, the constraint is given by:

$$1 - \hat{\theta}_1 u(k) - \hat{\theta}_2 u(k)^2 \neq 0, \quad 0 \leq u \leq 1 \quad (7.13)$$

If the constraint is violated, the solution is rejected.

Two data-sets with measured variables, *i.e.* T_p, T_e, X_e and u , of about 50 days with a sampling interval of 15 minutes were available. The data were obtained from the same location at the same season but for a different period within the season. All parameters are assumed to be constant during the whole season. The parameters are calibrated over one data set (calibration period). The predictive quality is then obtained by using an open loop prediction over the same data set and cross-validated over the second data set (validation period).

The MSE of the predictor (7.7) with parameters estimated by nonlinear least squares (\hat{p}) and the MSE of predictor (7.9) with the parameters estimated by truncated least squares and GTLS ($\hat{\theta}$), with data set 1 the calibration period and data set 2 the validation period, are presented in table 7.2. The predicted and measured temperatures of truncated least squares and GTLS are given in figures 7.3 and 7.4.

The same estimation, validation and cross-validation procedures were performed but now with the data-sets switched, *i.e.* the second data-set was used for estimation of \hat{p} and $\hat{\theta}$. The results are given in table 7.3 and figures 7.5 and 7.6.

Table 7.2: MSE of predictors with parameters estimated by nonlinear LS, truncated LS and GTLS in calibration (data set 1) and validation (data set 2) period.

	nonlinear LS	truncated LS	GTLS
calibration	0.027	0.019	0.023
validation	0.113	0.031	0.053

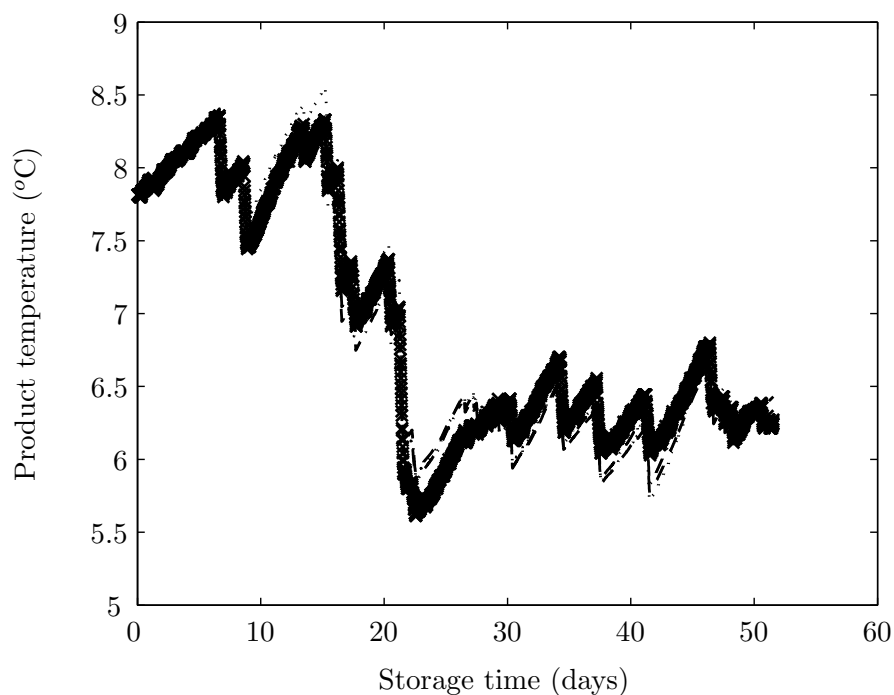


Figure 7.3: Measured (\times) and predicted product temperatures in the calibration period (data set 1) with parameters obtained by nonlinear least squares (\cdots), truncated least squares ($\cdot - \cdot$) and GTLS ($- - -$)

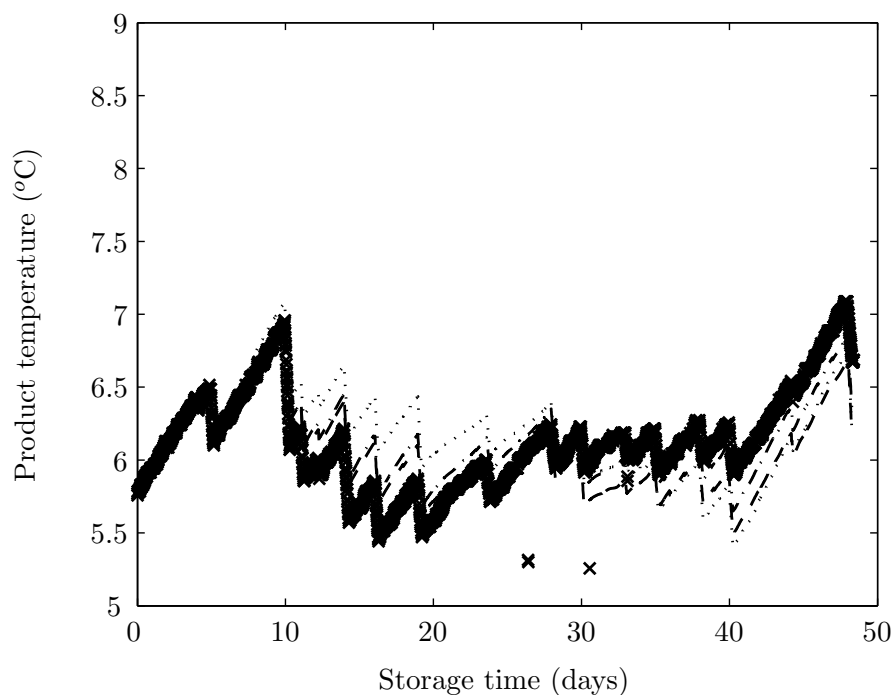


Figure 7.4: Measured (\times) and predicted product temperatures in the validation period (data set 2) with parameters obtained by nonlinear least squares (\cdots), truncated least squares ($\cdot - \cdot$) and GTLS ($- - -$).

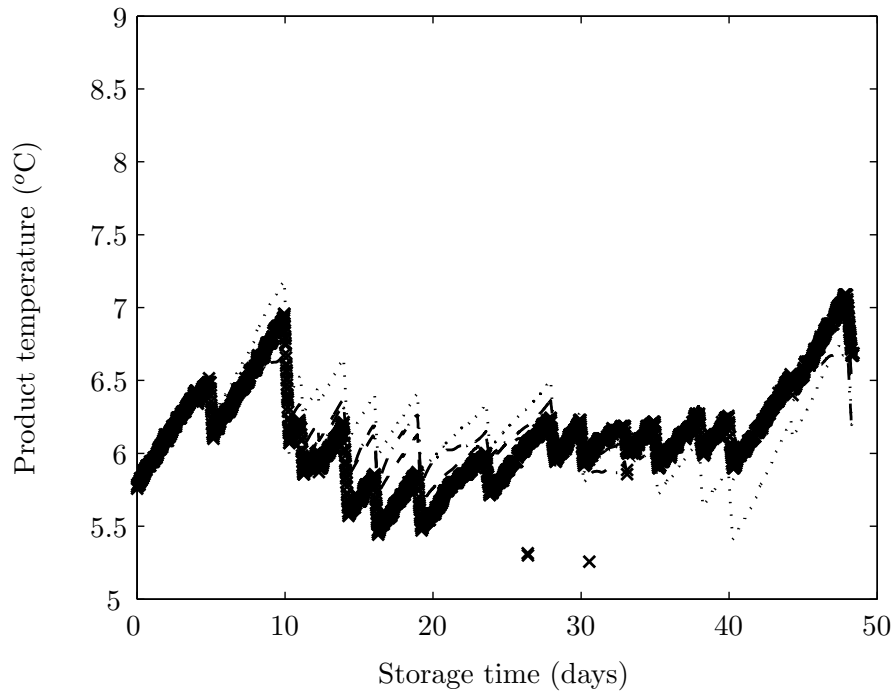


Figure 7.5: Measured (\times) and predicted product temperatures in the calibration period (data set 2) with parameters obtained by nonlinear least squares (\cdots), truncated least squares ($\cdot - \cdot$) and GTLS ($- - -$).

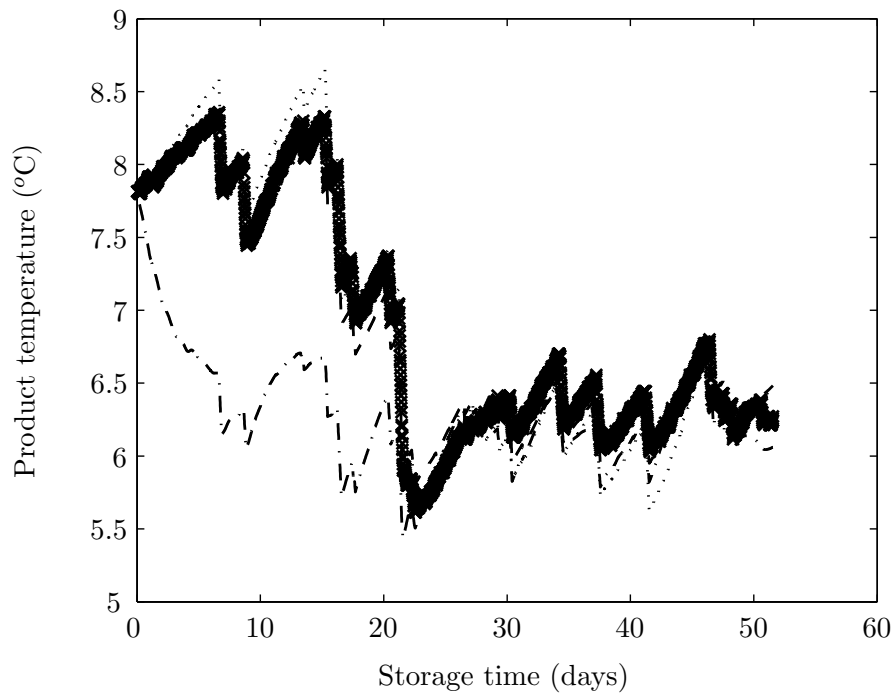


Figure 7.6: Measured (\times) and predicted product temperatures in the validation period (data set 1) with parameters obtained by nonlinear least squares (\cdots), truncated least squares ($\cdot - \cdot$) and GTLS ($- - -$).

Table 7.3: MSE of predictors with parameters estimated by nonlinear LS, truncated LS and GTLS in calibration (data set 2) and validation (data set 1) period.

	nonlinear LS	truncated LS	GTLS
calibration	0.097	0.035	0.017
validation	0.036	0.766	0.016

GTLS has a good performance in each of the four cases and clearly outperforms the nonlinear least squares estimates. The ordinary least squares estimate only has a good performance if the calibration data-set is informative enough. This can be clearly seen in the first part of figure 7.6 where the ordinary least squares estimate has a very poor performance. There appears to be some time variation in figures 7.3 to 7.6, with estimated values above the actual values early on and below actual values late on. This is particularly the case for the nonlinear least squares estimates.

If a nonlinear model is linearly regressive reparameterized (from (7.2) to (7.4)) not only ordinary least squares or total least squares methods become available but a whole set of linear identification tools. As we focus on long-term predictive quality, a prediction error approach may be a good alternative to the presented least squares methods.

7.5 Conclusions

Reparameterization of nonlinear discrete-time models with polynomial quotient structure in Z and p towards linear regression can result in estimates that lead to good predictive quality. In a simulation environment with low noise signals GTLS outperforms ordinary least squares in predictive context. In a real world example it is shown that if the calibration data set is informative, ordinary least squares estimation generates good results in predictive context. For a less informative dataset, the ordinary least squares estimate behaves badly in the region where no information is available. GTLS generates good results in both cases.

Chapter 8

Linear Parameter Estimation of Rational Biokinetic Functions

This chapter is submitted for publication as:
Doeswijk, T.G. and K.J. Keesman. Linear Parameter Estimation of Rational Biokinetic Functions

8.1 Introduction

Rational biokinetic functions appear frequently in biological studies such as metabolic pathway modeling (see *e.g.* [KYY⁺99]). Parameters that appear nonlinearly in a rational mathematical model can be estimated linearly by rearranging and reparameterizing the equations. A well-known example is the Michaelis-Menten model:

$$v = \frac{v_{\max}S}{K_m + S} \quad (8.1)$$

Several types of linearization like the so-called Lineweaver-Burk and Eady-Hofstee linearizations [Dor95] have been proposed in the past and are widely used in practice. It is recognized, however, that by rearranging the model in a linear-in-the-parameters form the error becomes distorted [CB02]. A nonlinear least squares (NLS) approach to the original model avoids this error distortion introduced by linearization and result, in general, in a more accurate and more facile determination of the kinetic constants [GKGE77]. Moreover, the accuracy can be further increased when applying an optimal input design [KS02, SK04]. Proper initial estimates, however, are still required as local minima do occur (see *e.g.* [YAOY03] and figure 6.1). Even with multi-start procedures, a global minimum can not be assured nor an approximation error can be given.

By rearranging a model structure in a linear regressive form, it can be investigated whether the parameters are globally identifiable [LG94]. For the Michaelis-Menten equation (8.1) the parameters were shown to be globally identifiable in chapter 6. For some compartmental Michaelis-Menten models it was shown that after reparameterizing the new parameters become globally identifiable [Sac04]. In general, the original parameters can then be recovered by nonlinear estimation of a static problem, with all the disadvantages of an NLS approach. If the model is used for prediction, however, the new parameters can be readily used in the reparameterized model, as will be demonstrated in this paper.

Generally, the rate v in (8.1) is assumed to be the uncertain variable. As the rate is obtained from measured substrate concentrations, it is reasonable to expect that S at the right-hand side of (8.1) is also uncertain. This assumption makes parameter estimation related to the Michaelis-Menten model an errors-in-variables problem. Errors-in-variables problems that are linear in the parameters can be solved with a total least squares (TLS) method [GVL80]. In chapter 7 based on simulation, this TLS estimator showed an improved performance in predictive context compared to the ordinary least squares (OLS) estimator.

As mentioned before, linear reparameterization usually leads to nonlinearities in the data. These nonlinearities in uncertain data in turn leads to biased estimates [Box71]. Therefore, bias compensation must be introduced. In [Van98] it is shown

that a bias compensated total least squares (CTLS) estimator is asymptotically unbiased for static polynomial functions.

As already mentioned, the Michaelis-Menten model has been linearized frequently in the past to linearly estimate the parameters with an ordinary least squares estimator [Dor95]. To our best knowledge, more advanced methods such as errors-in-variables or bias compensation methods, however, have not been reported. On the other hand, the CTLS was only used for static polynomial functions.

The objective of this paper is to illustrate the wide applicability of linearly reparameterizing rational biokinetic functions and to evaluate the performance of CTLS on these functions. The outline of the paper is as follows. First, a general approach to linearly reparameterize rational functions is proposed. Then, the Michaelis-Menten parameters are estimated linearly with CTLS. Furthermore, to show the wide applicability of model reparameterization and CTLS, the parameters of enzyme kinetics with substrate inhibition are estimated as well. Finally, the practical usability is presented by identifying the biokinetic model parameters of activated sludge from real respirometric data.

8.2 Preliminaries

8.2.1 Linear reparameterization

In chapter 6 it was outlined that for discrete-time rational models a reparameterization procedure can be applied such that the new parameters can be directly and uniquely estimated with ordinary least squares. This procedure is widely used in many different fields when only two parameters need to be estimated. The Lineweaver-Burk linearization is one example of this procedure. In example 6.1 the validity of the linear reparameterization procedure was verified with the Michaelis-Menten model. The solution $\hat{\theta} = \theta$ was obtained in the noise free case whereas the outcome of a nonlinear parameter estimation procedure was not guaranteed. It was concluded that, if no physical insight in terms of bounds is used and the estimation problem is merely treated as a fitting problem, one could easily end up in a local minimum. But also with knowledge about system or parameter bounds care must be taken and the solution due to possible singularities in (8.1) is not always obvious.

8.2.2 Handling data errors

The traditional linearizations of the Michaelis-Menten model have been criticized many times [GKGE77, CB02] because they distort the error. First, the distribution of the error is changed. Second, the relationship between the dependent and independent variables is changed. To see this more clearly, let us add an error term to

the basic Michaelis-Menten model (8.1):

$$v = \frac{v_{\max}S}{K_m + S} + e \quad (8.2)$$

Linearizing according to the Lineweaver-Burk plot leads then to:

$$\frac{1}{v} = \frac{K_m + S}{(v_{\max} + e)S + K_me} \quad (8.3)$$

Furthermore, it is also noted in [GKGE77] that the substrate concentrations are not free of error but they are generally believed to be much smaller than those of the velocities. Ignoring this last assumption, the Michaelis-Menten model becomes an *errors-in-variables* problem. To illustrate this, the Michaelis-Menten model (8.1) is used to describe substrate consumption in a batch reactor. Since, in practice, we most often have sampled substrate data, the substrate conversion rate is replaced by $v = \frac{S(k+1)-S(k)}{\Delta t}$. The discrete-time substrate consumption model can then be written as:

$$S(k+1) = S(k) - V_{\max} \frac{S(k)}{K_m + S(k)} \quad (8.4)$$

with $V_{\max} = v_{\max}\Delta t$. The variables $S(k+1)$ and $S(k)$ are both measured variables and as such subject to errors. Hence, the model clearly is errors-in-variables. A method to estimate parameters in such problems is total least squares [GVL80]. More details and algorithms can be found in [VHV91].

Now, let us proceed by linearly reparameterizing the substrate model (8.4) by multiplying both sides with the denominator and rearrange such that the model becomes linear in its parameters:

$$S(k)(S(k+1) - S(k)) = [-(S(k+1) - S(k)) \quad -S(k)] \begin{bmatrix} K_m \\ V_{\max} \end{bmatrix} \quad (8.5)$$

If the substrate concentrations are known, the parameters can be estimated with ordinary least-squares. Because the problem is errors-in-variables, however, the ordinary least squares solution is biased [Nor86]. Because the data are not independent (*e.g.* $S(k)$ appears in all three columns) an extension of the total least squares method is needed. The dependency of the columns can be expressed in a covariance matrix. If the covariance matrix of the errors in the data is known, generalized total least squares (GTLS) [VHV89] gives the solution.

The GTLS solution can be calculated easily if the regression variables are linear and the covariance matrix of the variables is known. Due to the linear reparameterization, however, this will rarely be the case. Not only the construction of the covariance matrix is complicated but also bias is introduced because uncertain variables appear nonlinearly in the regression matrix [Box71]. A simple example will

illustrate this:

$$x = x_0 + e \quad e \sim \mathcal{N}(0, \sigma^2) \quad (8.6)$$

$$E[x] = E[x_0 + e] = x_0 \quad (8.7)$$

$$E[x^2] = E[x_0^2 + 2x_0e + e^2] = x_0^2 + \sigma^2 \quad (8.8)$$

where x is a measured variable, x_0 the true but unknown variable, and e the measurement error. From this example it can be seen that the expected value of x^2 is biased, *i.e.* $b := E[x^2] - x_0^2 = \sigma^2$. Now, if parameter a in the model $y = ax^2$, with y and x uncertain variables, is to be estimated it will be biased if no bias compensation method is used. In [Van98] a bias compensated total least squares (CTLS) method is proposed. It is based on the GTLS approach and provides algorithms to calculate a bias compensated data matrix, to construct the covariance matrix and add bias compensation to this covariance matrix. A summary of the algorithm to calculate the matrices with an illustrative example can be found in appendix D.

8.3 Methods

In the following, three separate cases are investigated: estimation of parameters from Michaelis-Menten kinetics from simulated data; estimation of parameters from enzyme kinetics with substrate inhibition from simulated data; and respirometric data from an activated sludge experiment are used to estimate Monod parameters. The traditional parameter estimation methods based on OLS and NLS were compared with CTLS.

8.3.1 Michaelis-Menten kinetics

In this simulation experiment, the substrate concentration S^0 was simulated using (8.4) with nominal parameter values, $K_m^0 = 10$, $V_{\max}^0 = 5$ and the initial substrate concentration $S(0) = 50$, with $k = 1 \dots 15$. In addition, noise was added to the simulated substrate concentrations according to:

$$S = S^0 + e \quad e \sim \mathcal{N}(0, \sigma^2) \quad (8.9)$$

The parameters were then estimated from the noise corrupted data-set.

First, the parameters were estimated with NLS. Assuming that no prior information about the parameters was present, the initial parameter estimates were chosen as: $\hat{K}_m^i = 0$ and $\hat{V}_{\max}^i = 0$. Second, the traditional Lineweaver-Burk (L-B) transformation was used to estimate the parameters with OLS. Finally, CTLS was used with data matrix Φ and covariance matrix C , that is:

$$\Phi = \begin{bmatrix} S(k) - S(k+1) & -S(k) & S(k)(S(k+1) - S(k)) + \sigma^2 \end{bmatrix} \quad (8.10)$$

$$C = \frac{1}{m} \sum_{k=1}^m \sigma^2 \begin{bmatrix} 2 & -1 & S(k+1) - 3S(k) \\ -1 & 1 & 2S(k) - S(k+1) \\ S(k+1) - 3S(k) & 2S(k) - S(k+1) & 5S(k)^2 + S(k+1)^2 - 4S(k+1)S(k) - 3\sigma^2 \end{bmatrix} \quad (8.11)$$

Here it is assumed that the noise between $S(k+1)$ and $S(k)$ are independent and identically distributed as given in (8.9). The estimates were then used to predict the substrate concentration

8.3.2 Substrate inhibition

The discrete-time model of substrate inhibition of enzyme kinetics in a batch reactor is given by:

$$S(k+1) = S(k) - \frac{V_{\max} S(k)}{K_m + S(k) + \frac{S_{k-1}^2}{K_i}} \quad (8.12)$$

Normally, the parameters of this model are estimated with NLS. Nonetheless, methods have been presented that graphically analyzes these kinetics [WAO⁺99]. This procedure is based on a two-step procedure. First, the value of $\sqrt{\frac{K_m}{K_i}}$ was determined by visual inspection of a plot of the rate versus the substrate concentration. Hereafter, the remaining parameters were estimated by a linearized equation. Linear reparameterization according to theorem 6.1, on the other hand, is then a natural extension that provides an estimate in a single step. Linearly reparameterizing the model (8.12) leads to:

$$S(k)(S(k+1) - S(k)) = (S(k) - S(k+1))K_m + K_i' S(k)^2 (S(k) - S(k+1)) - V_{\max} S(k) \quad (8.13)$$

with $K_i' = \frac{1}{K_i}$.

In this experiment the substrate concentration is simulated with eqn. (8.12) and $K_m = 2$, $K_i = 5$, $V_{\max} = 1$ and $S(0) = 10$ with $k = 1 \dots 25$. In addition, gaussian distributed noise was generated and added to the simulated substrate concentration (S^0) similarly as in (8.9). Next, the parameters were estimated using the noisy dataset with NLS, OLS and CTLS. The bias compensated data matrix Φ and covariance matrix C for this case are given by (8.14) and (8.15) where, $u = S(k+1)$ and $v = S(k)$. Finally, The substrate concentration was predicted using the estimated parameters. The mean square error of the predicted substrate concentration minus the actual simulated concentration was calculated.

8.3.3 Respirometric data

The practical identifiability of the Monod parameters was tested with real respirometric data from an activated sludge process [KSv98]. The respirometric data are

$$\Phi = \begin{bmatrix} S(k) - S(k+1) & S(k)^2(S(k+1) - S(k)) + (S(k+1) - 3S(k))\sigma^2 & -S(k) & S(k)(S(k+1) - S(k)) + \sigma^2 \end{bmatrix} \quad (8.14)$$

$$C = \frac{1}{m} \sum_{k=1}^m \sigma^2 \begin{bmatrix} 2 & 4v^2 - 2uv - 4\sigma^2 & -1 & u - 3v \\ 4v^2 - 2uv - 4\sigma^2 & 10v^4 - 12v^3u - 4u^2v^2 - \sigma^2(-4v^2 - 24uv - 2u^2) - 24\sigma^4 & 2uv - 3v^2 - 3\sigma^2 & -7v^3 - 7uv^2 - 2u^2v - \sigma^2(15v - 5u) \\ -1 & 2uv - 3v^2 - 3\sigma^2 & 1 & -u - 2v \\ u - 3v & -7v^3 - 7uv^2 - 2u^2v - \sigma^2(15v - 5u) & -u - 2v & 5v^2 + u^2 - 4uv - 3\sigma^2 \end{bmatrix} \quad (8.15)$$

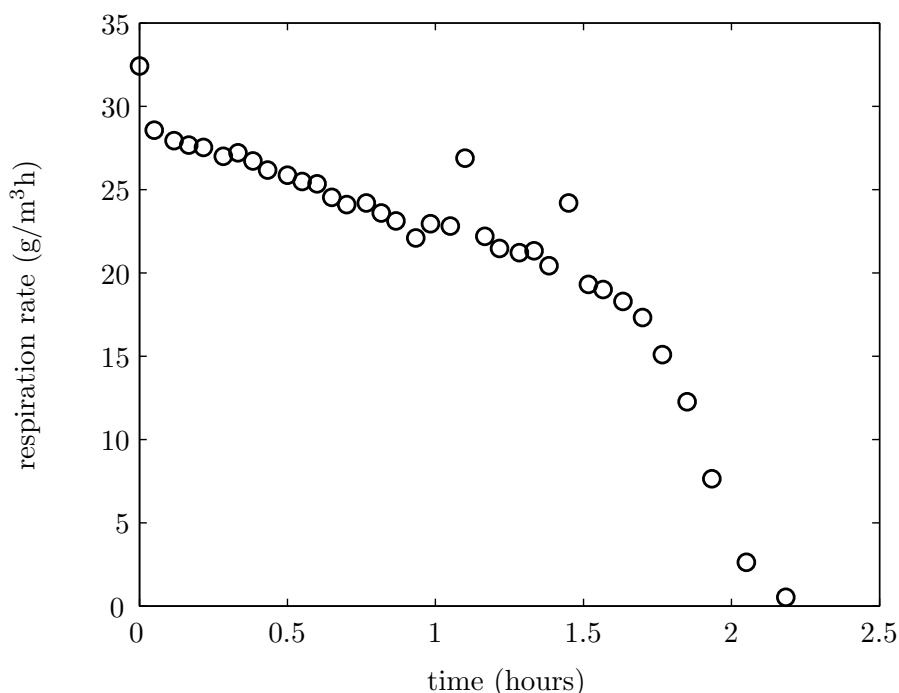


Figure 8.1: Measured respiration rates in activated sludge

presented in figure 8.1. In this figure, the endogenous respiration was subtracted from the total respiration. The data shown are also known as the exogenous respiration rates in activated sludge experiments. The model of the fast substrate response is given by:

$$r = \frac{1 - Y}{Y} \mu_m \frac{S}{K_s + S} X_{BH,0} \quad (8.16)$$

where r the respiration rate ($gm^{-3}h^{-1}$), Y the yield (gg^{-1}), μ_m the maximum specific growth rate (h^{-1}), K_s the Monod half-saturation constant (gm^{-3}), S the substrate concentration (gm^{-3}) and $X_{BH,0}$ the initial biomass concentration (gm^{-3}) which is assumed to be constant in this experiment. Furthermore, respiration rate and substrate consumption are related by:

$$(1 - Y) \frac{dS}{dt} = -r \quad (8.17)$$

If only respiration rates are available, structural identifiability is a problem in activated sludge processes [DVV95]. For example, $S(0)$ and Y cannot be estimated separately; only the combination $(1 - Y)S(0)$ can be estimated. Now, let us therefore rewrite the system with $\tilde{S} := (1 - Y)S$, so that

$$\frac{d\tilde{S}}{dt} = -\frac{(1 - Y)}{Y} \mu_m \frac{\tilde{S}}{(1 - Y)K_s + \tilde{S}} X_{BH,0} \quad (8.18)$$

If $X_{BH,0}$ is assumed to be constant, which holds for sufficiently small $\frac{S(0)}{X_{BH,0}}$ -ratios [CCC92], the model (8.18) can be further reduced by substituting $\tilde{\mu}_m = \frac{(1 - Y)}{Y} \mu_m X_{BH,0}$

and $\tilde{K}_s = (1 - Y)K_s$:

$$\frac{d\tilde{S}}{dt} = -\frac{\tilde{\mu}_m \tilde{S}}{\tilde{K}_s + \tilde{S}} \quad (8.19)$$

This model is analogous to the Michaelis-Menten model (8.1). Using (8.17), the modified substrate concentrations \tilde{S} can be estimated from the integral of the respiration rates. Under non-equidistant sampling the initial substrate concentration is determined by:

$$S(0) = \sum_{i=0}^{m-1} r_i(t(i+1) - t(i)) \quad (8.20)$$

with m the number of measurements. Given the initial substrate concentration and the respiration rates and discretizing the model (8.19) the substrate concentrations can be calculated by:

$$\tilde{S}(k+1) = \tilde{S}(k) - \tilde{\mu}_m \frac{\tilde{S}(k)(t_k - t_{k-1})}{\tilde{K}_s + \tilde{S}(k)} \quad (8.21)$$

If $\tilde{S}(k+1)$ is independent and identical distributed for all k then the data matrix Φ and the covariance matrix C are defined by (8.10)-(8.11) and all data are present to estimate the parameter vector $[\tilde{K}_s \quad \tilde{\mu}_m]^T$ with CTLS. Because the data are obtained by integrating the respiration rates, however, it cannot be assumed that the calculated substrate concentrations are independent even if the errors on the respiration rates were. This is illustrated by elaborating the variances of $\tilde{S}(0)$ and $\tilde{S}(1)$ and assuming an additive measurement error: $y(k) = r(k) + e(k)$ with e normally distributed with zero mean. In addition, it is assumed that $\text{cov}(\tilde{S}, e) = 0$. First, some variances are calculated.

$$\begin{aligned} \text{var}(\tilde{S}(0)) &= \sigma^2 \\ \text{var}(\tilde{S}(1)) &= E \left[\tilde{S}(1)^2 \right] - E \left[\tilde{S}(1) \right]^2 \\ &= E \left[(\tilde{S}(0) - (y(0) - e(0))\Delta t(0))^2 \right] - E \left[(\tilde{S}(0) - (y(0) - e(0))\Delta t(0)) \right]^2 \\ &= \sigma^2 + \Delta t(0)^2 E(e(0)^2) \end{aligned}$$

If $\Delta t(0)^2 E(e(0)^2) \ll \sigma^2$, which will generally be the case because Δt is small, then the variance does not change in time. Next, the covariance between $\tilde{S}(0)$ and $\tilde{S}(1)$ is calculated.

$$\begin{aligned} \text{cov}(\tilde{S}(0), \tilde{S}(1)) &= E \left[\tilde{S}(0)\tilde{S}(1) \right] - E \left[\tilde{S}(0) \right] E \left[\tilde{S}(1) \right] \\ &= E \left[\tilde{S}(0)(\tilde{S}(0) - (y(0) - e(0))\Delta t(0)) \right] - E \left[\tilde{S}(0) \right] E \left[\tilde{S}(0) - (y(0) - e(0))\Delta t(0) \right] \\ &= \sigma^2 \end{aligned}$$

These results can be generalized to:

$$\text{var}(\tilde{S}(k+1)) \approx \text{var}(\tilde{S}(k)) \quad (8.22)$$

$$\text{cov}(\tilde{S}(k+1), \tilde{S}(k)) = \text{var}(\tilde{S}(k)) \quad (8.23)$$

Hence, if $\text{var}(\tilde{S}(k)) = \sigma^2 \forall k$ then $\tilde{S}(k+1)$ and $\tilde{S}(k)$ are fully correlated for all k .

The data matrix as well as the covariance matrix to be used in the CTLS estimation can now be calculated. The data matrix is given by:

$$\Phi = \begin{bmatrix} \tilde{S}(k) - S(k+1) & -\tilde{S}(k) & \tilde{S}(k)(\tilde{S}(k+1) - \tilde{S}(k)) \end{bmatrix} \quad (8.24)$$

Recall that the covariance matrix is defined by (8.11) if $\tilde{S}(k+1)$ and $\tilde{S}(k)$ are independent. If they are not independent the covariance matrix is given by:

$$C = \frac{1}{m} \sum_{k=1}^m \sigma^2 \begin{bmatrix} (1-\rho) & 0 & (\rho-1)\tilde{S}(k) \\ 0 & \rho\Delta t^2 & \rho\Delta t (\tilde{S}(k) - \tilde{S}(k+1)) \\ (\rho-1)\tilde{S}(k) & \rho\Delta t (S(k) - \tilde{S}(k+1)) & c_{33} \end{bmatrix} \quad (8.25)$$

with $c_{33} = -2(1-\rho)\tilde{S}(k+1)\tilde{S}(k) + \tilde{S}(k)^2 + (1-\rho)\tilde{S}(k+1)^2 - \rho(5\rho-1)\sigma^2$ and

$\rho = \frac{\text{cov}(\tilde{S}(k+1), \tilde{S}(k))}{\sqrt{\text{var}(\tilde{S}(k+1))\text{var}(\tilde{S}(k))}} = \frac{\sigma_{\tilde{S}(k)}^2}{\sigma_{\tilde{S}(k+1)}\sigma_{\tilde{S}(k)}} = \frac{\sigma_{\tilde{S}(k)}}{\sigma_{\tilde{S}(k+1)}}$. With the data and covariance matrix given, the parameter vector $[\tilde{K}_s \quad \tilde{\mu}_m]^T$ can be calculated with CTLS.

8.4 Results and discussion

In figure 8.2 a typical result for a specific noise realization of the predictor (8.1) is presented with parameter estimates found by the several fitting procedures. It can be seen that NLS has the best predictor performance in terms of mean squared error (MSE). The peak at noise level 1 indicates that the procedure ended up in a *local* minimum. Furthermore, it is clear that CTLS outperforms the (unweighed) Lineweaver-Burk reparameterization and has an MSE that is less than the variance of the noise.

As the estimates depend on the noise sequence, the noise was generated 1000 times. The mean and standard deviation of the estimates from NLS and CTLS are given in table 8.1. It should be noted that the NLS approach, for each of the estimations, used the initial parameter vector $[K_m \quad V_{max}]^T = [0 \quad 0]^T$. The results show that in this case the NLS approach usually generates the best estimates. Furthermore, it can be seen that CTLS generates estimates that are well suited. The estimates are, however, widely spread at high noise levels. This indicates that for several noise sequences the estimates are far out of range.

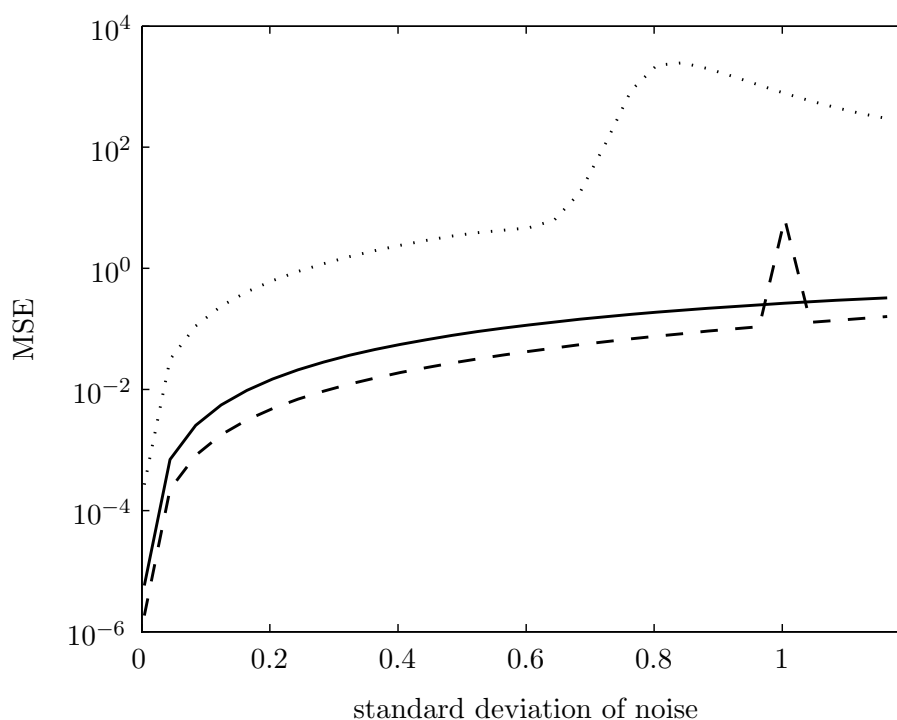


Figure 8.2: MSE of simulated output with nominal parameters ($K_m = 10$ and $V_m = 5$) minus simulated output with estimated parameters with NLS, CTLS and L-B at several noise levels.

The parameters of the substrate inhibition model (8.12) are very hard to identify. The difficulties in the practical identifiability of the comparable Andrews kinetics based on sensitivity equations were shown by [SKS03]. In a noise-free simulation study the true parameters cannot be obtained with NLS. In figure 8.3 the many local minima and long valleys can be clearly seen. By nature, under noise-free conditions and linear reparameterization, OLS finds the exact solution. This also holds for CTLS.

In figure 8.4 a typical realization of the predictive performance of the model (8.12) with parameters estimated with different estimators is presented. It can be clearly seen, as expected from figure 8.3, that NLS does not find the global minimum in general. The CTLS estimator on the other hand finds the global minimum. With low noise levels this results in quite accurate estimates. With increasing noise levels

Table 8.1: Mean \pm standard deviation of parameter estimates at three noise levels with $K_m^0 = 10$, $V_{max}^0 = 5$

	$\sigma = 0.01$		$\sigma = 0.1$		$\sigma = 1$	
	NLS	CTLS	NLS	CTLS	NLS	CTLS
\hat{K}_m	10.00 \pm 0.022	10.00 \pm 0.059	10.00 \pm 0.22	9.97 \pm 0.60	9.89 \pm 2.86	10.92 \pm 114.97
\hat{V}_{max}	5.00 \pm 0.003	4.99 \pm 0.091	5.00 \pm 0.029	4.99 \pm 0.091	4.98 \pm 0.39	5.13 \pm 13.09

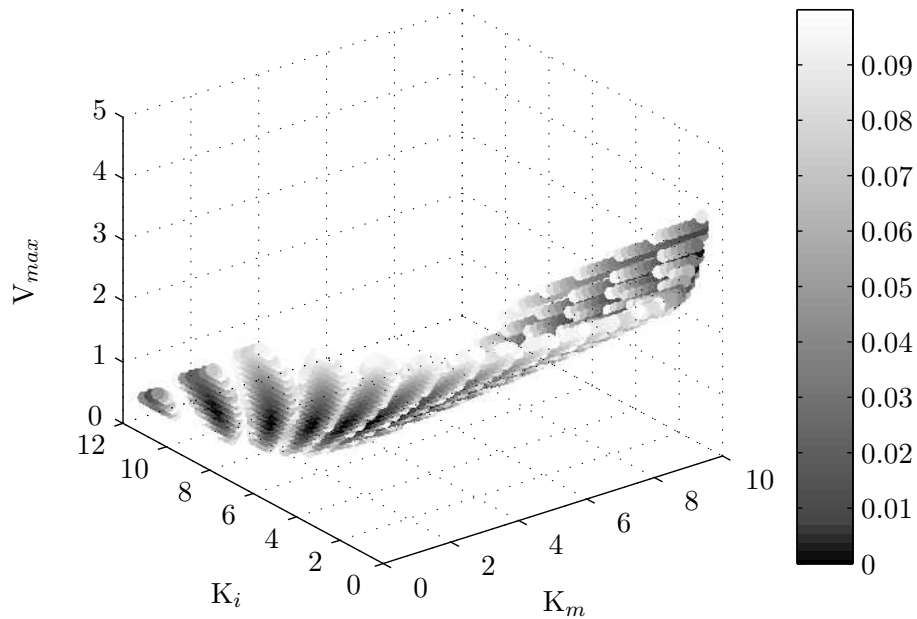


Figure 8.3: MSE of simulated output with nominal parameters ($K_m = 2$, $K_i = 5$ and $V_m = 1$) minus simulated output with estimated parameters in the 3-dimensional parameter space (black = low MSE, white = high MSE)

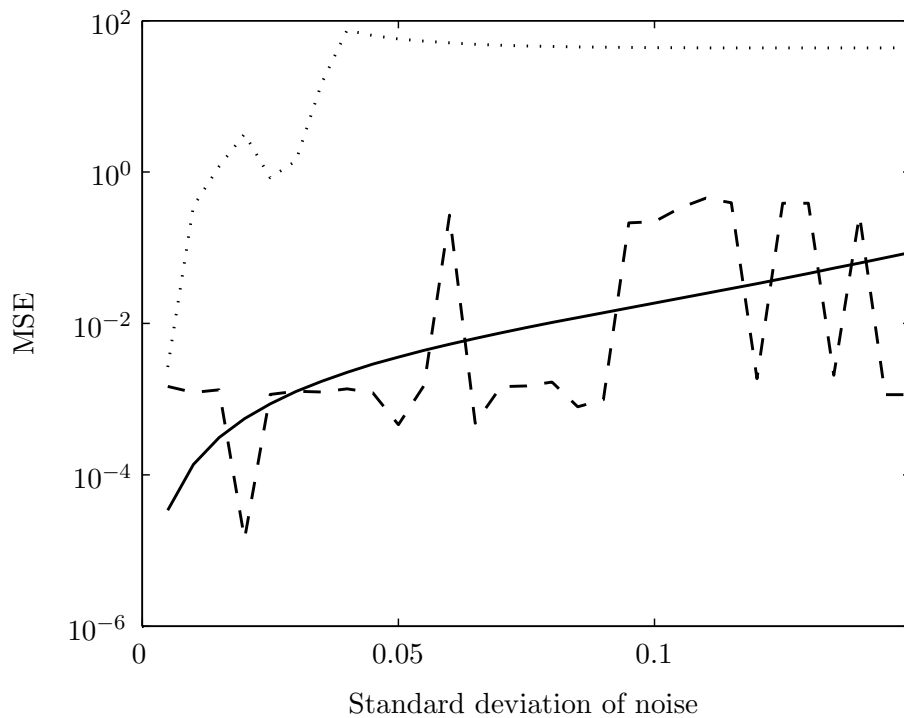


Figure 8.4: MSE of simulated output with nominal parameters ($K_m = 2$, $K_i = 5$ and $V_m = 1$) minus simulated output with parameters estimated with NLS, CTLS and OLS at several noise levels.

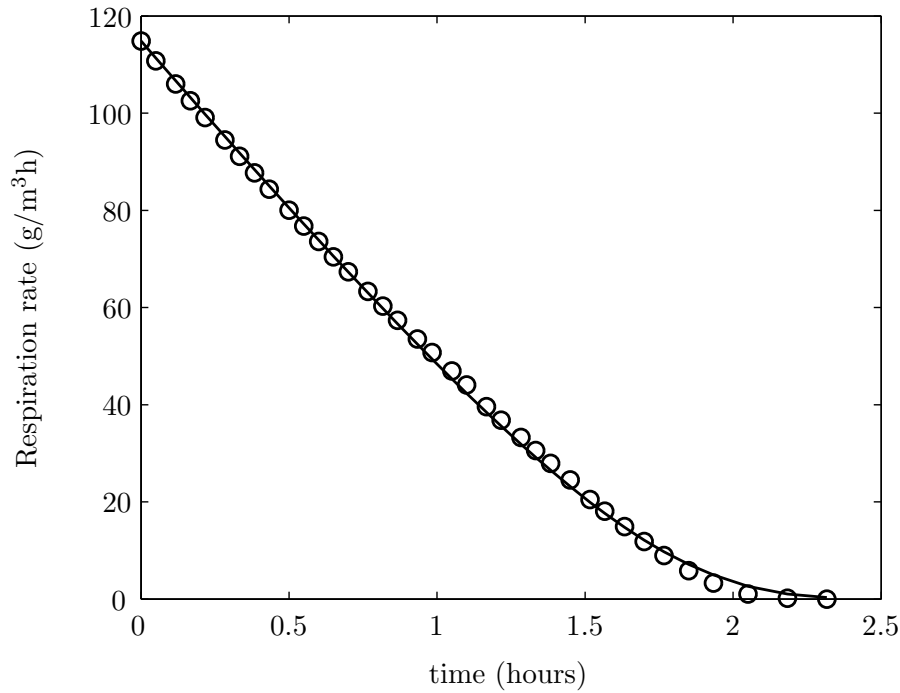


Figure 8.5: Predicted vs. measured substrate concentration in waste water

the accuracy decreases but up to the noise levels presented, the estimates are frequently more reliable than the estimates of the NLS estimator. The OLS estimator shows to be reliable only for very low noise levels. This estimator is therefore not recommended.

In figure 8.5 the predicted substrate concentrations from the activated sludge experiment with the estimates obtained with CTLS are presented. It is clear from this figure that the predicted values are close to reconstructed substrate concentrations from the measured respirometric data.

It is also possible to estimate the initial substrate concentration together with the kinetic parameters using linear regression. Therefore, the differential equation:

$$\frac{d\tilde{S}}{dt} = -\frac{\tilde{\mu}_m \tilde{S}}{\tilde{K}_s + \tilde{S}} \quad (8.26)$$


must be integrated. This method is also known as the integrated rate equation (see *e.g.* [Sch69]). After integration and linear reparameterization the following linear regressive form can be obtained:

$$\tilde{S} = \begin{bmatrix} 1 & -t & \ln \tilde{S} \end{bmatrix} \begin{bmatrix} \tilde{S}(0) + \tilde{K}_s \ln \tilde{S}(0) \\ \tilde{\mu}_m \\ \tilde{K}_s \end{bmatrix} \quad (8.27)$$

Given (8.27), the bias compensated data and covariance matrices can be derived and CTLS can be applied. Note that the first two elements (1 and $-t$) of (8.27) are free of error.

Just as the traditional linearization approaches, *i.e.* Lineweaver-Burk, Eady-Hofstee and Langmuir linearizations, the general linear reparameterization method given by theorem 6.1 does not yield a unique reparameterization. Further research is required to obtain the most suitable reparameterization for each rational function that is to be estimated linearly (see also chapter 7).

The estimation results depend on the chosen parameters, initial conditions and the number of measurements used in the simulation. The results, however, are generally the same: a global optimum in the noise-free case for CTLS and OLS; less biased estimates for CTLS than OLS with increasing noise levels; NLS might end up in a local minimum.

Although, calculation of the data and covariance matrices for CTLS are somewhat involved, Matlab software is available from the authors for calculating these matrices. Gaussian distributed noise sequences are assumed in these calculations.

8.5 Conclusions

It has been illustrated in this paper that for a simple rational biokinetic function like the Michaelis-Menten equation the NLS estimator gives good results in general. It is possible, however, that a local minimum is found. Traditional linearization methods ensures a global minimum but become biased if noise is present. CTLS also ensures a global minimum and provides less biased estimates than the traditional linearization methods with OLS. For somewhat more complex rational functions that entail many local minima like enzyme kinetics with substrate inhibition, CTLS provides estimates that are closer to the nominal values than NLS for low noise levels. At higher noise levels, NLS might end up in a local minimum (see figure 8.3) but performs better in predictive context. Hence, no physical interpretation should be given to the estimates found with NLS. Additionally, CTLS showed to be practically usable in an example with respirometric data from a waste water treatment plant. Finally, it is concluded that CTLS is a powerful alternative to the existing least-squares methods for biokinetic models.

Chapter 9

Conclusions and general discussion

In this chapter conclusions are drawn of the main research questions postulated in section 1.5. In addition to the conclusions some thoughts are discussed for possible further research on each of the subjects.

9.1 Weather forecasts

From the results found in chapters 2 and 3 it is clear that local measurements are crucial for local short term forecasting. It was already shown by [TVWVS96] that for a forecast up to one hour ahead the local lazy man weather prediction, *i.e.* the weather for the next hour remains constant, outperformed regional commercial weather forecasts updated twice a day. Despite the improvements made in weather forecasting it is not expected that the very short term regional commercial weather forecasts become more reliable than the local lazy man weather prediction. This is due to the fact that (commercial) weather forecasts are based on measurements that are obtained at least several hours earlier. These observations are then used in atmospheric models to predict the future weather. Model deficiencies result in prediction errors, hence, even the zero hour ahead forecast contains errors. Furthermore, site specific information is generally not present in atmospheric models but can effect the local climate *e.g.* [Mak06]. Hence, research question 1 is answered positively, *i.e.*, data assimilation with local measurements results in better local forecasts provided that the local measurements are of good quality.

Generally, in data-assimilation a trade-off is made between predicted model states and measurements. Agricultural control systems are directly influenced by the weather. For instance, cooling of a storage facility can only be done if the ambient air temperature is low enough [GNF03]. In this case well calibrated measurement devices are important for proper control. Hence, if these devices are used for direct control they can also be used for updating local weather forecasts. Reversely, the deviation between measurements and weather forecasts can also be used to alert for an anomaly in the measurement device.

As many agricultural sites use weather observations and/or weather forecasts for management purposes, the electronic measurements of each of the agricultural sites can be integrated into a network of sensors. Hence, instead of local updates this network can be used to adjust the weather forecast for the local sites.

9.2 Uncertainty and optimal control

In chapter 4 the relation between uncertainty in weather forecasts and the uncertainty of the predicted potato temperature in a storage facility was investigated. Minimum variance control was not suitable because the disturbance entering the system (*i.e.*

weather forecast uncertainty) was related to the control variable (ventilation rate). Furthermore, because the main objective was to use weather forecasts to minimize energy costs, open loop optimal control based on a predefined cost function was used as a general tool for calculating the future controls. However, if uncertainty is present the future controls as computed in the open loop calculations are generally not optimal anymore but only a best guess. The consequences of a deviation in the system states on the cost function should therefore be analyzed. Furthermore, the possibility of constraint violation must be verified.

It has been shown that given the controls a linear time-varying system arises. Furthermore, uncertainty is only entering the system if it is ventilated. The disturbance propagating through the system can be calculated in order to obtain the uncertainty of the system states. The 2σ -band provides a 95% confidence region if the disturbances are normally distributed. This confidence region can then be used for risk assessment or it can be integrated in the objective function to calculate the optimal control trajectory. This answers the research question 2. The answer to research question 3 specified to the storage facility is given by: because uncertainty is entering the system if the storage facility is ventilated with ambient air, the control input (ventilation) generally shifts towards parts where the weather forecast contains less uncertainty. This approach is especially interesting if the uncertainty of the disturbances is time-varying as in ensemble weather forecasts.

If the disturbances are not normally distributed the variance can still be calculated and used in the optimization procedure. If the distributions are given, such as in ensemble weather forecasts, the maximum (\bar{d}) and minimum (\underline{d}) values at each time instant can be calculated. This then results in an unknown-but-bounded approach [Sch73]. The effect of these two extreme outer bounding profiles are then evaluated in the cost function. If state constraints are present both extreme profiles need to be evaluated against these constraints. Alternatively, both profiles are part of the cost function and a trade-off will be made. As an example an unweighted cost function in addition to section 4.4 can be formulated as

$$J = \sum_{k=0}^H [(\bar{x}(k) - x_{ref})^2 + (\underline{x}(k) - x_{ref})^2 + u(k)^2] \quad (9.1)$$

where $\bar{x} = f(\bar{d})$ the maximum possible temperature and $\underline{x} = f(\underline{d})$ the minimum possible temperature. This approach only holds for linear systems because only then \bar{d} leads to the maximum possible temperature.

The weather forecast ensembles generated by atmospheric models are generally not normally distributed. A numerical approach by predicting the system states for each of the ensemble members can be used. The output ensemble then contains an uncertainty measure and this uncertainty measure can be used in the the cost

function. This leads to a less conservative control than the unknown-but-bounded approach but is computationally demanding for an increasing number of ensemble members. An advantage over the unknown but bounded approach is that it can also be used for nonlinear systems.

For a more robust performance of the controller feedback is applied. In chapter 5 the effect of the receding horizon optimal controller on the realized costs was investigated. It has been shown that the uncertainty in the weather forecast ensemble has only a minor effect on the total costs if the optimal controls are recalculated every 24 hours or less. Furthermore, for the arbitrary chosen range of parameter uncertainty the effect on the total costs quite low if the optimal controls are recalculated every 6 hours. This answers research question 4. This does not imply, however, that no weather forecasts are needed. In the papers [KPL03] and [LvdVdKC06] the effect of using weather forecasts in an optimal control approach showed an improvement by reducing costs if the forecast horizon increases. Furthermore, using the weather of "today" as a forecast for the coming days resulted in much less cost decrease than using the commercial weather forecast.

9.3 Parameter Estimation

9.3.1 Linear reparametrization

In part II calibration of a mathematical model which is an important procedure to optimize the predictive quality of the model is the main subject of interest. Usually a nonlinear least squares approach is used to solve this issue. Whereas the main weakness of nonlinear least squares algorithms is that they can easily end up in local minima and have a high computational demand, their strength is the general applicability. The main strength of OLS and TLS is that they give directly the global minimum. Unfortunately, these methods are only applicable to systems with outputs that are linear in the parameters. Therefore, if the system is reparameterized such that its output becomes linear in its new parameters, OLS and TLS can be used and the parameters are globally identifiable [LG94]. In the noise-free case unbiased parameter estimates are obtained with ordinary least squares. If, however, noise is present difficulties arise. Additional methods like GTLS and CTLS proved to be more robust in these cases.

Rational systems or systems with a polynomial quotient structure as used in this study are defined by numerator and denominator polynomials in inputs, outputs and parameters. If a rational system is not globally identifiable, the system can be reparameterized such that the newly defined system is linear in its parameters and linear estimation methods can be applied to estimate the parameters. The system

structure is preserved in this case. The new parameters, however, might become correlated.

In a recent study [Zhu05a], also based on putting rational model parameters in a linear regressive form, an iterative solution was presented to cope with noise and bias compensation. The main advantage of GTLS and CTLS, as applied in this study, is that *no* iterative procedures are needed and a global minimum is found directly.

In answering research questions 5 and 6 the following can be concluded: (i) rational models can be linearly reparameterized such that the newly obtained parameters can be estimated with linear least squares but does not necessarily lead to a good predictor, (ii) estimation based on total least squares techniques generally outperforms ordinary least squares for linearly reparameterized models.

After linear reparametrization and estimating the newly defined linear parameters the original parameters can be recovered by backtransformation. This backtransformation is a *static* nonlinear estimation problem. As an example, the storage system with parameter vector p from eqn. (6.14) is given. After reparameterization the new system (6.15) with the parameter vector θ is obtained. The parameters are related as follows:

$$\begin{aligned}
 \theta_1 &= \frac{\tilde{p}_7}{\tilde{p}_6} + \frac{\tilde{p}_4}{\tilde{p}_3} \\
 \theta_2 &= \frac{\tilde{p}_4\tilde{p}_7}{\tilde{p}_3\tilde{p}_6} \\
 \theta_3 &= \tilde{p}_1 + \frac{\tilde{p}_2}{\tilde{p}_3} + \frac{\tilde{p}_5}{\tilde{p}_6} \\
 \theta_4 &= \frac{\tilde{p}_1\tilde{p}_3\tilde{p}_7 + \tilde{p}_1\tilde{p}_4\tilde{p}_6 + \tilde{p}_2\tilde{p}_7 + \tilde{p}_4\tilde{p}_5}{\tilde{p}_3\tilde{p}_6} \\
 \theta_5 &= \frac{\tilde{p}_1\tilde{p}_4\tilde{p}_7}{\tilde{p}_3\tilde{p}_6} \\
 \theta_6 &= \frac{\tilde{p}_8}{\tilde{p}_3} \\
 \theta_7 &= \frac{\tilde{p}_6\tilde{p}_9 + \tilde{p}_7\tilde{p}_8}{\tilde{p}_3\tilde{p}_6} \\
 \theta_8 &= \frac{\tilde{p}_7\tilde{p}_9}{\tilde{p}_3\tilde{p}_6} \\
 \theta_9 &= \frac{\tilde{p}_{10}}{\tilde{p}_6} \\
 \theta_{10} &= \frac{\tilde{p}_3\tilde{p}_{11} + \tilde{p}_4\tilde{p}_{10}}{\tilde{p}_3\tilde{p}_6} \\
 \theta_{11} &= \frac{\tilde{p}_4\tilde{p}_{11}}{\tilde{p}_3\tilde{p}_6} \\
 \theta_{12} &= \tilde{p}_{12} + \frac{\tilde{p}_{13}}{\tilde{p}_3}
 \end{aligned}$$

$$\theta_{13} = \frac{\tilde{p}_{12}\tilde{p}_4\tilde{p}_6 + \tilde{p}_{12}\tilde{p}_3\tilde{p}_7 + \tilde{p}_{13}\tilde{p}_4}{\tilde{p}_3\tilde{p}_6}$$

$$\theta_{14} = \frac{\tilde{p}_{12}\tilde{p}_4\tilde{p}_7}{\tilde{p}_3\tilde{p}_6}$$

Clearly, this is a set of rational functions and the parameter vector p can be estimated with a nonlinear least squares approach if θ is given. An advantage over the original nonlinear least squares estimation problem is that it now is a static problem. Because the main purpose in this thesis was to obtain a good predictor this backtransformation was not needed.

In chapter 7 the covariance matrix of the storage system used in the GTLS algorithm was given as a diagonal matrix with the variance of the columns of the data matrix as the diagonal entries. This resulted in an excellent performance on both the calibration data and the validation data set. To provide a general approach of tuning the covariance matrix in chapter 8 CTLS was used to calculate the covariance matrix. Here, the results obtained with the covariance matrix of chapter 7 (eqn. 9.2) are compared with the the results obtained with the covariance matrix (eqn. 9.3) generated with the CTLS approach with $\rho_{ux} = \rho_{uy} = \rho_{vy} = \rho_{wy} = \rho_{xy} = \rho_{xv} = \rho_{xw} = \rho_{yv} = \rho_{yw} = \rho_{vw} = 0$, *i.e.* independent noise, and variances $\sigma_u^2 = 0.001$, $\sigma_x^2 = \sigma_y^2 = \sigma_v^2 = 0.1$ and $\sigma_w^2 = 10^{-7}$, the subscripts denoting $y = \Delta T_p(k+1)$, $x = \Delta T_p(k)$, $u = \Delta u(k)$, $v = \Delta T_e(k)$ and $w = \Delta X_e(k)$ where Δ the measurement error. The simulation results of table 7.2 and 7.3 extended with CTLS results are given in table 9.1. In this table it can be seen that both methods have comparable

Table 9.1: MSE of predictors with parameters estimated by nonlinear LS, GTLS and CTLS in accordance with tables 7.2 and 7.3.

	calibration set 1	validation set 2	calibration set 2	validation set 1
nls	0.027	0.113	0.097	0.036
gtls	0.022	0.053	0.017	0.016
ctls	0.021	0.050	0.024	0.071

predictive performance except for the second cross-validation (last column of table 9.1). The difficulty, of course, is how to find the correct covariance matrix. Tuning is based on trial and error whereas CTLS gives a deterministic covariance matrix. If the noise characteristics are not known, however, the variances and covariances between the errors of the variables need to be approximated. Furthermore, the calculated covariance matrix is not always positive definite with an arbitrary choice of the variances and covariances.

The linear reparameterization method as introduced in chapter 6 is not limited to rational functions. Linear reparameterization methods also exist for other non-linear structures. For instance, exponential functions can be linearized by taking the logarithm. To illustrate this, the Freundlich isotherm [Dor95] that describes the adsorption of material on a particular surface is given as:

$$q_e = K_f C_e^{\frac{1}{n}} \quad (9.4)$$

$$C^{\text{tuning}} = \begin{bmatrix} 2.13 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.82 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.685 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 2.14 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 1.83 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 14.8 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.981 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.904 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 1.95 * 10^{-6} & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 1.09 * 10^{-6} & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 9.46 * 10^{-7} & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0423 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0351 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.685 \end{bmatrix} \quad (9.2)$$

$$C^{\text{CTLS}} = \begin{bmatrix} 0.05261 & 0.0094 & 0 & 0.0482 & 0.0054 & 0 & 0.0327 & 0.0033 & 0 & 3.52 * 10^{-5} & 3.77 * 10^{-6} & 0.0069 & 7.63 * 10^{-4} & 0.0055 \\ 0.00935 & 0.0127 & 0 & 0.00536 & 0.00903 & 0 & 0.00325 & 0.00578 & 0 & 3.77 * 10^{-6} & 6.39 * 10^{-6} & 7.63 * 10^{-4} & 0.00126 & 0.00444 \\ 0 & 0 & 0.100 & 0.00551 & 0.00444 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0482 & 0.00536 & 0.00551 & 0.0526 & 0.00939 & 0 & 0.0327 & 0.00326 & 0 & 3.52 * 10^{-5} & 3.78 * 10^{-6} & 0.00689 & 7.66 * 10^{-4} & 0 \\ 0.00536 & 0.00903 & 0.00444 & 0.00939 & 0.0128 & 0 & 0.00326 & 0.00580 & 0.0 & 3.78 * 10^{-6} & 6.41 * 10^{-6} & 7.66 * 10^{-4} & 0.00127 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.1 & 0.00551 & 0.00444 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0327 & 0.00325 & 0 & 0.0327 & 0.00326 & 0.00551 & 0.0409 & 0.0623 & 0 & 2.88 * 10^{-5} & 2.37 * 10^{-6} & 0.00465 & 4.59 * 10^{-4} & 0 \\ 0.00325 & 0.00578 & 0 & 0.00326 & 0.00580 & 0.00444 & 0.0623 & 0.00775 & 0 & 2.37 * 10^{-6} & 4.22 * 10^{-6} & 4.59 * 10^{-4} & 8.08 * 10^{-4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.00 * 10^{-7} & 5.51 * 10^{-9} & 4.44 * 10^{-9} & 0 & 0 & 0 \\ 3.52 * 10^{-5} & 3.77 * 10^{-6} & 0 & 3.52 * 10^{-5} & 3.78 * 10^{-6} & 0 & 2.88 * 10^{-5} & 2.37 * 10^{-6} & 5.51 * 10^{-9} & 3.23 * 10^{-8} & 6.70 * 10^{-9} & 5.09 * 10^{-6} & 5.40 * 10^{-7} & 0 \\ 3.77 * 10^{-6} & 6.39 * 10^{-6} & 0 & 3.78 * 10^{-6} & 6.41 * 10^{-6} & 0 & 2.37 * 10^{-6} & 4.22 * 10^{-6} & 4.44 * 10^{-9} & 6.70 * 10^{-9} & 8.29 * 10^{-9} & 5.40 * 10^{-7} & 9.02 * 10^{-7} & 0 \\ 0.00689 & 7.63 * 10^{-4} & 0 & 0.00689 & 7.66 * 10^{-4} & 0 & 0.00465 & 4.59 * 10^{-4} & 0 & 5.09 * 10^{-6} & 5.40 * 10^{-7} & 0.001 & 1.10 * 10^{-4} & 0 \\ 7.63 * 10^{-4} & 0.00126 & 0 & 7.66 * 10^{-4} & 0.00127 & 0 & 4.59 * 10^{-4} & 8.08 * 10^{-4} & 0 & 5.40 * 10^{-7} & 9.02 * 10^{-7} & 1.10 * 10^{-4} & 1.79 * 10^{-4} & 0 \\ 0.00551 & 0.00444 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1 \end{bmatrix} \quad (9.3)$$

with q_e the equilibrium concentration of e on the adsorbent, C_e the equilibrium concentration of e in the fluid phase and K_f and n are the parameters that describe the characteristics of the adsorption system. Taking the logarithm of (9.4) gives

$$\ln q_e = \ln K_f + \frac{1}{n} \ln C_e \quad (9.5)$$

Reparameterization of (9.5) leads to:

$$\ln q_e = [1, \quad \ln C_e] \begin{bmatrix} \tilde{K}_f \\ \tilde{n} \end{bmatrix} \quad (9.6)$$

with $\tilde{K}_f = \ln K_f$ and $\tilde{n} = \frac{1}{n}$. Both variables q_e and C_e are measured and, hence, subject to errors. Because the data variables q_e and C_e appear nonlinearly in eqn. (9.5) CTLS according to chapter 8 and appendix D is applied. This gives rise to the following data and covariance matrices:

$$[A, y] = \left[1, \quad \ln C_e + \frac{\sigma_C^2}{2C_e^2}, \quad \ln q_e + \frac{\sigma_q^2}{2q_e^2} \right] \quad (9.7)$$

$$C = \begin{bmatrix} \frac{\sigma_c^2(2c^4 - 5\sigma_c^2c^2 - 10\sigma_c^4)}{2c^6} & \frac{\rho_{cq}(c^2q^2 - \sigma_c^2q^2 - \rho_{cq}cq - \sigma_q^2c^2)}{c^3q^3} \\ \frac{\rho_{cq}(c^2q^2 - \sigma_c^2q^2 - \rho_{cq}cq - \sigma_q^2c^2)}{c^3q^3} & \frac{\sigma_q^2(2q^4 - 5\sigma_q^2q^2 - 10\sigma_q^4)}{2q^6} \end{bmatrix} \quad (9.8)$$

with σ_c^2 the variance of the error of C_e , σ_q^2 the variance of the error of q_e and ρ_{cq} the covariance between the errors of C_e and q_e . With these matrices the parameters \tilde{K}_f and \tilde{n} can be uniquely estimated.

This method using logarithms, however, cannot directly be applied to all nonlinear functions. An example of an exponential function that cannot be linearly reparameterized directly is given by:

$$y = a + e^{bx} \quad (9.9)$$

There are some possibilities, however, to come to a linear in the parameters expression. First, the exponential function can be approximated with a Taylor series expansion. Taking too few elements would lead to a relatively large error in the approximation. Taking too many elements leads to an ill-conditioned data-matrix. Another approach can be found in taking the derivative of y to x . The following expression is then obtained:

$$\frac{dy}{dx} = be^{bx} = by - ba \quad (9.10)$$

This can be further worked out, *e.g.* via

$$\tilde{y} = by - \tilde{a} \quad (9.11)$$

with $\tilde{y} = \frac{dy}{dx}$ and $\tilde{a} = ba$. The difficulty in this case is that the error in $\frac{dy}{dx}$ increases compared to the error in y .

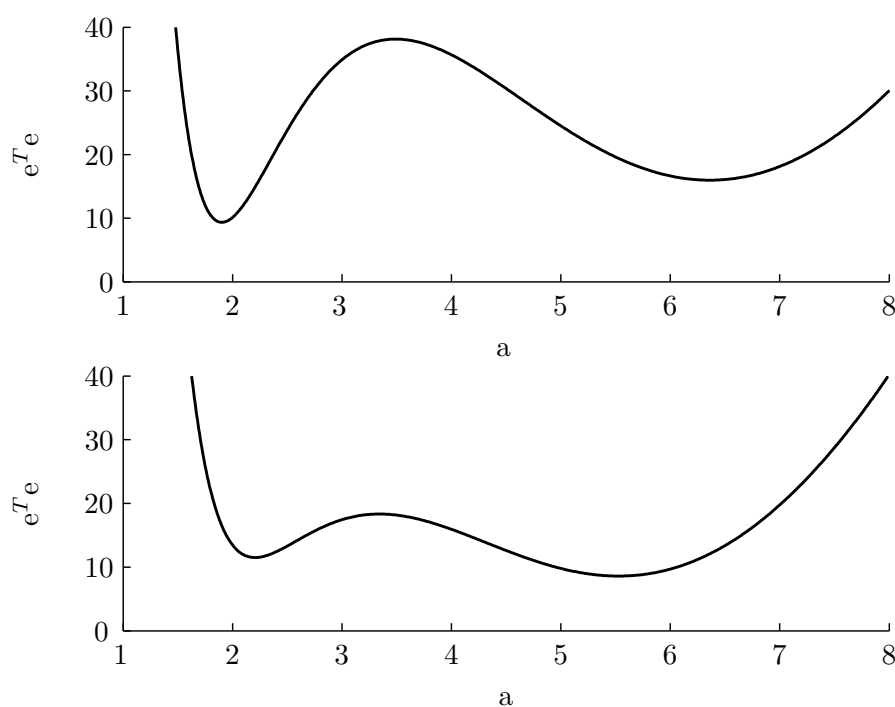


Figure 9.1: sum of squares of two different noise sequences $\{e\}_{k=1}^N$

9.3.2 Parameter uncertainty

The parameter vector that is found is the solution of a parameter estimation problem. Depending on the method an uncertainty measure, *i.e.* the second moment, can also be obtained. In general, however, it is not guaranteed that the global minimum results in the true parameters. This merely depends on the correspondence between the chosen noise model and the actual noise sequence.

If multiple local solutions exist with almost equal sum of squares it is worthwhile to investigate the probability density function of the parameter estimates. This is illustrated with the following single parameter system:

$$x(k+1) = x(k) - \frac{ax(k)}{a^2 + x(k)} \quad (9.12a)$$

$$y(k) = x(k) + e(k) \quad (9.12b)$$

Two data sets are generated with $a = 2$, $x_0 = 20$ and $\sigma_e^2 = 1$. In figure 9.1 it can be seen that there are two minima. Furthermore, the nonlinear least squares search routine from Matlab[®], *i.e.* `lsqnonlin`, only finds the global optimum if the initial guess of a is: $a \lesssim 4$. Using a bayesian approach [CL00, Nor86] for parameter estimation gives the conditional parameter distribution:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta} \quad (9.13)$$

with $p(\theta)$ the prior distribution of the parameters, $p(y|\theta)$ the data distribution given the parameter values and $\int p(y|\theta)p(\theta)d\theta$ the normalization factor. If $p(\theta)$ is assumed

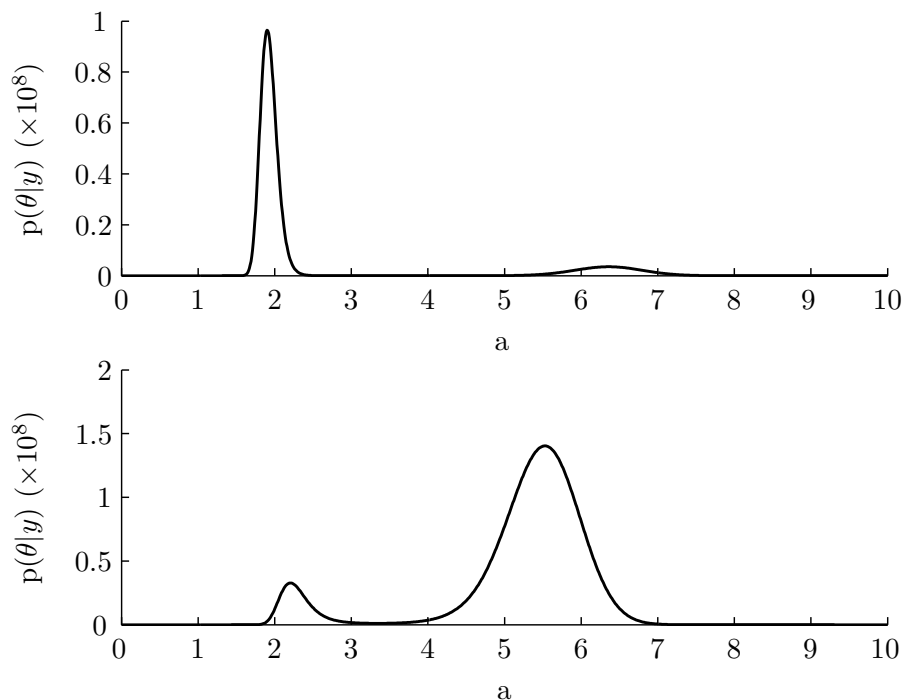


Figure 9.2: Probability distributions $p(\theta|y)$ of two different noise sequences $\{e\}_{k=1}^N$

to be uniformly distributed with $-\infty < \theta < \infty$ the distinction between conditional and unconditional parameter distributions vanishes [Nor86]. The unnormalized likelihood functions $p(\theta|y)$ from figure 9.1 are presented in figure 9.2.

In practice $p(\theta|y)$ can generally not be solved analytically. A numerical approach by sampling $p(\theta)$ is therefore required. The number of samples taken from $p(\theta)$, however, is limited especially with higher dimensions of θ . Given the sampled data a posterior probability density can be obtained. This posterior can then be used to sample new values of θ and create a new posterior distribution. More research is needed to see if $p(\theta|y)$ converges. If the posterior distribution is non-Gaussian as in figure 9.2 the probability can be approximated with a weighted set of normal distributions, *i.e.* Parzen densities [Par62]. Samples can be drawn easily from the normal distributions. Finally, given a sampled set of the parameter distribution a set of model predictions can be made and an uncertainty measure of the model prediction is directly available.

9.4 Epilogue

Two problems that are of interest in agricultural context have been studied in this thesis. The major contributions of this work are, however, of wider interest. The fact that weather forecasts can be improved for local situations can be relevant for a wide range of weather forecast users such as power plant managers, airports and harbors, water management etc. The possibility to linearly reparameterize a rational model is not limited to agricultural or biokinetic models. Together with total least squares for

estimation this approach is interesting in almost all applications of modeling. Further research must be done to generalize this approach to other types of nonlinearities as mentioned in section 9.3.1. Also multiple output problems are not dealt with in this thesis and are subject of further research.

Bibliography

- [Abe26] N.H. Abel. Beweis der unmöglichkeit algebraische gleichungen van höheren graden als dem vierten allgemein aufzulösen. *Journal für die reine und angewandte Mathematik*, 1:65–84, 1826.
- [AMH91] T. J. Abatzoglou, J. M. Mendel, and G. A. Harada. The constrained total least-squares technique and its applications to harmonic super-resolution. *Ieee Transactions on Signal Processing*, 39(5):1070–1087, 1991.
- [Åst70] KJ Åström. *Introduction to stochastic control theory*. 1970.
- [Atz95] A. J. Atzema. Determination of the quality of forecast moisture content of cut grass using forecast weather elements. *Journal of Agricultural Engineering Research*, 61(1):1–10, 1995.
- [BBT06] A. Beck and A. Ben-Tal. On the solution of the tikhonov regularization of the total least squares problem. *Siam Journal on Optimization*, 17(1):98–118, 2006.
- [Ber79] P. Berliner. *Psychrometrie: D. Thermodynamik d. lufttechn. Prozesse, Mischung, Kühlung, Erwärmung, Verdunstung, Entfeuchtung, Taubildung, Diffusion, Verdichtung, Nebelbildung, Aufstieg erwärmter Luft*. C.F. Mueller, Karlsruhe, 1979.
- [Bha00] K. Bhattacharya. Forecast surface temperature correction by estimation theory. *Indian Journal of Pure and Applied Mathematics*, 31(8):921–932, 2000.
- [BJ76] G. E. P. Box and M. J. Jenkins. *Time Series Analysis: forecasting and control*. Time series analysis and digital processing. Holden-Day, Oakland, 1976.
- [Box71] M. J. Box. Bias in nonlinear estimation. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 33(2):171–201, 1971.
- [CB02] A. Cornish-Bowden. Statistical analysis of enzyme kinetic data. In R. Eisenthal and M.J. Danson, editors, *Enzyme Assays*, pages 249–268. Oxford University Press, New York, 2nd edition, 2002.

- [CBCdMO02] J. P. Coelho, J. Boaventura Cunha, and P. B. de Moura Oliveira. Solar radiation prediction methods applied to improve greenhouse climate control. In *World Congress on Computers in Agriculture and Natural Resources*, pages 154–160, Iguaca falls, Brazil, 2002.
- [CBW96] Z. S. Chalabi, B. J. Bailey, and D. J. Wilkinson. A real-time optimal control algorithm for greenhouse heating. *Computers and Electronics in Agriculture*, 15(1):1–13, 1996.
- [CCC92] P. Chudoba, B. Capdeville, and J. Chudoba. Explanation of biological meaning of the S0/X0 ratio in batch cultivation. *Water Science and Technology*, 26(3-4):743–751, 1992.
- [CG01] M. K. Chourasia and T. K. Goswami. Losses of potatoes in cold storage vis-a-vis types, mechanism and influential factors. *Journal of Food Science and Technology-Mysore*, 38(4):301–313, 2001.
- [CHE93] JQ CHEN. Contrasting microclimates among clear-cut, edge, and interior of old-growth douglas-fir forest. *Agricultural and forest meteorology*, 63(3-4):219–237, 1993.
- [CL00] B. P. Carlin and T. A. Louis. *Bayes and empirical bayes methods for data analysis*. Chapman & Hall/CRC, Boca Raton, 2nd edition, 2000.
- [CLW⁺01] D. B. Crawley, L. K. Lawrie, F. C. Winkelmann, W. F. Buhl, Y. J. Huang, C. O. Pedersen, R. K. Strand, R. J. Liesen, D. E. Fisher, M. J. Witte, and J. Glazer. Energyplus: creating a new-generation building energy simulation program. *Energy and Buildings*, 33(4):319–331, 2001. Sp. Iss. SI.
- [COC05] J. P. Coelho, P. B. D. Oliveira, and J. B. Cunha. Greenhouse air temperature predictive control using the particle swarm optimisation algorithm. *Computers and Electronics in Agriculture*, 49(3):330–344, 2005.
- [DAPS06] P. I. Daskalov, K. G. Arvanitis, G. D. Pasgianos, and N. A. Sigrimis. Non-linear adaptive temperature and humidity control in animal buildings. *Biosystems Engineering*, 93(1):1–24, 2006.
- [Dem93] B. Demoor. Structured total least-squares and l2 approximation-problems. *Linear Algebra and Its Applications*, 188:163–205, 1993.
- [DK05] T. G. Doeswijk and K. J. Keesman. Adaptive weather forecasting using local meteorological information. *Biosystems Engineering*, 91(4):421–431, 2005.
- [Dor95] P. M. Doran. *Bioprocess Engineering Principles*. Elsevier Academic Press, London, UK, 1995.

- [DS96] J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, volume 16 of *Classics in Applied Mathematics*. Prentice Hall, Inc, Englewood Cliffs, NJ, 1996.
- [Duf06] P. Dufour. Control engineering in drying technology: Review and trends. *Drying Technology*, 24(7):889–904, 2006.
- [DVV95] D. Dochain, P. A. Vanrolleghem, and M. Vandaele. Structural identifiability of biokinetic models of activated-sludge respiration. *Water Research*, 29(11):2571–2578, 1995.
- [EJM+97] D. A. Epp, D. S. Jayas, W. E. Muir, N. D. G. White, and D. St George. Near-ambient drying of stored wheat using variable airflow - a simulation study. *Canadian Agricultural Engineering*, 39(4):297–302, 1997.
- [Fie97] S Fierro, RD. Regularization by truncated total least squares. *SIAM journal on scientific computing*, 18(4):1223–1241, 1997.
- [GA02] G. Galanis and M. Anadranistakis. A one-dimensional kalman filter for the correction of near surface temperature forecasts. *Meteorological Applications*, 9:437–441, 2002.
- [GAS84] S. Gal, A. Angel, and I. Seginer. Optimal-control of greenhouse climate - methodology. *European Journal of Operational Research*, 17(1):45–56, 1984.
- [Gei61] R. Geiger. *Das Klima der bodennahen Luftschicht : ein Lehrbuch der Mikroklimatologie*, volume 78 of *Die Wissenschaft*. Braunschweig, 4 edition, 1961.
- [Gel74] A. Gelb. *Applied Optimal Estimation*. MIT Press, Cambridge, Massachusetts, 1974.
- [GKGE77] L. Garfinkel, M. C. Kohn, D. Garfinkel, and L. Endrenyi. Systems-analysis in enzyme-kinetics. *Crc Critical Reviews in Bioengineering*, 2(4):329–361, 1977.
- [GM03] F. Grabowski and M. Marcotte. Drying of fruits, vegetables, and spices. In A. Chakraverty, A.S. Mujumdar, G.S.V. Raghavan, and H.S. Ramaswamy, editors, *Handbook of Postharvest Technology*, page 884. Marcel Dekker Inc., New York, 2003.
- [GNF03] K. Gottschalk, L. Nagy, and I. Farkas. Improved climate control for potato stores by fuzzy controllers. *Computers and Electronics in Agriculture*, 40(1-3):127–140, 2003.
- [God83] K Godfrey. *Compartmental models and their application*. Academic Press, London, 1983.

- [Gol99] GH Golub. Tikhonov regularization and total least squares. *SIAM journal on matrix analysis and applications*, 21(1):185–194, 1999.
- [GVL80] G. H. Golub and C. F. Van Loan. An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, 17(6):883–893, 1980.
- [Han94] Per Christian Hansen. Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems. *Numer. Algorithms*, 6(1-2):1–35, 1994.
- [HKFK04] G. P. Henze, D. E. Kalz, C. Felsmann, and G. Knabe. Impact of forecasting accuracy on predictive optimal control of active and passive building thermal storage inventory. *Hvac & R Research*, 10(2):153–178, 2004.
- [Hol82] A. Holmberg. On the practical identifiability of microbial-growth models incorporating michaelis-menten type nonlinearities. *Mathematical Biosciences*, 62(1):23–43, 1982.
- [Hom95] M. Homleid. Diurnal corrections of short-term surface temperature forecasts using the kalman filter. *Weather and Forecasting*, 10(4):689–707, 1995.
- [Kai80] T. Kailath. *Linear systems*. Prentice-Hall Information and System Sciences Series. Prentice-Hall, Englewood cliffs, New Jersey, 1980.
- [Kal60] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, 82(Series D):35–45, 1960.
- [Kee97] K. J. Keesman. Weighted least-squares set estimation from l_∞ -norm bounded-noise data. *IEEE Trans. Autom. Control*, 42(10):1456–1459, 1997.
- [KG91] J. C. Kabouris and A. P. Georgakakos. Stochastic-control of the activated-sludge process. *Water Science and Technology*, 24(6):249–255, 1991.
- [KPL03] K. J. Keesman, D. Peters, and L. J. S. Lukasse. Optimal climate control of a storage facility using local weather forecasts. *Control Engineering Practice*, 11(5):505–516, 2003.
- [KS72] H. Kwakernaak and R. Sivan. *Linear optimal control systems*. Wiley. Interscience, New York [etc.], 1972. Isn356840en.
- [KS02] K. J. Keesman and J. D. Stigter. Optimal parametric sensitivity control for the estimation of kinetic parameters in bioreactors. *Mathematical Biosciences*, 179(1):95–111, 2002.
- [KSv98] K. J. Keesman, H. Spanjers, and G. vanStraten. Analysis of endogenous process behavior in activated sludge. *Biotechnology and Bioengineering*, 57(2):155–163, 1998.

- [KYY⁺99] T. Katoh, D. Yuguchi, H. Yoshii, H. D. Shi, and K. Shimizu. Dynamics and modeling on fermentative production of poly (beta-hydroxybutyric acid) from sugars via lactate by a mixed culture of *Lactobacillus delbrueckii* and *Alcaligenes eutrophus*. *Journal of Biotechnology*, 67(2-3):113–134, 1999.
- [LB06] L. J. S. Lukasse and J. Bontsema. Weer in control ii. Openbaar eindrapport, Wageningen UR, 18 mei 2006 2006.
- [LdKCvdV06] L. J. S. Lukasse, J.E. de Kramer-Cuppen, and A.J. van der Voort. A physical model to predict climate dynamics in ventilated bulk-storage of agricultural produce. *International Journal of Refrigeration*, 2006. Accepted for publication.
- [LEI74] Max LEITH, CE. Theoretical skill of monte-carlo forecasts. *Monthly weather review*, 102(6):409–418, 1974.
- [LG94] L. Ljung and T. Glad. On global identifiability for arbitrary model parametrizations. *Automatica*, 30(2):265–276, 1994.
- [Lju87] L. Ljung. *System Identification, Theory for the User*. Information and System Sciences. Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1987.
- [LKvS96] L. J. S. Lukasse, K. J. Keesman, and G. van Straten. Grey-Box Identification of Dissolved Oxygen Dynamics in an Activated Sludge Process. In *Proc. 13th IFAC world congress, San Francisco, USA, Vol. N*, pages 485–490, 1996.
- [LOR63] Max LORENZ, EN. Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141, 1963.
- [LvdVdKC06] L. J. S. Lukasse, A.J. van der Voort, and J.E. de Kramer-Cuppen. Optimal climate control to anticipate future weather and energy tariffs. In K. Gottschalk, editor, *4th IFAC Workshop on Control Applications in Post - Harvest and Processing Technology*, volume 55 of *Bornimer Agrartechnische Berichte*, pages 109–122, Potsdam, Germany, 2006.
- [LVHDM97] P. Lemmerling, S. Van Huffel, and B. De Moor. Structured total least squares problems : formulations, algorithms and applications. In S. Van Huffel, editor, *Recent Advances in Total Least Squares Techniques and Errors-in-Variables modeling*, pages 215–223. SIAM, Philadelphia, 1997.
- [LVHDM02] P. Lemmerling, S. Van Huffel, and B. De Moor. The structured total least-squares approach for non-linearly structured matrices. *Numerical Linear Algebra with Applications*, 9(4):321–332, 2002.
- [Mak06] PA Makar. Heat flux, urban properties, and regional weather. *Atmospheric environment*, 40(15):2750–2766, 2006.

- [Mar06] C Marzban. Mos, perfect prog, and reanalysis. *Monthly weather review*, 134(2):657–663, 2006.
- [Mur93] A. H. Murphy. What is a good forecast? an essay on the nature of goodness in weather forecasting. *Weather and Forecasting*, 8(2):281–293, 1993.
- [NCE05] NCEP. History of recent modifications to the global forecast/analysis system, 2005.
http://wwwt.emc.ncep.noaa.gov/gmb/STATS/html/model_changes.html.
- [Nor86] J. P. Norton. *An Introduction to Identification*. Academic Press, London, 1986.
- [Par62] E. Parzen. Estimation of a probability density-function and mode. *Annals of Mathematical Statistics*, 33(3):1065–1076, 1962.
- [PBBP97] T. N. Palmer, J. Barkmeijer, R. Buizza, and T. Petroliaigis. The ecmwf ensemble prediction system. *Meteorological Applications*, 4(4):301–304, 1997.
- [PCKP05] S. Pinon, E. F. Camacho, B. Kuchen, and M. Pena. Constrained predictive control of a greenhouse. *Computers and Electronics in Agriculture*, 49(3):317–329, 2005.
- [PGM97] R. H. Perry, D. W. Green, and J. O. Maloney. *Perry’s chemical engineers’ handbook*. McGraw-Hill, New York, 1997.
- [RPG96] J. B. Rosen, H. S. Park, and J. Glick. Total least norm formulation and solution for structured problems. *Siam Journal on Matrix Analysis and Applications*, 17(1):110–126, 1996.
- [RPG98] J. B. Rosen, H. Park, and J. Glick. Structured total least norm for nonlinear problems. *Siam Journal on Matrix Analysis and Applications*, 20(1):14–30, 1998.
- [RPM06] A. Rynieccki, A. Pawlowska, and K. Molinski. Stochastic analysis of grain drying with unheated air under two different climates. *Drying Technology*, 24(9):1147–1152, 2006.
- [Sac04] M.P. Saccomani. Some results on parameter identification of nonlinear systems. *Cardiovascular Engineering: an International Journal*, 4(1):95–102, 2004.
- [Sch69] G. W. Schwert. Use of integrated rate equations in estimating kinetic constants of enzyme-catalyzed reactions. *Journal of Biological Chemistry*, 244(5):1278–1284, 1969.
- [Sch73] F. C. Schweppe. *Uncertain Dynamic Systems*. Prentice Hall, Englewood cliffs, New Jersey, 1973.

- [SHD06] G. S. Srzednicki, R. L. Hou, and R. H. Driscoll. Development of a control system for in-store drying of paddy in northeast china. *Journal of Food Engineering*, 77(2):368–377, 2006.
- [SK04] J. D. Stigter and K. J. Keesman. Optimal parametric sensitivity control of a fed-batch reactor. *Automatica*, 40(8):1459–1464, 2004.
- [SKS03] E. A. Seagren, H. Kim, and B. F. Smets. Identifiability and retrievability of unique parameters describing intrinsic Andrews kinetics. *Applied Microbiology and Biotechnology*, 61(4):314–322, 2003.
- [SL93] J.D. Simmons and B.D. Lott. Automatic fan control to reduce fan run time during warm weather ventilation. *Journal of Applied Poultry Research*, 2(4):314–323, 1993.
- [Son90] D. Sonntag. Important new values of the physical constants of 1986, vapour pressure formulations on the ITS-90, and psychrometer formulae. *Zeitschrift fuer Meteorologie*, 40(5):340–344, 1990.
- [SRH01] L. A. Smith, M. S. Roulston, and J. Hardenberg. End to end ensemble forecasting: Towards evaluating the economic value of the ensemble prediction system. Technical memorandum 336, ECMWF, 2001.
- [Ste94] R. F. Stengel. *Optimal Control and Estimation*. Dover publications, New York, 1994.
- [Tap00] F Tap. *Economics-based optimal control of greenhouse tomato crop production*. PhD thesis, Wageningen University, 2000.
- [TCR02] D. J. Tanner, A. C. Cleland, and T. R. Robertson. A generalised mathematical modelling methodology for design of horticultural food packages exposed to refrigerated conditions: Part 3, mass transfer modelling and testing. *International Journal of Refrigeration-Revue Internationale Du Froid*, 25(1):54–65, 2002.
- [Tei05] TJ Teisberg. The economic value of temperature forecasts in electricity generation. *Bulletin of the American Meteorological Society*, 86(12):1765, 2005.
- [TGB⁺95] M. B. Timmons, R. S. Gates, R. W. Bottcher, T. A. Carter, J. Brake, and M. J. Wineland. Simulation analysis of a new temperature control method for poultry housing. *Journal of Agricultural Engineering Research*, 62(4):237–245, 1995.
- [TVWVS96] R. F. Tap, L. G. Van Willigenburg, and G. Van Straten. Receding horizon optimal control of greenhouse climate using the lazy man weather prediction. In *Proceedings of the 13th IFAC World Congress*, pages 387–392, San Francisco, USA, 1996.

- [TZM01] Zoltan Toth, Yuejian Zhu, and Timothy Marchok. The use of ensembles to identify forecasts with small and large uncertainty. *Weather and Forecasting*, 16(4):463–477, 2001.
- [Van76] C. F. Vanloan. Generalizing singular value decomposition. *Siam Journal on Numerical Analysis*, 13(1):76–83, 1976.
- [Van98] G. Vandersteen. On the use of compensated total least squares in system identification. *IEEE Transactions on Automatic Control*, 43(10):1436–1441, 1998.
- [VHL02] S. Van Huffel and P. Lemmerling, editors. *Total Least Squares and Error-in-Variables modeling: Analysis, Algorithms and Applications*. Kluwer Academic Publishers, Dordrecht, 2002.
- [VHV89] S. Van Huffel and J. Vandewalle. Analysis and properties of the generalized total least squares problem $Ax \approx b$ when some or all columns in A are subject to error. *SIAM Journal on Matrix Analysis and Applications*, 10:294–315, 1989.
- [VHV91] S. Van Huffel and J. Vandewalle. *The Total Least Squares Problem: Computational Aspects and Analysis*, volume 9 of *Frontiers in Applied Mathematics*. SIAM, Philadelphia, 1991.
- [VPR96] S. VanHuffel, H. Park, and J. B. Rosen. Formulation and solution of structured total least norm problems for parameter estimation. *Ieee Transactions on Signal Processing*, 44(10):2464–2474, 1996.
- [vSCB00] G. van Straten, H. Challa, and F. Buwalda. Towards user accepted optimal control of greenhouse climate. *Computers and Electronics in Agriculture*, 26(3):221–238, 2000.
- [vSvWT02] G. van Straten, L. G. van Willigenburg, and R. F. Tap. The significance of crop co-states for receding horizon optimal control of greenhouse climate. *Control Engineering Practice*, 10(6):625–632, 2002.
- [VWAVB⁺05] A. V. Van Wagenberg, J. M. Aerts, A. Van Brecht, E. Vranken, T. Leroy, and D. Berckmans. Climate control based on temperature measurement in the animal-occupied zone of a pig room with ground channel ventilation. *Transactions of the Asae*, 48(1):355–365, 2005.
- [WAO⁺99] J. S. Wang, T. Araki, T. Ogawa, M. Matsuoka, and H. Fukuda. A method of graphically analyzing substrate-inhibition kinetics. *Biotechnology and Bioengineering*, 62(4):402–411, 1999.
- [WIL91] TB WILLIAMS. Microclimatic temperature relationships over different surfaces. *The Journal of geography*, 90(6):285–291, 1991.
- [Wil98] DS Wilks. Optimal use and economic value of weather forecasts for lettuce irrigation in a humid climate. *Agricultural and forest meteorology*, 89(2):115–129, 1998.

- [WMO96] World Meteorological Organization WMO. *Guide to Meteorological Instruments and Methods of Observation*, volume 8. WMO, Geneva, sixth edition, 1996.
- [YAOY03] N. Yildirim, F. Akcay, H. Okur, and D. Yidirim. Parameter estimation of nonlinear models in biochemistry: a comparative study on optimization methods. *Applied Mathematics and Computation*, 140(1):29–36, 2003.
- [Zhu02] YJ Zhu. The economic value of ensemble-based weather forecasts. *Bulletin of the American Meteorological Society*, 83(1):73, 2002.
- [Zhu05a] Q. M. Zhu. An implicit least squares algorithm for nonlinear rational model parameter estimation. *Applied Mathematical Modelling*, 29(7):673–689, 2005.
- [Zhu05b] YJ Zhu. Ensemble forecast: A new approach to uncertainty and predictability. *Advances in atmospheric sciences*, 22(6):781–788, 2005.
- [Zie01] C. Ziehmann. Skill prediction of local weather forecasts based on the ecmwf ensemble. *Nonlinear Processes in Geophysics*, 8(6):419–428, 2001.

Appendices

Appendix A

Nomenclature

Chapter 1

P	State covariance matrix
V	loss function
Ξ	error matrix $[e w]$
a	parameter
b	parameter
d	disturbance
e	measurement error
k	discrete time index
l	l -steps ahead prediction
m	number of measurements
p	parameter vector
u	control input
y	measurement vector
\hat{y}	model output prediction
w	measurement error
τ	time delay
\mathbb{Z}^*	set of nonnegative integers

Chapters 2 and 3

A	system matrix
B	input matrix
C	observation matrix
E	expectation value
G	noise matrix
I	identity matrix
K	Kalman gain matrix
M	final forecast horizon
P	covariance matrix of the estimated system states
Q	system or input noise covariance matrix
R	observation noise covariance matrix
V	standard deviation of the observation noise
W	correlation matrix
Φ	system transition matrix
i	index number
k	discrete-time index
p	length of output vector
u	input vector
v	observation noise

w	system or input noise
x	state vector
y	output vector
σ	standard deviation
σ_e	standard deviation of the forecast error
σ_w	standard deviation of the system or input noise
0	null matrix
\mathbb{N}	set of natural numbers
\mathbb{R}	set of real numbers

Chapter 4

A	discrete-time system matrix
E	expectation
F	continuous time system matrix
G	disturbance matrix
I	identity matrix
J	total costs
P	state covariance matrix
P_{tot}	atmospheric pressure
Q	covariance matrix of the disturbances
T	temperature
X	absolute humidity
a	parameter vector
b	parameter vector
c	colored noise vector
d	disturbance vector
k	discrete-time index
p	partial pressure
$p_1 \cdots p_{22}$	parameters
t	time
x	state vector
y	output vector
u	input vector
w	system noise
Δ	discretization interval
Φ	terminal costs
α	hatch position

γ	constant
μ	mean
ϕ	ventilation speed
σ	standard deviation
\mathcal{N}	set of natural numbers
subscripts	
a	air around the product
e	ambient air
p	product
s	saturation value
0	nominal value
ref	reference value

Chapter 5

x	state vector
k	discrete-time index
p	parameter vector
u	control input
d	disturbance input vector
T	temperature
H	maximum forecast horizon
J	total costs
E	expectation
L	total costs
l	time between two consecutive optimal control runs
R	total costs
r	forecast horizon
subscripts	
ref	reference value
vent	ventilation
mix	hatch position
obs	observation
fct	forecast
tot	total
superscripts	
rel	relative

Part II

C	covariance matrix
E	expectation
K_i	inhibition constant
K_m	Michaelis-Menten constant

K_s	monod half-saturation constant
$\mathcal{N}(\cdot)$	normal distribution
S	substrate concentration
T	temperature
X	absolute humidity
V_{\max}	maximum substrate conversion rate (discrete-time)
$X_{BH,0}$	biomass concentration
Y	yield
e	vector of measurement errors
k	discrete time index
m	number of measurements
n	number of parameters
p	parameter vector
r	respiration rate
t	time
x	state vector
y	measurement
u	vector of control variables
v	substrate conversion rate
v_{max}	maximum substrate conversion rate
Φ	data matrix
θ	reparameterized parameter vector
σ	standard deviation
μ_m	maximum specific growth rate
\mathbb{R}	set of real numbers
Subscripts	
p	product
e	ambient air
Superscripts	
0	nominal value
^	estimated value
~	modified

Appendix B

Storage model

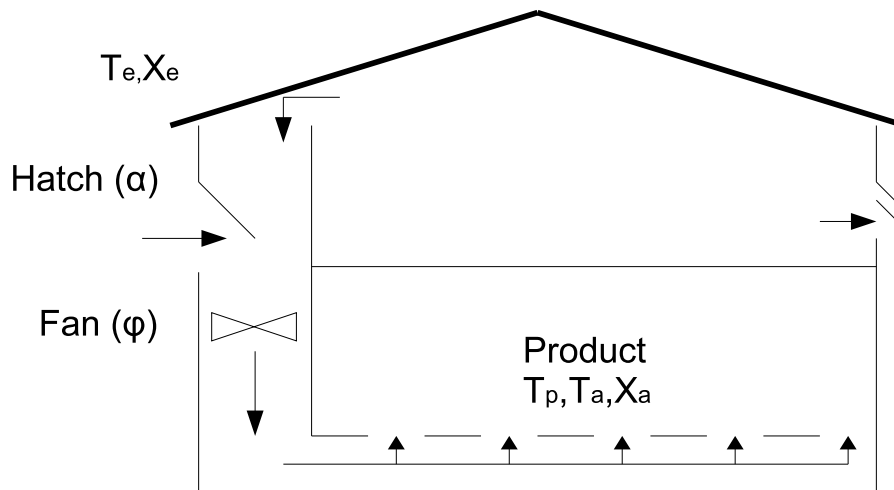


Figure B.1: Storage facility

In chapters 4, 6 and 7 the storage model adopted from [KPL03] is used with the lumped parameter vectors p (eqn. 4.4) and \tilde{p} (eqns. 6.14 and 7.7). In this appendix, the physical background of this model is presented. Furthermore, the relationship between the physical parameters and the parameter vector p of chapter 4 and \tilde{p} of chapters 6 and 7 is given.

B.1 Layout

The storage facility to be modeled consists of a potato pile, and a fan and hatches with which the temperature of the potatoes can be controlled. A schematic overview of the modeled storage facility is presented in figure B.1. Since the main purpose is to predict the *temperature* of the potatoes, the model is limited to respiration (with respect to heat production), heat transfer and evaporation. This results in the following states:

T_p potato temperature

T_a temperature of the air around the potatoes

X_a absolute humidity of the air around the potatoes

The following assumptions are given beforehand:

- The potato pile is spatially homogeneous distributed, *i.e.* each of the states T_p , T_a and X_a have only one value.
- The relative humidity at the potato surface is always 100%.
- Transpiration heat is completely removed from the potatoes.

- Condensation is not incorporated and thus a relative humidity greater than 100% is possible.
- The mixing ratio between ambient air and recirculation air is linearly related with the hatch position.
- Heat production of the fan is neglected.
- Respiration, heat capacity and density of the potatoes are constant during the whole period of storage.

B.2 Modeling

The general dynamical model of the storage facility is given in eqns. (B.1)-(B.3). A list of symbols is given in table B.2.

The product temperature (T_p) dynamics are described by

$$\frac{dT_p}{dt} = \frac{\left(\frac{T_a - T_p}{R_{pa}(\phi)}\right) + P_{resp} - rk_v A_{sp} \left(100 \left(-1.7011 + 7.7835e^{\frac{T_p}{17.0798}}\right) - \frac{X_a P_{tot}}{0.622 + X_a}\right)}{\rho_p V_p c_{p,p}} \quad (\text{B.1})$$

The first term on the right-hand side is related to heat transfer between the product (potatoes) and the air around the product. The second term reflects the produced heat due to respiration of the product and the last term represents the heat loss due to transpiration. The dynamics of the air temperature (T_a) is described by

$$\frac{dT_a}{dt} = \frac{\left(\frac{T_p - T_a}{R_{pa}(\phi)}\right) + \left(\frac{T_e - T_a}{R_{ea}}\right) + \frac{(\alpha\phi + \phi_l)\rho_a c_{p,a}(T_e - T_a)}{h}}{\rho_a V_a c_{p,a}} \quad (\text{B.2})$$

where the first term on the right-hand side is related to the heat transfer between the product and the air around the product. The second term reflects the heat transfer between ambient air (T_e) and the air around the product through the walls. The last term represents heat transfer due to forced ventilation with flow ϕ and leakage with flow ϕ_l , respectively. Finally, the water content of the air inside the storage facility is given by

$$\frac{dX_a}{dt} = \frac{k_v A_{sp} \left(100 \left(-1.7011 + 7.7835e^{\frac{T_p}{17.0798}}\right) - \frac{X_a P_{tot}}{0.622 + X_a}\right) + \frac{(\alpha\phi - \phi_l)\rho_a (X_e - X_a)}{h}}{\rho_a V_a} \quad (\text{B.3})$$

where the first term on the right-hand side reflects the transpiration of the product. The second term represents the water flux due to forced ventilation and leakage, respectively.

In short hand notation the model (B.1)-(B.3) can be written as:

$$\dot{x} = \begin{pmatrix} p_6 x_2 - p_6 x_1 + p_7 - p_8 e^{p_4 x_1} + p_9 x_3 (p_5 + x_3)^{-1} \\ p_{10} x_1 + (-p_{10} - p_{11} - p_{12} - p_{13} u) x_2 + (p_{11} + p_{12} + p_{13} u) d_1 \\ (-p_{12} - p_{13} u) x_3 + (p_{12} + p_{13} u) d_2 - p_{14} + p_{15} e^{p_4 x_1} - p_{16} x_3 (p_5 + x_3)^{-1} \end{pmatrix} \quad (\text{B.4})$$

which is equal to equation (4.4). The parameter vector p is defined by:

$$\begin{aligned} p_1 &= 100 \\ p_2 &= 1.7011 \\ p_3 &= 7.7835 \\ p_4 &= 17.0798^{-1} \\ p_5 &= 0.6228 \\ p_6 &= (R_{pa} \rho_p V_p c_{p,p})^{-1} \\ p_7 &= (P_{resp} + p_1 p_2 r k_v A_{sp} (\rho_p V_p c_{p,p})^{-1})^{-1} \\ p_8 &= r k_v A_{sp} p_1 p_3 (\rho_p V_p c_{p,p})^{-1} \\ p_9 &= r k_v A_{sp} P_{tot} (\rho_p V_p c_{p,p})^{-1} \\ p_{10} &= (R_{pa} \rho_a V_a c_{p,a})^{-1} \\ p_{11} &= (R_{ea} \rho_a V_a c_{p,a})^{-1} \\ p_{12} &= \phi_l (h V_a)^{-1} \\ p_{13} &= \phi_{max} (h V_a)^{-1} \\ p_{14} &= p_1 p_2 k_v A_{sp} (\rho_a V_a)^{-1} \\ p_{15} &= p_1 p_3 k_v A_{sp} (\rho_a V_a)^{-1} \\ p_{16} &= k_v A_{sp} P_{tot} (\rho_a V_a)^{-1} \\ p_{17} &= P_{tot} \\ p_{18} &= p_2 * p_5 \\ p_{19} &= p_3 * p_5 \end{aligned}$$

Model reduction and discretization in chapter 4 led to the resulting model given by eqn. (4.16). For clarity this result is again presented in eqn. (B.5).

$$\begin{aligned} T_p(k+1) &= \left(1 + p_{22} \left(\frac{p_6 p_{10}}{(p_{10} + p_{11} + p_{12} + p_{13} \alpha(k) \phi(k))} + \frac{p_{20} p_9 p_{15}}{p_5 (p_{12} + p_{13} \alpha(k) \phi(k)) + p_{16}} - p_6 - p_{20} p_8 \right) \right) T_p(k) \\ &\quad + p_{22} \frac{p_6 (p_{11} + p_{12} + p_{13} \alpha(k) \phi(k))}{(p_{10} + p_{11} + p_{12} + p_{13} \alpha(k) \phi(k))} T_e(k) + p_{22} \frac{p_9 (p_{12} + p_{13} \alpha(k) \phi(k))}{p_5 (p_{12} + p_{13} \alpha(k) \phi(k)) + p_{16}} X_e(k) \\ &\quad + p_{22} \left(p_7 + \frac{p_9 (p_{21} p_{15} - p_{14})}{p_5 (p_{12} + p_{13} \alpha(k) \phi(k)) + p_{16}} - p_{21} p_8 \right) \end{aligned} \quad (\text{B.5})$$

In chapters 6 and 7 the same model is used for parameter estimation but then with

an even more densed notation as in eqns. (6.14), (7.7) and (B.6)

$$\begin{aligned}
T_p(k+1) = & \left(\tilde{p}_1 + \frac{\tilde{p}_2}{\tilde{p}_3 + \tilde{p}_4 u(k)} + \frac{\tilde{p}_5}{\tilde{p}_6 + \tilde{p}_7 u(k)} \right) T_p(k) + \frac{\tilde{p}_8 + \tilde{p}_9 u(k)}{\tilde{p}_3 + \tilde{p}_4 u(k)} T_e(k) \\
& + \frac{\tilde{p}_{10} + \tilde{p}_{11} u(k)}{\tilde{p}_6 + \tilde{p}_7 u(k)} X_e(k) + \left(\tilde{p}_{12} + \frac{\tilde{p}_{13}}{\tilde{p}_6 + \tilde{p}_7 u(k)} \right)
\end{aligned} \tag{B.6}$$

Equations (B.5) and (B.6) are equivalent with

$$\tilde{p}_1 = 1 - p_{22}p_6 - p_{22}p_{20}p_8$$

$$\tilde{p}_2 = p_{22}p_6p_{10}$$

$$\tilde{p}_3 = p_{10} + p_{11} + p_{12}$$

$$\tilde{p}_4 = p_{13}$$

$$\tilde{p}_5 = p_{22}p_{20}p_9p_{15}$$

$$\tilde{p}_6 = p_5p_{12} + p_{16}$$

$$\tilde{p}_7 = p_5p_{13}$$

$$\tilde{p}_8 = p_{22}p_6(p_{11} + p_{12})$$

$$\tilde{p}_9 = p_{22}p_6p_{13}$$

$$\tilde{p}_{10} = p_{22}p_9p_{13}$$

$$\tilde{p}_{11} = p_{22}p_9p_{13}$$

$$\tilde{p}_{12} = p_{22}p_7 - p_{22}p_{21}p_8$$

$$\tilde{p}_{13} = p_{22}p_9(p_{21}p_{15} - p_{14})$$

Table B.1: List of symbols used in eqns. (B.1)-(B.3).

Symbol	Description	Units
α	relative position hatch	-
ϕ	air flow	$m^3 m^{-2} h^{-1}$
ϕ_l	air flow	$m^3 m^{-2} h^{-1}$
ρ_a	air density	$kg m^{-3}$
ρ_p	product density	$kg m^{-3}$
$c_{p,a}$	specific heat capacity of air	$kJ kg^{-1} ^\circ C^{-1}$
$c_{p,p}$	specific heat capacity of product	$kJ kg^{-1} ^\circ C^{-1}$
h	store room height	m
k_v	evaporation constant	$kg m^{-2} Pa^{-1} h^{-1}$
r	evaporation heat of water	$kJ kg^{-1}$
t	time	h
A_{sp}	specific area	$m^2 m^{-3}$
P_{resp}	respiration heat	$kJ m^{-3} h^{-1}$
P_{tot}	air pressure	Pa
R_{pa}	heat transfer resistance between product and air	$^\circ C m^3 h kJ^{-1}$
R_{ea}	heat transfer resistance between ambient air and air around the product	$^\circ C m^3 h kJ^{-1}$
T_a	air temperature around the product	$^\circ C$
T_e	ambient air temperature	$^\circ C$
T_p	product temperature	$^\circ C$
V_a	air volume per volume store room	$m^3 m^{-3}$
V_p	product volume per volume store room	$m^3 m^{-3}$
X_a	water content air around the product	$kg kg^{-1}$
X_e	water content ambient air	$kg kg^{-1}$

Appendix C

Stability of filter

The proposed forecasting system in Section 2.3 is checked on stability. Recall from systems theory that if (A, G) is stabilizable and (C, A) is detectable then $\lim_{t \rightarrow \infty} P(t) = P_1$, where P_1 is the solution of an algebraic Riccati equation. A sufficient condition for detectability of (C, A) is that it is observable. A sufficient condition for stabilizability of (A, Q) is that it is controllable.

As the system noise matrix G from Eqn (2.1) is time-varying the following theorem should hold for complete controllability [KS72]: the system is controllable if and only if for every i_0 there exists an $i_1 \geq i_0 + 1$ such that the symmetric nonnegative definite matrix:

$$W(i_0, i_1) = \sum_{i=i_0}^{i_1-1} \Phi(i_1, i+1)G(i)G^T(i)\Phi^T(i_1, i+1) \quad (\text{C.1})$$

is nonsingular. Here, Φ is the transition matrix of the system and is defined by:

$$\Phi(i+1, i_0) = A(i)\Phi(i, i_0) \quad i \geq i_0 \quad (\text{C.2})$$

$$\Phi(i_0, i_0) = I \quad (\text{C.3})$$

The system noise matrix G swaps between the values I and 0 . If there exists a time instant k for which $G(k) = I$ then, for all $i_0 \leq k$ and $i_1 = k + 1$, Eqn (C.1) can be rewritten as:

$$\begin{aligned} W(i_0, k+1) &= \sum_{i=i_0}^k \Phi(k+1, i+1)G(i)G^T(i)\Phi^T(k+1, i+1) \\ &= \Phi(k+1, k+1)G(k)G^T(k)\Phi^T(k+1, k+1) + \\ &\quad \sum_{i=i_0}^{k-1} \Phi(k+1, i+1)G(i)G^T(i)\Phi^T(k+1, i+1) \\ &= I + \sum_{i=i_0}^{k-1} \Phi(k+1, i+1)G(i)G^T(i)\Phi^T(k+1, i+1) \end{aligned} \quad (\text{C.4})$$

The matrix $W(i_0, k+1)$ is nonsingular and therefore the system is controllable when future inputs are available.

As the matrix A is time-varying the following result should hold for complete observability [KS72]: the system is observable if and only if for every i_1 there exists an $i_0 \leq i_1 - 1$ for which the nonnegative definite matrix:

$$M(i_0, i_1) = \sum_{i=i_0+1}^{i_1} \Phi^T(i, i_0+1)C^T(i)C(i)\Phi(i, i_0+1) \quad (\text{C.5})$$

is nonsingular. Transition matrix Φ is again defined by Eqns (C.2) and (C.3). Complete observability cannot be obtained through this theorem because of the term $C^T C = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & 0 \end{array} \right] \forall i$. The matrix M shall therefore always be singular. Hence, stability cannot be proven by these classical methods. However, the following theorem can be proven:

Theorem C.1. *Given discrete-time state-space system (2.1-2.2) with matrices defined in equations (2.10-2.14), i.e. short term forecasting system, and given no future inputs, i.e. no external forecasts, the following holds:*

$$\lim_{k \rightarrow \infty} P(k) = P_\infty$$

The solution of the discrete algebraic Riccati equation is: $P_\infty = 0$ [Sch73]:

$$P_\infty = P_\infty(1|0) - P_\infty(1|0)C^T [CP_\infty(1|0)C^T + R]^{-1}CP_\infty(1|0)$$

with

$$P_\infty(1|0) = AP_\infty A^T + GQG^T$$

Given future inputs at $k = k^$, the following holds: $P(k^* + 1|k^* + 1) = Q - QC^T(CQC^T + R)^{-1}CQ$*

Proof. Given $k \neq k^* \forall k$, matrix A is time-invariant of the form of Eqn (2.10) and $G = 0$ (Eqn 2.14). Consequently, Eqn (2.6) reduces to $P(k + 1|k) = AP(k|k)A^T$. After substituting $P_\infty(1|0) = AP_\infty A^T$, because $G = 0$, P_∞ is the solution of the algebraic Riccati equation

$$\begin{aligned} P_\infty &= AP_\infty A^T - AP_\infty A^T C^T (CAP_\infty A^T C^T + R)^{-1} CAP_\infty A^T \\ &= AP_\infty A^T \left[I - C^T (CAP_\infty A^T C^T + R)^{-1} CAP_\infty A^T \right] \end{aligned} \quad (C.6)$$

Consequently, the solution of (C.6) is: $P_\infty = 0$.

At $k = k^*$, $A = 0$ (2.11) and $G = I$ (2.13). Hence, Eqn. (2.6) reduces to $P(k^* + 1|k^*) = Q$. Consequently, the updated error covariance matrix is: $P(k^* + 1|k^* + 1) = Q - QC^T [CQC^T + R]^{-1}CQ$. Consequently, this covariance matrix only depends on the user-defined matrices Q and R . □

Appendix D

Bias compensation

This appendix presents an outline of the algorithm to calculate bias compensated data and covariance matrices according to the theory developed in [Van98]. One of the key aspects is linearization by Taylor series expansion. Before presenting the algorithm, two matrix operations are introduced.

Kronecker product Given $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$, the kronecker product denoted by \otimes is defined by:

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

vec-operator The vec-operator on matrix A is defined by:

$$\text{vec}(A) = [a_{11}, \dots, a_{m1}, a_{12}, \dots, a_{m2}, \dots, a_{1n}, \dots, a_{m.n}]^T$$

In the left column the general case is described whereas in the right column a specific example is elaborated.

$y = X\theta$ $\phi = [X, y]$ $\Phi^T(z) = [\phi(z), \frac{\partial \phi(z)}{\partial z^T}, \frac{\partial \phi(z)}{\partial z^{T[2]}}, \frac{\partial \phi(z)}{\partial z^{T[J]}}]$	$y = ax^2$ $\phi = [y, x^2]$ $\Phi^T(z) = \begin{bmatrix} x^2 & 2x & 0 & 2 & 0 & 0 & 0 \\ y & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$
--	--

where $X \in \mathbb{R}^{K \times p}$ the data matrix, $y \in \mathbb{R}^K$ the output vector, z the vector with the variables *e.g.* $z = [x, y]^T$, z^T the transpose of z , J the mixing order of the uncertain variables and $z^{[J]} = z \otimes z \otimes \cdots \otimes z$ with \otimes the kronecker product. Given Δz the mean error in z :

$N(\Delta z) = \begin{bmatrix} 1 & 0 & 0 & \cdots \\ \Delta z & I & 0 & \cdots & 0 \\ \Delta z^{[2]}/2! & I \otimes \Delta z & I^{[2]} & & 0 \\ \vdots & \vdots & & \ddots & \\ \Delta z^{[J]}/J! & I \otimes \Delta z^{[j-1]} & \cdots & \cdots & I \end{bmatrix}$	$N(\Delta z) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta x & 1 & 0 & 0 & 0 & 0 & 0 \\ \Delta y & 0 & 1 & 0 & 0 & 0 & 0 \\ 1/2\Delta x^2 & \Delta x & 0 & 1 & 0 & 0 & 0 \\ 1/2\Delta xy & \Delta y & 0 & 0 & 1 & 0 & 0 \\ 1/2\Delta xy & 0 & \Delta x & 0 & 0 & 1 & 0 \\ 1/2\Delta y^2 & 0 & \Delta y & 0 & 0 & 0 & 1 \end{bmatrix}$ <p>Assume $x \sim \mathcal{N}(0, \sigma_x^2)$ and $y \sim \mathcal{N}(0, \sigma_y^2)$</p> $t = [1 \quad 0 \quad 0 \quad -\frac{1}{2}\sigma_x^2 \quad -\frac{1}{2}\rho_{xy} \quad -\frac{1}{2}\rho_{xy} \quad -\frac{1}{2}\sigma_y^2]^T$
---	---

with σ the standard deviation and ρ_{xy} the covariance of x and y . In the example it is assumed that $J = 2$ and $\rho_{xy} = 0$.

$\tilde{\phi} = t^T \Phi \quad \tilde{\phi} \in \mathbb{R}^{1 \times p}$	$\tilde{\phi} = [x^2 - \sigma_x^2, y]$
--	--

Incorporating the data X and y gives the bias compensated data matrix $\tilde{\phi}$. Next, the covariance matrix of the bias compensated functions is calculated. The covariance matrix is calculated element by element.

$\text{cov}(\tilde{\phi}_k \tilde{\phi}_l) = t^{T[2]} E [N^{T[2]}(\Delta z)] \text{vec}(\Phi_k \Phi_l) - \phi_k \phi_l$	$\text{cov}(\tilde{\phi}_1 \tilde{\phi}_1) = 4x^2 \sigma_x^2 + 2\sigma_x^4$
---	---

with $\text{vec}(\cdot)$ the vec-operator. Finally, bias compensation is needed in the covariance matrix as in the example nonlinearities in the data appear, *i.e.* x^2 . This results in:

$C = \frac{1}{K} \sum_{k=1}^K \begin{bmatrix} \text{cov}(\tilde{\phi}_1 \tilde{\phi}_1) & \cdots & \text{cov}(\tilde{\phi}_1 \tilde{\phi}_{p+1}) \\ \vdots & & \vdots \\ \text{cov}(\tilde{\phi}_{p+1} \tilde{\phi}_1) & \cdots & \text{cov}(\tilde{\phi}_{p+1} \tilde{\phi}_{p+1}) \end{bmatrix}$	$C = \frac{1}{K} \sum_{k=1}^K \begin{bmatrix} 4x_k^2 \sigma_x^2 + 2\sigma_x^4 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}$
--	--

Summary

Many systems in agricultural practice are influenced by the weather. When these systems are closed they are usually controlled in order to keep the climate inside the system on a reference trajectory. In this context the weather is a disturbance of the system. In practice, the weather does not only act as a disturbance but is also necessary as a resource, *e.g.* the effect of global radiation on plant growth. An enhancement would be to use weather forecasts in control strategies. This idea was implemented in the EET-project "weather in control". The main purpose of the project "weather in control" was (i) to compute *optimal control* strategies that anticipate to changing environmental conditions, and (ii) to analyze the effects of weather forecast and parameter uncertainty on the predicted process states and calculated optimal inputs. In this context, the weather is no longer a disturbance but an external input driving the system. Weather forecast uncertainty in relation to model based control has been studied according to the following research questions:

1. Can data assimilation techniques be used to improve *local* weather forecasts?
2. How is it possible to integrate the state uncertainty in the goal function?
3. How are the control variable and state uncertainty affected if the uncertainty of the system states is part of the goal function?
4. What are the effects of weather forecast uncertainty and parameter uncertainty on the calculated costs in a receding horizon optimal control framework?

Not only input uncertainty is relevant to control but also the quality of the model itself. Therefore, calibration of models to data is a crucial step. In this thesis the following questions have been studied.

5. Is it possible to linearly reparameterize a rational model such that its parameters can be estimated with linear regression and what effects does it have on the predictive performance of the model?
6. Is total least squares a suitable technique to estimate the new parameters obtained with linear reparameterization such that the predictive performance of the predictor increases?

In chapters 2 and 3 a closer look is taken at the local weather forecasts, *i.e.* the weather forecast at one specific location. In chapter 2 weather forecasts and measurements from a meteorological station are compared. In chapter 3 the forecasts used for a greenhouse from the most nearby meteorological station, are compared with local measurements. A framework is presented in which local weather forecasts are updated using local measurements. Kalman filtering is used for this purpose as data assimilation technique. This method is compared and combined with diurnal bias correction. It is shown that the standard deviation of the forecast error can be reduced up to six hours ahead for temperature, up to 31 hours ahead for wind speed, and up to three hours for global radiation using local measurements. Diurnal bias correction is able to remove most of the the bias but has only a minor effect on the standard deviation. Combining both methods leads to a further increase in performance in terms of both bias and standard deviation.

In chapter 4 linearization and discretization techniques are used on a nonlinear model of a storage facility with forced ventilation using ambient air. Finally, a linear system describing the behavior of the potato temperature is obtained for a given trajectory of control inputs, *i.e.* ventilation rate. In order to predict the future temperature of the potatoes weather forecasts are needed. Because weather forecasts inherit uncertainty, the storage model is amenable to uncertainty. Based on a regionally linearized version of the storage model standard error propagation rules have been used to predict the system uncertainty analytically under the assumption of gaussian distributed noise.

Optimal control algorithms calculate controls in such a way that a prespecified cost criterion is minimized. As the uncertainty of the storage system is dependent on ventilation with outside air, the uncertainty of the potato temperature increases with increased ventilation. The medium-range weather forecast, up to 10 days ahead, consists of an ensemble of forecasts. From this ensemble the mean and the variance are obtained at each of the forecast intervals. The mean is used as the nominal input and the variance is used to calculate the error propagation. The predicted uncertainty is integrated into the cost criterion to allow a trade off between an optimal nominal solution and a minimum variance control solution. Because uncertainty is entering the system if the storage facility is ventilated with ambient air, the control input (ventilation) generally shifts towards parts where the weather forecast contains less uncertainty.

Another approach for assessing the uncertainty in weather forecasts and its effect optimal control solutions is presented in chapter 5. This numerical evaluation is based on an optimal control trajectory of the mean of the ensemble weather forecast with expected total costs. The total costs based on the optimal control trajectory with each of the ensemble members are calculated and the risks are analyzed. Furthermore,

the model parameters are not known exactly and hence subject to uncertainties. An ensemble of parameter vectors is generated. The total costs of each of the ensemble members is calculated in the same way as for the weather forecast ensemble. The worst-case scenarios show that if closed loop optimal control strategies are used the costs are more sensitive for model uncertainty than for weather uncertainty. Furthermore, if a new optimal control trajectory is calculated within 24 hours the cost increase is smaller than 5%.

Mathematical models are always an approximation of the real system. Model uncertainties arise in the model structure and/or in unknown parameter values. If measurements from the system are present, the model is fitted to the data by changing parameter values. Generally, parameters that are nonlinear in system model output are estimated by nonlinear least-squares (NLS) optimization algorithms. A major drawback of the NLS estimator is that it can end up in a local minimum. In chapters 6 and 8, a method is proposed for nonlinear discrete-time models with a polynomial quotient structure in input, output, and parameters ($x(k+1) = f(Z, p)$, with Z containing the previous inputs and outputs) to re-parameterize the model such that the model becomes linear in its new parameters. The new parameters (θ) can then be estimated by ordinary least squares and a global minimum is ensured. Finally, the model is rewritten in predictor form. This leads to the modified predictor: $\hat{x}(k+1) = \tilde{f}(Z, \hat{\theta})$. In chapter 6 this approach was used on the simplified storage model obtained in chapter 4. Real data was available to demonstrate the applicability of the procedure and the predictive quality of the modified predictor.

Rearranging and linearizing rational functions for parameter estimation is common practice (*e.g.* Lineweaver-Burk linearization). By rearranging, however, the error is distorted. In addition, the rearranged model frequently becomes errors-in-variables *i.e.* not only the output vector contains errors but also the regression vectors contain errors. The total least squares (TLS) approach provides a natural solution for these problems. Generalized total least squares (GTLS) is an extension to this that provides the covariance matrix between the output vector and the regression vector. In chapter 7 first the well-known Michaelis-Menten kinetics are used as an illustration with simulated (noisy) data. For low noise levels the predictive quality of the Michaelis-Menten model is better if the parameters are estimated with GTLS than with OLS. In addition to the Michaelis-Menten example, the simplified model of the storage facility containing biological products as obtained in chapter 4 is used as a real life example. Here, it is seen that ordinary least squares generates estimates with good predictive quality if the calibration data-set is informative enough. Otherwise, the predictive quality is poor. GTLS, however, provides estimates with good predictive quality for both data-sets.

In chapter 7 the covariance was merely used as a tuning parameter. In chapter 8 the covariance matrix is generated with an algorithm called bias compensated total least squares (CTLS). Nonlinearities in the data that contain errors may lead to biased estimates. CTLS accounts for these nonlinearities and generates a bias compensated data matrix and an accompanying covariance matrix. Two simulation examples that show the applicability of linear reparameterization and the advantages of CTLS over OLS are provided. The examples contain Michaelis-Menten kinetics and enzyme kinetics with substrate inhibition. In addition, the existence of local minima was clearly shown for the substrate inhibition example. Finally, the applicability of CTLS was demonstrated with real data of an activated sludge experiment. It is concluded that CTLS is a powerful alternative to the existing least-squares methods for rational biokinetic models.

Samenvatting

Veel systemen in de agrarische praktijk staan onder invloed van het weer. Als de systemen gesloten zijn worden ze geregeld om het klimaat binnen het systeem op een gewenst niveau te houden. Vanuit deze zienswijze is het weer een verstoring op het systeem. Het weer werkt niet alleen als verstoring op het systeem maar kan ook dienen als bron, bijvoorbeeld zoninstraling is nodig voor plantengroei. Vanuit dit kader bezien zou het een vooruitgang zijn om weersverwachtingen te gebruiken voor regelstrategieën. Dit idee is geïmplementeerd in het EET-project "weer in control". De hoofddoelen van dit project betreffen: (i) het berekenen van optimale stuurprofielen die anticiperen op verwachte veranderingen van het weer, en (ii) het analyseren van de effecten van onzekerheid in weersverwachtingen en van onzekerheid in de geschatte parameters op de onzekerheid van de voorspelde model toestanden. In dit geval wordt het weer niet behandeld als een verstoring maar als externe ingangsvaariabele die fungeert als drijvende kracht van het systeem. Onzekerheid in weersverwachtingen in relatie tot model gebaseerd regelen is onderzocht naar aanleiding van de volgende onderzoeksvragen:

1. Kunnen data-assimilatie technieken gebruikt worden om lokale weersverwachtingen te verbeteren?
2. Op welke wijze kan de onzekerheid van modeltoestanden worden meegenomen in de doelfunctie?
3. Hoe worden de stuurvariabele en de onzekerheid van de modeltoestanden beïnvloed als de onzekerheid van de modeltoestanden wordt meegenomen in de doelfunctie?
4. Wat zijn de effecten van onzekerheid in weersverwachtingen en de onzekerheid van model parameters op de berekende kosten in een receding horizon optimal control raamwerk

Niet alleen onzekerheid van de ingangsvaariabelen is van invloed op de regeling maar ook de kwaliteit van het model zelf. Daarom is calibratie van modellen op data een cruciale stap. In dit proefschrift zijn de volgende vragen bestudeerd:

5. Is het mogelijk om rationele modellen linear te reparameteriseren zodat de nieuwe parameters geschat kunnen worden met lineaire regressie en wat voor effect heeft dit op de voorspellende kracht van het model?
6. Is total least squares een bruikbare techniek om de parameters van het lineair gereparameteriseerde model te schatten zodat de voorspellende kracht van het model wordt vergroot?

In hoofdstukken 2 en 3 wordt de aandacht gericht op de lokale weersverwachtingen, dat wil zeggen lokale metingen en weersverwachtingen uitgegeven voor die lokatie worden met elkaar vergeleken. In hoofdstuk 2 betreffen zowel de metingen als verwachtingen een specifiek meteorologisch station, te weten "De Bilt". In hoofdstuk 3 worden de verwachtingen voor een tuinbouwkas van de dichtst bijzijnde meteorologische locatie vergeleken met lokale metingen van de betreffende kas. Een raamwerk waarin lokale metingen worden gebruikt om de weersverwachtingen aan te passen wordt gepresenteerd. Kalman filtering wordt hierbij gebruikt als data-assimilatie techniek. De methode is vergeleken met en gecombineerd met een methode genaamd "diurnal bias correction". De standaard deviatie van de voorspelfout kan gereduceerd worden tot 6 uur vooruit voor temperatuur, tot 3 uur vooruit voor globale straling en tot 31 uur vooruit voor windsnelheid met behulp van de voorgestelde methode. Diurnal bias correction verwijdert wel de systematische afwijking van de voorspelfout maar leidt nauwelijks tot vermindering van de standaard deviatie. Een combinatie van beide methodes leidt tot een vermindering van de systematische afwijking en tot een verlaging van de standaard deviatie.

In hoofdstuk 4 worden linearisatie en discretisatie technieken gebruikt op een model van een aardappelbewaarpplaats die wordt geventileerd met buitenlucht. Het model beschrijft het gedrag van de aardappeltemperatuur in de bewaarplaats. Uiteindelijk wordt een lineair model verkregen voor een gegeven stuurpatroon (ventilatiesnelheid). Om de toekomstige aardappeltemperatuur te kunnen beschrijven zijn weersverwachtingen nodig. Omdat in weersverwachtingen nogal wat onzekerheden zitten zijn de modelvoorspellingen ook vatbaar voor onzekerheid. Standaard foutpropagatie technieken zijn gebruikt op een regionaal gelineariseerd aardappel bewaarmodel om de onzekerheid van de aardappeltemperatuur analytisch te kunnen berekenen. Hierbij is aangenomen dat de fout normaal verdeeld is.

Optimale besturingsalgoritmen berekenen de stuurvariabele dusdanig dat een bepaald kostencriterium wordt geminimaliseerd. Omdat het bewaarsysteem wordt geventileerd met buitenlucht zal de onzekerheid van de voorspelde aardappeltemperatuur toenemen indien er actief geventileerd wordt. De middellange termijn weersverwachting, tot 10 dagen vooruit, bestaat uit een ensemble van verwachtingen. Vanuit dit ensemble kunnen het gemiddelde en de variantie op elk van de verwachtingsinter-

vallen worden berekend. Omdat onzekerheid alleen het model in komt als er actief geventileerd wordt met buitenlucht verschuiven de stuurvariabelen naar tijdstippen waar de onzekerheid van de weersverwachtingen kleiner is.

Een andere aanpak om de onzekerheid in weersverwachtingen en de effecten hiervan op de optimale besturing te benaderen wordt gegeven in hoofdstuk 5. Deze numerieke benadering is gebaseerd op een berekend optimaal stuurprofiel aan de hand van het gemiddelde van de ensemble weersverwachting en verwachte totale kosten. Voor elk ensemble lid afzonderlijk worden de totale kosten berekend en worden de risico's geanalyseerd. Modelonzekerheid is in dit hoofdstuk benaderd door de parameters te laten variëren. Hierbij wordt een ensemble van parameter vectoren gegenereerd die elk het "echte" systeem kunnen voorstellen. De totale kosten van elk van de parameter vectoren worden berekend. In het geval de hoogste extra kosten worden vergeleken voor beide gevallen blijkt dat de kosten gevoeliger zijn voor modelonzekerheid dan voor weersonzekerheid. In een gesloten lus regelstrategie (RHOC) blijken de uiteindelijke kosten minder dan 5% te stijgen voor weersonzekerheid als de optimale sturingen zijn herberekend met een interval van 24 uur of minder.

Wiskundige modellen zijn altijd een benadering van de werkelijkheid. Modelonzekerheid komt naar voren in model structuur en /of model parameters. Als metingen van het systeem beschikbaar zijn wordt het model gefit op de metingen door het variëren van de parameters. Normaal gesproken worden parameters die niet lineair voorkomen in modellen geschat met behulp van niet-lineaire kleinste kwadraten (NLS) algoritmes. Een groot nadeel van NLS schatters is dat ze in lokale minima kunnen uitkomen. In hoofdstuk 8 wordt een methode geïntroduceerd om niet lineaire discrete tijd modellen met een polynomiale quotient structuur in ingangsvARIABLEN, uitgangsvARIABLEN en parameters ($x(k+1) = f(Z, p)$, waarin Z de voorgaande inputs en outputs bevat) te reparameteriseren zodanig dat een vergelijking ontstaat die lineair is in de parameters. De nieuwe parameters (θ) kunnen dan geschat worden met de gewone kleinste kwadraten methode en een minimum is gegarandeerd. Uiteindelijk wordt de voorspelling berekend op basis van het model in gemodificeerde voorspellende vorm en wordt geschreven als: $\hat{x}(k+1) = \tilde{f}(Z, \hat{\theta})$. Deze aanpak is toegepast in hoofdstuk 8 op het vereenvoudigde aardappel bewaarmodel uit hoofdstuk 4. Echte meetdata was beschikbaar om de toepasbaarheid van de procedure en voorspellende kwaliteit van het gemodificeerde model aan te tonen.

Het herschikken en lineariseren van rationale functies wordt veelvuldig toegepast (bv Lineweaver-Burk linearizatie). Door het herschikken wordt echter ook de foutstructuur verstoord. Verder kunnen de fouten zowel aan de linkerzijde als de rechterzijde van de vergelijking terecht komen. Total least squares (TLS) is een methode voor het oplossen van dit type problemen. Generalized total least squares (GTLS) is een uitbreiding hierop waarbij een covariantie matrix van de regressoren kan worden

opgegeven. In hoofdstuk 7 wordt deze methode geïllustreerd aan de hand van de Michaelis-Menten vergelijking en gesimuleerde data. Voor lage ruisniveau's presteert de voorspeller beter als de parameters geschat zijn met GTLS dan met OLS. Hierna is de GTLS methode toegepast op het vereenvoudigde aardappel bewaarmodel van hoofdstuk 4. In dit geval presteert het model met schattingen uit OLS goed als de data van de calibratie set voldoende informatie bevatten. Is dit niet het geval dan is de voorspellende kwaliteit slecht. Voor GTLS geldt dat in beide gevallen een goede voorspellende kwaliteit van het gemodificeerde model wordt gevonden.

In hoofdstuk 7 werd de covariantie matrix bepaald door tuning. In hoofdstuk 8 wordt de covariantie matrix bepaald aan de hand van een algoritme genaamd: bias compensated total least squares (CTLS). Niet-lineariteiten in de data kunnen leiden tot onzuivere schattingen. CTLS houdt rekening met deze niet-lineariteiten en genereert een "bias compensated" data matrix en een bijbehorende "bias compensated" covariantie matrix. In twee simulatie voorbeelden wordt de toepasbaarheid van CTLS weergegeven en de voordelen ten opzichte van OLS aangetoond. Deze voorbeelden betreffen de Michaelis-Menten vergelijking en een model voor enzym kinetiek met substraat-inhibitie. Het bestaan van lokale minima was duidelijk zichtbaar voor het voorbeeld van substraat-inhibitie. Uiteindelijk is de toepasbaarheid van CTLS aangetoond met echte data uit een actief slib experiment. Er wordt geconcludeerd dat CTLS een krachtig alternatief vormt voor de bestaande kleinste kwadraten methodes voor rationele biokinetische modellen.

Dankwoord

Tijdens mijn studie had ik mezelf niet zo snel verder zien gaan in een promotie-onderzoek. Hoewel een van mijn sollicitaties op een aio plaats gericht was ben ik uiteindelijk in de procesautomatisering terecht gekomen. De aanraking met industriële toepassing was boeiend en ik heb hier dan ook veel van geleerd. Echter, de inhoudelijke kant van de processen bleef op de achtergrond en uiteindelijk heb ik de kans aangegrepen om dit op te pakken middels het voor u liggende promotie-onderzoek.

Voor de meeste aio's geldt dat dit hun eerste echte baan is (sommige onder ons vinden het echter een betaalde studie) en de problemen die hierbij ondervonden worden resulteren dan in de bekende aio-dip. Waarschijnlijk door de voorgaande banen en de daaruit voortvloeiende welbewuste keuze voor onderzoek is mij deze dip bespaart gebleven. Dit was echter niet gelukt wanneer mijn begeleider, Karel Keesman, mij niet met raad en daad had bijgestaan. Karel, bedankt voor al je opmerkingen en suggesties en de wekelijkse discussies die inderdaad met de tijd in lengte toenamen. Verder wil ik graag Gerrit bedanken. In het begin was je alleen via de projectvergaderingen van weer in control actief betrokken en vertrouwde je de verdere inhoud toe aan Karel en mij. Een retourtje Potsdam, hoofdstukken met betrekking op optimal control en het afronden van het proefschrift leverden echter nog voldoende op om je befaamde kritische noten te verwerken.

Dit brengt mij op het zojuist genoemde EET-project "weer-in-control II". Dit EET-project (EETK01120) is uitgevoerd in samenwerking met de onderzoeksgroepen Greenhouse technology en QUIC van WUR en de bedrijven Tolsma Techniek, PRIVA BV en Weathernews Benelux. Allereerst moet ik Leo bedanken die dit project geïnitieerd heeft. Naast je inzet voor het project en het kritisch doorlezen van mijn papers heeft deze samenwerking ook nog een fiets opgeleverd. Verder wil ik Peter bedanken voor alle informatie over de meteorologie en de weersverwachtingen. Ook alle overige leden van de werkgroep wil ik bij deze bedanken voor de prettige werksfeer en de altijd interessante en daarom uitlopende vergaderingen. Als blijk van de goede samenwerking kan dan ook wel het afsluitende diner in "De Lier" genoemd worden. Omdat een van de voornaamste commentaren op mijn artikelen de acknowledgements betrof zal ik hiermee afsluiten: Leo, Aart-Jan, Janneke, Eldert, Jan, Enrico, Jan, Gerrit, Theo, Jorrit, Hendrik, Peter en de stuurgroepleden Hans, Sjaak,

Simon, Wil, Elbert en Roger bedankt!

Een goede werksfeer werkt stimulerend en mede hierdoor heb ik mijn onderzoek tijdig kunnen afronden. Dit begint bij de kamergenoten waarvan Bas (toen nog als student) de eerste was. Later heb ik met Martijn de IMAG-tijd vol gemaakt. We hebben samen onze eerste stappen het onderzoek gezet na enige jaren in de automatisering te hebben gewerkt. De langste tijd heb ik met Dirk doorgebracht. Vanaf het moment dat we het technotron betraden hebben we leuke discussies gehad over onze beide onderwerpen. Verder heb je me op het pad gezet van velerlei alternatieve software programma's. Ik vind het dan ook erg leuk dat je de zware taak van paranimf op je wilt nemen. Ook de overige (ex-)medewerkers en -aio's wil ik bij deze bedanken. Marja, Rachel, Tijmen, Hady, Djaeni, Kees, Kees, Wilko, Ton, Gerard, Johan, Hans, Zita, Klaas, Gert, Stefan en Ilse de discussies aan de koffietafel waren vaak interessant en mede daardoor kostten ze ook af en toe meer tijd dan gepland.

Dan kom ik bij mijn tweede paranimf: Roy. Ik vind het erg leuk dat je voor deze gelegenheid even terugkomt naar Nederland. We zien elkaar niet vaak maar je komt wel trouw aan als je weer in de buurt bent. Samen met Paul komen we zo af en toe zelfs tot concrete afspraken. Hopelijk komt er niets tussen bij een volgend weekendje naar Engeland.

Mijn ouders wil ik graag bedanken voor het mogelijk maken van mijn studie waardoor ik uiteindelijk ook aan dit promotie-onderzoek heb kunnen beginnen. Het is jullie misschien niet duidelijk geweest wat dit onderzoek inhield maar jullie houden er een mooi boekje aan over.

Terugkomend op de aio-dip, een van de grootste redenen om hier niet in terecht te komen zijn onze lieve dochters Tosca en Blanca en zoon Milo. Door jullie geklets, gelach (en gehuil) was ik de problemen en tegenslagen van het onderzoek snel weer vergeten en zorgden jullie voor de nodige gezelligheid. Als laatste wil ik natuurlijk Chantal bedanken. Zonder haar had ik nooit de stap terug naar het onderzoek durven nemen. Lieve Chantal, je hebt me leren openhartig te zijn waardoor we alles kunnen delen. Je bent hierdoor mijn steun en toeverlaat en als zodanig is dit resultaat ook jouw resultaat.

Curriculum Vitae

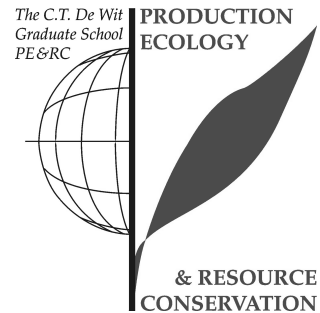
Timotheus Gerardus (Timo) Doeswijk was born November 17, 1972 in Lisse, The Netherlands. He completed his secondary school in 1991 at the Fioretti College in Lisse. From 1991 until 1997 he studied Bioprocess Technology at the agricultural university in Wageningen. He conducted two master's theses. The first was about 'Optimal control evaluation of a solid-state fermentation' at the systems and control group. The second was about 'Optimization of the catch of noisy populations with Kalman filtering' at the mathematics department. After his graduation in 1997 he started working at Process Management and Control B.V. He was Mainly posted at Yokogawa Netherlands BV where he worked at several projects with Yokogawa distributed control systems (DCS). Implementation and testing of process control software and assisting with start-up of plants where his main responsibilities. The main projects were in the chemical/petrochemical industries (*e.g.* Pasmenco Budel Zinc, Nederlandse Aardolie Maatschappij). In 2000 he started at Vertis B.V. where he worked on several projects of process control and automation. Besides these projects he assisted in the implementation of Oracle databases. Finally, in 2002, he started his Ph.D. project 'Weather in Control' at the systems and control group of Wageningen University. The results of this research is described in this thesis.

List of Publications

- [1] T. G. Doeswijk and K. J. Keesman. Adaptive weather forecasting using local meteorological information. *Biosystems Engineering*, 91(4):421–431, 2005.
- [2] T. G. Doeswijk and K. J. Keesman. Improving local weather forecasts for agricultural applications. In *Signal and Image Processing, Second IASTED Intern. Multi-Conference on Automation, Control, and Information Technology*, pages 421–431, Novosibirsk, Russia, 2005. IASTED.
- [3] T. G. Doeswijk and K. J. Keesman. Parameter estimation and prediction of a nonlinear storage model: an algebraic approach. In *Proceedings of the 2005 International Conference on Control & Automation*, Budapest, 2005. IEEE.
- [4] T. G. Doeswijk and K. J. Keesman. Linear parameter estimation of rational biokinetic functions. 2006. Submitted for publication.
- [5] T. G. Doeswijk and K. J. Keesman. Parameter estimation and prediction of nonlinear biological systems: some examples. In *14th IFAC Symposium on System Identification*, Newcastle, Australia, 2006. Elsevier.
- [6] T. G. Doeswijk, K. J. Keesman, and G. vanStraten. Impact of weather forecast uncertainty in optimal climate control of storehouses. In K. Gottschalk, editor, *4th IFAC Workshop on Control Applications in Post - Harvest and Processing Technology*, volume 55 of *Bornimer Agrartechnische Berichte*, pages 46–57, Potsdam, Germany, 2006.
- [7] T.G. Doeswijk, K.J. Keesman, and G. van Straten. Uncertainty analysis of a storage facility under optimal control. 2007. Submitted for publication.

PE&RC PhD Education Statement Form

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 22 credits (= 32 ECTS = 22 weeks of activities)



Review of Literature (5.6 credits)

- Total least squares

Writing of Project (4.2 credits)

- Weather in control; PE&RC 41 04a

Post-Graduate Courses (7 credits)

- Design methods for control systems; DISC (2003)
- Uncertainly analysis; Wimek (2004)

Competence Strengthening / Skills courses (3 credits)

- Techniques for writing and presenting a scientific paper; WGS (2005)
- Career perspectives; WGS (2006)

PhD Discussion Groups (4.5 credits)

- Math&Stat (2004/5/6)
- Dutch institute for systems and control (2004)

PE&RC Annual Meetings, Seminars and Introduction Days (1 credit)

- PE&RC day (2004/5/6)
- Computational modelling for plant biology (2006)

International Symposia, Workshops and Conferences (10.5 credits)

- IASTED Automation control and information technology (2005)
- IEEE International conference on control and automation (2005)
- Benelux meeting (2004/5/6)