Paper F

Application of a Nonlinear Mechanistic Model and an Infinite Horizon Predictive Controller on Paper Machine 6 at Norske Skog Saugbrugs

Hauge, T.A., Slora, R., and Lie, B. (2002). *Application of a Nonlinear Mechanistic Model and an Infinite Horizon Predictive Controller on Paper Machine 6 at Norske Skog Saugbrugs*, Submitted to Journal of Process Control.

Extended version.

Application of a Nonlinear Mechanistic Model and an Infinite Horizon Predictive Controller to Paper Machine 6 at Norske Skog Saugbrugs

Tor Anders Hauge[*], Roger Slora[†], and Bernt Lie[‡]

Contents

[*]Telemark University College.
[†]Norske Skog Saugbrugs.
[‡]Corresponding author: Bernt.Lie@hit.no, Telemark University College, P.b. 203, 3901 Porsgrunn, Norway.

Abstract

A mechanistic nonlinear model of the wet end of paper machine 6 (PM6) at Norske Skog Saugbrugs, Norway has been developed, and used in an industrial MPC implementation. The MPC uses an infinite horizon criterion, successive linearization of the model, and estimation of states and parameters by an augmented Kalman filter. Variation in important quality variables and consistencies in the wet end have been reduced substantially, compared to the variation prior to the MPC implementation. The MPC also provides better efficiency through faster grade changes, control during sheet breaks and start ups, and better control during periods of poor measurements. From May 2002 the MPC has been the preferred controller choice for the process operators at PM6.

1 Introduction

Norske Skog Saugbrugs in Halden, Norway, is one of the largest manufacturers of uncoated super calendered magazine paper in the world. The total production capacity of the mill is 550, 000 ton per year. The largest paper machine (PM) at the Saugbrugs mill is PM6, accounting for more than half the total production capacity. PM6 was built in the early 1990s and produce paper with 8.62 meters width, and with a typical velocity of 1500 meters per minute.

Previous work A clear distinction is usually made between CD (Cross Direction) and MD (Machine Direction) control, when discussing the control of a paper machine. CD control refers to the profile across the paper web, while MD refers to the average value. In this paper we only consider the MD control problem.

Several MPC implementations using multivariable empiric paper machine models are reported, e.g. (McQuillin & Huizinga 1995), (Lang, Tian, Kuusisto & Rantala 1998), (Mack, Lovett, Austin, Wright & Terry 2001), (Kosonen, Fu, Nuyan, Kuusisto & Huhtelin 2002), and (Austin, Mack, Lovett, Wright & Terry 2002). To the best of the authors' knowledge, there exists no reported industrial MPC implementations utilizing a multivariable mechanistic model of the wet-end of the paper machine. Some industrial implementations of MPC with mechanistic models are known in other industry areas, e.g. (Qin & Badgwell 1998) and (Badgwell & Qin 2001) have reported a few implementations. Papers describing industrial implementations of MPC with mechanistic models are few, however (Hillestad & Andersen 1994) and (Glemmestad, Ertler & Hillestad 2002) report several applications to industrial polymer reactors. Several simulated examples exist, e.g. (Lee, Lee, Yang & Mahoney 2002), (Prasad, Schley, Russo & Bequette 2002), (Amin, Mehra & Arambel 2001), and (Schei & Singstad 1998), and also some applications to experimental test stands, e.g. (Ahn, Park & Rhee 1999) and (Park, Hur & Rhee 2002).

Process description A simplified drawing of PM6 is shown in Figure 1. Cellulose, TMP (thermomechanical pulp) and broke (repulped fibers and filler from sheet breaks and edge trimmings) are blended in the mixing chest. The stock is fed to the machine chest with a controlled total consistency[1]. Filler is added between the mixing and machine chests. The fillers used in paper production depend on the end-user requirements; typical fillers are kaolin, chalk, talc, and titanium dioxide (Bown 1996). About two thirds of the filler particles used at PM6 are added to the thick stock; the rest is added at the outlet of the white water tank. The flow to the machine chest is large in order to keep the level of the machine chest constant, and an overflow is returned to the mixing chest. The total consistency in the mixing and machine chests are typically around $3-4\%$, which is considerably higher than consistencies later on in the process, and thus the stock from the machine chest is denoted the "thick stock".

The thick stock enters the "short circulation" in the white water tank. Here, the thick stock is diluted to 1-1.5% total consistency by white water[2] and a recirculation flow from the deculator. Filler is added to the stock just after the white water tank. The first cleaning process is a five stage hydrocyclone arrangement, mainly intended to separate heavy particles (e.g. sand and stones) from the flow. The *accept* from the first stage of the hydrocyclones goes to the deculator where air is separated from the stock. The second cleaning process consists of two parallel screens, which separate larger particles (e.g. bark) from the stock. Retention aid is added to the stock at

[1] The total consistency is the weight of solids (i.e. filler particles and fiber) divided by the total weight of solids and water.

[2] White water, which is stored in the white water tank, is the drainage from the wire.

the outlet of the screens. The retention aid is a cationic polymer which, amongst others, adsorb onto anionic fibers and filler particles and cause them to flocculate. The flocculation is a key process for retaining small filler particles and small fiber fragments on the wire, although the significance of mechanical entrapment of non-flocculated filler and fines seems to be somewhat controversial in the literature. For example (Van de Ven 1984) found (theoretically) that mechanical entrapment was low, while (Bown 1996) reports that mechanical entrapment can be a dominant mechanism. In the headbox, the pulp is distributed evenly onto the finely meshed woven wire cloth. Most of the water in the pulp is recirculated to the white water tank, while a share of fiber material and filler particles form a network on the wire which will soon become the paper sheet. The pulp flow from the white water tank, through the hydrocyclones, deculator, screens, headbox, onto the wire and back to the white water tank is denoted the "short circulation".

In the wire section, most of the water is removed by drainage. In the press section, the paper sheet is pressed between rotating steel rolls, thus making use of mechanical forces for water removal. Finally, in the dryer section, the paper sheet passes over rotating and heated cast iron cylinders, and most of the water left in the sheet is removed by evaporation. The paper is then rolled up on the reel before it is moved on to further processing.

Motivation for multivariable model based control Magazine paper is characterized by its glossy appearance due to a high content of filler in the paper. The finished magazine paper typically consists of 65% fiber, 30% filler, and 5% water. The main difference between magazine paper and e.g. newsprint is the high content of filler. For newsprint the amount of filler is typically 0-10%. Due to the high filler content in magazine paper, the couplings between important input and output variables are relatively strong. A project called "Stabilization of the wet end at PM6" was initiated in 1999 based on the experience of strong couplings and oscillating behavior in the process. A key goal was to reduce variation in certain variables, such as consistencies in the short circulation, basis weight, filler content, and more. Based on experience and reported results from competitive mills (e.g. (McQuillin & Huizinga 1995), and (Lang et al. 1998)), it was decided to develop a model of the process and utilize this in a model predictive controller (MPC). Three input and three output variables were selected

$$\bar{u} = \begin{bmatrix} \bar{u}_{TS} \\ \bar{u}_F \\ \bar{u}_{RA} \end{bmatrix}, \ \bar{y} = \begin{bmatrix} \bar{y}_{BW} \\ \bar{y}_{PA} \\ \bar{y}_{WC} \end{bmatrix}, \tag{1}$$

where the inputs $\bar{u}$ are the amount of *thick stock*, *filler* added at the outlet of the white water tank, and *retention aid* added at the outlet of the screens, and where the outputs $\bar{y}$ are the *basis weight* (weight per area), *paper ash* content (content of filler in the paper), and *wire tray consistency* in the recirculation flow from the wire to the white water tank. The basis weight and paper ash outputs are direct quality variables, while the wire tray consistency is an indirect quality variable having significant effect on variation in other short circulation variables.

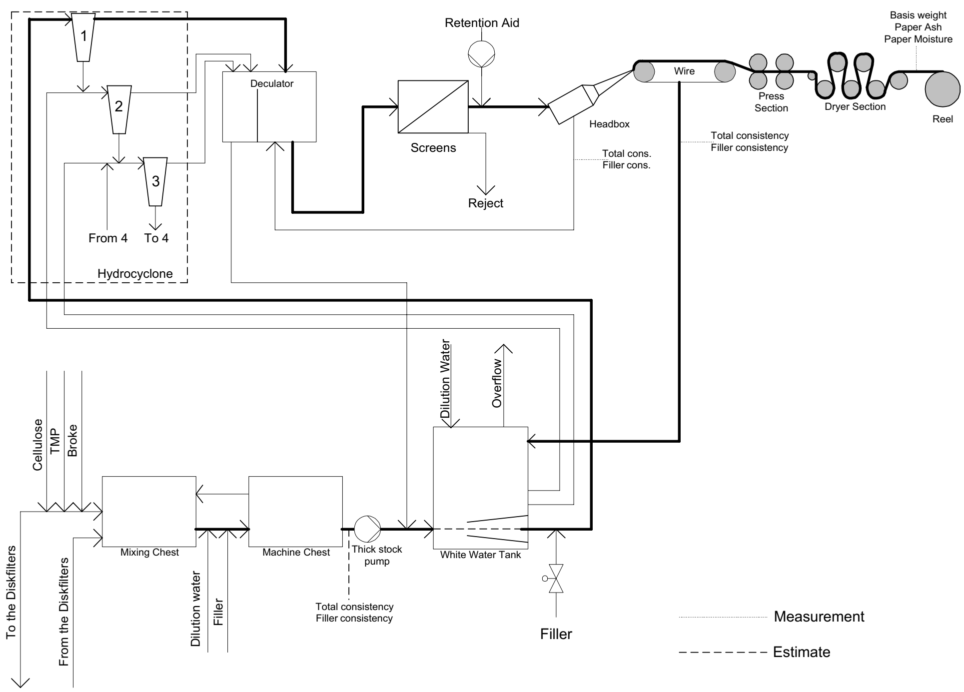

Figure 1: A simplified drawing of PM6. More details are available in (Hauge & Lie 2002).

Before the project started, single loop controllers and manual control were used. Grade changes were carried out manually or partly manually by the operators: the setpoints were changed a number of times before they were equal to the new grade. During start ups, the controllers were kept in manual mode until the measurements were close to the desired specifications. In addition, during sheet breaks the basis weight and paper ash measurements were lost and the inputs controlling these variables were set equal to the values that they had prior to the sheet break. The controllers were kept in manual mode until the paper was back on the reel. Thus, it was also a key goal in the project to be able to have the controllers in automatic mode during grade changes, sheet breaks, and start ups.

The process model The process model is described in detail in e.g. (Hauge & Lie 2002) and only a brief description will be given here. Note that some modifications have been introduced to the model detailed in (Hauge & Lie 2002), as compared to the model implemented at PM6. The most prominent modification is that a first order empiric model that was added to capture neglected and unknown dynamics in the process, has been removed.

The model was originally developed with several ordinary and partial differential equations. The model was then simplified, and eventually fitted to experimental and operational mill data. The implemented PM6 model consists of a third order nonlinear mechanistic model based on physical and chemical laws. The structure of the developed process model is

$$\begin{aligned} \dot{\bar{x}} &= \bar{f}(\bar{x}, \bar{u}, \bar{d}, \bar{\theta}) \\ \bar{y} &= \bar{g}(\bar{x}, \bar{u}, \bar{d}, \bar{\theta}), \end{aligned} \tag{2}$$

with $\bar{x} \in \mathbb{R}^n = \mathbb{R}^3$, $\bar{y} \in \mathbb{R}^m = \mathbb{R}^3$, $\bar{u} \in \mathbb{R}^r = \mathbb{R}^3$ and $\bar{d} \in \mathbb{R}^g = \mathbb{R}^4$. The bar above the variable names indicates that these are the variables in their original units and size. $\bar{\theta}$ consists of several model parameters, tuned to fit the model outputs to experimental and operational data.

The manipulated inputs $\bar{u}$ and the outputs $\bar{y}$ are shown in eq. 1, while the states and measured disturbances are

$$\begin{aligned} \bar{x}^T &= \begin{bmatrix} \bar{C}_{R,fil}, & \bar{C}_{WT,fil}, & \bar{C}_{D,fib} \end{bmatrix} \\ \bar{d}^T &= \begin{bmatrix} \bar{C}_{\mathrm{TS,tot}}, & \bar{C}_{\mathrm{TS,fil}}, & \bar{v}, & \bar{f} \end{bmatrix}, \end{aligned} \tag{3}$$

where $\bar{C}_{R,fil}$ is the concentration of filler in a *reject tank* in the hydrocyclones, $\bar{C}_{WT,fil}$ is the concentration of filler in the *white water tank*, and $\bar{C}_{D,fib}$ is the concentration of fiber in the *deculator*. The measured disturbances accounted for in the model, are the total and filler thick stock concentrations $\bar{C}_{\mathrm{TS,tot}}$ and $\bar{C}_{\mathrm{TS,fil}}$, the paper machine velocity $\bar{v}$, and the paper moisture percentage $\bar{f}$.

Note that the total- and filler concentrations in the thick stock flow are called "measured disturbances", although they are not measured. A model of the thick stock area has been developed (Slora 2001), and implemented at PM6, providing *estimates* of total- and filler concentrations in the thick stock.

Model Predictive Controller (MPC) A commercial MPC developed by Prediktor AS (www.prediktor.no), was chosen by Norske Skog for implementation. The choice of MPC was based on (i) cost, and (ii) the ability to add and develop certain features that were important. Important features were the abilities to

- utilize the non-linear model;
- specify future reference changes. This means that the process operators can specify a setpoint change some time into the future, see how the controller will respond, and let the controller do the grade change if they are satisfied with the response. In many systems, the setpoint is constant into the future, so once a change in setpoint is made, the controller will respond immediately, giving the operators no time to consider how wise the response is;
- make an interface suitable for gaining operator acceptance of the MPC;
- use the MPC during grade changes, sheet breaks, and start ups.

The commercial MPC is part of a software package named Apis (Advanced Process Improvement System), which also consists of a Kalman filter, subspace system identification, and more. The Apis MPC was intended for linear models, based on the infinite horizon objective function presented in (Muske & Rawlings 1993). For the predictive controller implemented at PM6, several extensions were made to the original MPC, such as

- on-line linearization at each sample;
- on-line estimation of key model parameters/biases;
- future setpoint changes, i.e. the process operators can submit new setpoints to the controller some time before the actual grade change;
- addition of a direct input to output term;
- inclusion of measured disturbances.

These extensions will be further discussed in later sections. Note that the extensions discussed in this paper are based on the authors' work, and the actual implementation in Apis may be based on other solutions and ideas. The use of MPC, a nonlinear model, extended Kalman filter, and linearization at each sample, has also been suggested by (Lee & Ricker 1994), although with a finite horizon criterion. Similarly, (Gattu & Zafiriou 1992) proposed an algorithm for nonlinear MPC, with linearization at each sample, but with computation of the steady state Kalman gain at each sample.

Organization of paper In Section 2 an overview of the algorithm for infinite horizon MPC with augmented Kalman filter is given. Linearization of the model is discussed in Section 3, and the MPC equations are outlined in Section 4. In Section 5 the augmented Kalman filter is discussed, and results from using MPC on paper machine 6 (PM6) at Norske Skog Saugbrugs is presented and discussed in Section 6. Finally, some conclusions are given in Section 7.

A few notes about the notation are given in Appendix A. Details on how to find the steady state values are given in Appendix B, and proofs for the finite horizon criteria are given in Appendix C. State and parameter estimation is detailed in Appendix D.

2 Overview of algorithm

At time k we have available the process model (eq. 2) in its discrete version

$$\begin{aligned} \bar{x}_{k+1} &= f(\bar{x}_k, \bar{u}_k, \bar{d}_k, \bar{\theta}_k) \\ \bar{y}_k &= g(\bar{x}_k, \bar{u}_k, \bar{d}_k, \bar{\theta}_k), \end{aligned} \tag{4}$$

as well as the following past measurements and estimates

$$\left.\begin{array}{c} \bar{u}_{k-i} \\ \bar{y}_{k-i} \\ \bar{d}_{k-i} \\ \hat{\bar{x}}_{k-i+1} \end{array}\right\}, \forall i = 1, 2, 3..., \tag{5}$$

where $\hat{\bar{x}}$ is an estimated state vector. The following step by step algorithm for controlling the process is suggested, assuming the present time to be k.

1) Linearization of model based on conditions at time $k-1$ For the chosen MPC we need a linear model. The linearization is based on the most up-to-date information about the process, i.e. the variable values at time $k-1$. Note that we have no information about $\bar{u}_k$ yet, so we can not linearize based on variable values at time k. The resulting model is

$$\begin{aligned} \bar{\bar{x}}_{k+1} &= A_k \bar{\bar{x}}_k + B_k \bar{\bar{u}}_k + E_k \bar{\bar{d}}_k \\ \bar{\bar{y}}_k &= C_k \bar{\bar{x}}_k + D_k \bar{\bar{u}}_k + F_k \bar{\bar{d}}_k. \end{aligned} \tag{6}$$

The linearization is discussed more thoroughly in Section 3. See also notes about notation in Appendix A.

2) Obtain current measured disturbances and future setpoints and disturbances The measured disturbances obtained from the process are $\bar{d}_k$. The future disturbances and references are

$$\begin{aligned} &\bar{r}_{k+j},\ j = 0, ..., N-1 \\ &\bar{d}_{k+j},\ j = 0, ..., N-1, \end{aligned} \tag{7}$$

which must be provided by the process operators or simply taken as an extention of the present values into the future.

3) Shift variables corresponding to the linearized model The references, disturbances, and constraints will be used with the linearized model in eq. 6 for calculation of target values. The references, disturbances, and constraints must then be shifted along with the model so that all variables have the same origin before the calculation of target values.

4) Calculate steady state values The calculation of steady state values is carried out for several reasons. The steady state values are used as targets in the optimization criterion. One could use e.g. reference values directly as targets in the criterion. However, the calculation of steady state values is a way of ensuring that the targets are feasible. In addition, by calculating steady state values one has the opportunity to add e.g. an economic type criterion if there are additional degrees of freedom. Finally, for the special case of an infinite horizon criterion with possibility of changing future references and measured disturbances, we need the steady state values at the end of the horizon in order to shift the origin of the model.

5) Shift the origin of the model to the steady state values at time $k+N-1$ This is a step taken in order to reformulate the criterion to a finite horizon criterion.

6) Shift measured and estimated variables The variables must be shifted along with the model so that they have the same origin.

7) Update MPC matrices and vectors The matrices and vectors in the MPC formulation that contain time variant variables, such as linear model matrices, input variables, estimated states, etc., must be updated.

8) Optimization An optimization algorithm is used to calculate optimal inputs.

9) Apply $\bar{u}_k$ to the process Note that only the first input is used.

10) Obtain $\bar{y}_k$ from the process

11) Estimate $\widehat{\bar{x}}_{k+1}$ Unless all states are measured, we need to estimate them (or some of them). Typically an extended Kalman filter is used for this purpose.

12) Set $k := k+1$ and go to step 1

3 Linearization of model

Consider perturbations $\bar{\bar{x}}_k$, $\bar{\bar{u}}_k$, and $\bar{\bar{d}}_k$ near the variable values at time $k-1$

$$\begin{aligned}\bar{x}_k &= \bar{x}_{k-1} + \bar{\bar{x}}_k \\ \bar{u}_k &= \bar{u}_{k-1} + \bar{\bar{u}}_k \\ \bar{d}_k &= \bar{d}_{k-1} + \bar{\bar{d}}_k.\end{aligned} \tag{8}$$

Inserting these perturbations into eq. 4, and neglecting the parameter vector $\bar{\theta}_k$, gives

$$\begin{aligned}\bar{x}_{k+1} &= f(\bar{x}_k, \bar{u}_k, \bar{d}_k) = f(\bar{x}_{k-1} + \bar{\bar{x}}_k, \bar{u}_{k-1} + \bar{\bar{u}}_k, \bar{d}_{k-1} + \bar{\bar{d}}_k) \\ \bar{y}_k &= g(\bar{x}_k, \bar{u}_k, \bar{d}_k) = g(\bar{x}_{k-1} + \bar{\bar{x}}_k, \bar{u}_{k-1} + \bar{\bar{u}}_k, \bar{d}_{k-1} + \bar{\bar{d}}_k).\end{aligned} \tag{9}$$

Define $s = \{\bar{x}_{k-1}, \bar{u}_{k-1}, \bar{d}_{k-1}\}$. A first order expansion, with center corresponding to s, then gives

$$\begin{aligned}\bar{x}_{k+1} &\approx f(\bar{x}_{k-1}, \bar{u}_{k-1}, \bar{d}_{k-1}) + \frac{\partial f}{\partial \bar{x}_k}|_s(\bar{x}_{k-1} + \bar{\bar{x}}_k - \bar{x}_{k-1}) \\ &+ \frac{\partial f}{\partial \bar{u}_k}|_s(\bar{u}_{k-1} + \bar{\bar{u}}_k - \bar{u}_{k-1}) + \frac{\partial f}{\partial \bar{d}_k}|_s(\bar{d}_{k-1} + \bar{\bar{d}}_k - \bar{d}_{k-1}) \\ \bar{y}_k &\approx g(\bar{x}_{k-1}, \bar{u}_{k-1}, \bar{d}_{k-1}) + \frac{\partial g}{\partial \bar{x}_k}|_s(\bar{x}_{k-1} + \bar{\bar{x}}_k - \bar{x}_{k-1}) \\ &+ \frac{\partial g}{\partial \bar{u}_k}|_s(\bar{u}_{k-1} + \bar{\bar{u}}_k - \bar{u}_{k-1}) + \frac{\partial g}{\partial \bar{d}_k}|_s(\bar{d}_{k-1} + \bar{\bar{d}}_k - \bar{d}_{k-1})\end{aligned} \tag{10}$$

Defining

$$\begin{aligned}A_k &= \frac{\partial f}{\partial \bar{x}_k}|_s,\ B_k = \frac{\partial f}{\partial \bar{u}_k}|_s,\ E_k = \frac{\partial f}{\partial \bar{d}_k}|_s \\ C_k &= \frac{\partial g}{\partial \bar{x}_k}|_s,\ D_k = \frac{\partial g}{\partial \bar{u}_k}|_s,\ F_k = \frac{\partial g}{\partial \bar{d}_k}|_s \\ \bar{y}_k &= g(\bar{x}_{k-1}, \bar{u}_{k-1}, \bar{d}_{k-1}) + \bar{\bar{y}}_k\end{aligned} \tag{11}$$

and inserting the definitions in eq. 8 into eq. 10 gives

$$\begin{aligned}\bar{\bar{x}}_{k+1} &\approx A_k\bar{\bar{x}}_k + B_k\bar{\bar{u}}_k + E_k\bar{\bar{d}}_k \\ \bar{\bar{y}}_k &\approx C_k\bar{\bar{x}}_k + D_k\bar{\bar{u}}_k + F_k\bar{\bar{d}}_k,\end{aligned} \tag{12}$$

where we have assumed that $\bar{x}_k = f(\bar{x}_{k-1}, \bar{u}_{k-1}, \bar{d}_{k-1})$, in accordance with the original nonlinear model equation. In the remainder of this paper, eq. 12 will be used with equality sign ("=") instead of approximation ("≈").

Note that the linearization was carried out with center corresponding to variable values at time $k-1$. If the linearization is carried out before the computation of the optimal input $\bar{u}_k$, then the linearized model must have center corresponding to variable values at time $k-1$. If the input $\bar{u}_k$ is available at the time of linearization, the center can correspond to variable values at time k, however by the time the linearized model is used in the MPC the time is $k+1$. Thus, it is not important whether the linearization is carried out prior to, or after, the computation of optimal inputs since the linearized model will be centered on the variable values at the previous sample in any case.

The linear system matrices $A_k, B_k, ..., F_k$ can be found by analytic or numeric methods. These methods are presented next.

3.1 Analytic linearization

Analytic linearization is carried out e.g. by hand, automatic differentiation, see e.g. (Griewank 2000), (Griewank & Corliss 1991), and (Solberg 1988), or by software capable of symbolic computation, e.g. Maple, or Matlab with the symbolic toolbox. For small and not too complicated systems this is a convenient method. Consider e.g. matrices A_k and B_k, computed element by element according to

$$A_k = \frac{\partial f}{\partial \bar{x}_k}|_s = \begin{bmatrix} \frac{\partial f_1}{\partial \bar{x}_{k,1}}|_s & \frac{\partial f_1}{\partial \bar{x}_{k,2}}|_s & \cdots & \frac{\partial f_1}{\partial \bar{x}_{k,n}}|_s \\ \frac{\partial f_2}{\partial \bar{x}_{k,1}}|_s & \frac{\partial f_2}{\partial \bar{x}_{k,2}}|_s & \cdots & \frac{\partial f_2}{\partial \bar{x}_{k,n}}|_s \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial \bar{x}_{k,1}}|_s & \frac{\partial f_n}{\partial \bar{x}_{k,2}}|_s & \cdots & \frac{\partial f_n}{\partial \bar{x}_{k,n}}|_s \end{bmatrix}$$

$$B_k = \frac{\partial f}{\partial \bar{u}_k}|_s = \begin{bmatrix} \frac{\partial f_1}{\partial \bar{u}_{k,1}}|_s & \frac{\partial f_1}{\partial \bar{u}_{k,2}}|_s & \cdots & \frac{\partial f_1}{\partial \bar{u}_{k,r}}|_s \\ \frac{\partial f_2}{\partial \bar{u}_{k,1}}|_s & \frac{\partial f_2}{\partial \bar{u}_{k,2}}|_s & \cdots & \frac{\partial f_2}{\partial \bar{u}_{k,r}}|_s \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial \bar{u}_{k,1}}|_s & \frac{\partial f_n}{\partial \bar{u}_{k,2}}|_s & \cdots & \frac{\partial f_n}{\partial \bar{u}_{k,r}}|_s \end{bmatrix},$$

where $s = \bar{x}_{k-1}, \bar{u}_{k-1}, \bar{d}_{k-1}$ is the center of the linearization, $\bar{x}_{k,i}$ means the i'th state variable at time k in the nonlinear model, and similar for other variables. The other system matrices are computed similar to this. Note that the B matrix consists of n rows and r columns and is not in general a square matrix.

3.2 Numeric linearization

Numeric linearization is carried out by perturbing the variables and thus finding the derivatives in the system matrices, see e.g. (Dennis & Schnabel 1996) and (Gill, Murray & Wright 1981). Assuming $\bar{x}_{k-1}$, $\bar{u}_{k-1}$ and $\bar{d}_{k-1}$ are available, one would typically use an algorithm similar to the following:

1. Using the nonlinear model and known variables, compute $\bar{x}_k$ and $\bar{y}_{k-1}$.

2. For every state variable:

   (a) Perturb state variable $\bar{x}_{k-1,i}$, by adding a small number $\Delta\bar{x}_{k-1,i}$ to its value. For example one may increase its value by one percent, or a predetermined minimum perturbation (e.g. if the variable value is zero we can not use the one percent rule).
   (b) Using the nonlinear model, compute $\bar{x}_k^{\text{pert}}$ and $\bar{y}_{k-1}^{\text{pert}}$.
   (c) The i'th column in matrix A_k is then

$$\frac{\bar{x}_k^{\text{pert}} - \bar{x}_k}{\Delta\bar{x}_{k-1,i}}$$

   and the i'th column in matrix C_k is

$$\frac{\bar{y}_{k-1}^{\text{pert}} - \bar{y}_{k-1}}{\Delta\bar{x}_{k-1,i}}$$

3. For every input variable:

   (a) Perturb input variable $\bar{u}_{k-1,i}$, by adding a small number $\Delta\bar{u}_{k-1,i}$ to its value.
   (b) Using the nonlinear model, compute $\bar{x}_k^{\text{pert}}$ and $\bar{y}_{k-1}^{\text{pert}}$.
   (c) The i'th column in matrix B_k is then

$$\frac{\bar{x}_k^{\text{pert}} - \bar{x}_k}{\Delta\bar{u}_{k-1,i}}$$

   and the i'th column in matrix D_k is

$$\frac{\bar{y}_{k-1}^{\text{pert}} - \bar{y}_{k-1}}{\Delta\bar{u}_{k-1,i}}$$

4. For every measured disturbance variable:

   (a) Perturb measured disturbance variable $\bar{d}_{k-1,i}$, by adding a small number $\Delta\bar{d}_{k-1,i}$ to its value.
   (b) Using the nonlinear model, compute $\bar{x}_k^{\text{pert}}$ and $\bar{y}_{k-1}^{\text{pert}}$.
   (c) The i'th column in matrix E_k is then

$$\frac{\bar{x}_k^{\text{pert}} - \bar{x}_k}{\Delta\bar{d}_{k-1,i}}$$

   and the i'th column in matrix F_k is

$$\frac{\bar{y}_{k-1}^{\text{pert}} - \bar{y}_{k-1}}{\Delta\bar{d}_{k-1,i}}$$

3.3 Example: Linearization of a chemical reactor model

A model of a chemical reactor (Lie 1995) is

$$\begin{aligned}\dot{C}_A &= \frac{q}{V}\cdot(C_{Ai}-C_A) - k_0\cdot e^{-\frac{E}{R\cdot T}}\cdot C_A \\ \dot{T} &= \frac{q}{V}\cdot(T_i - T) + \frac{1}{\bar{C}_V}\cdot\left(-\Delta U_r\cdot k_0\cdot e^{-\frac{E}{R\cdot T}}\cdot C_A\cdot V + Q\right)\end{aligned} \tag{13}$$

The A matrix in the linearized model is then

$$\begin{aligned}A &= \begin{bmatrix} \frac{\partial \dot{C}_A}{\partial C_A} & \frac{\partial \dot{C}_A}{\partial T} \\ \frac{\partial \dot{T}}{\partial C_A} & \frac{\partial \dot{T}}{\partial T} \end{bmatrix} \\ &= \begin{bmatrix} -\frac{q}{V} - k_0 e^{-\frac{E}{RT}}, & -\frac{E}{RT^2}k_0 C_A e^{-\frac{E}{RT}} \\ \frac{1}{\bar{C}_V}\cdot -\Delta U_r\cdot k_0\cdot e^{-\frac{E}{R\cdot T}}\cdot V, & -\frac{q}{V} - \frac{1}{\bar{C}_V}\cdot\Delta U_r\cdot k_0\cdot\frac{E}{RT^2}\cdot e^{-\frac{E}{R\cdot T}}\cdot C_A\cdot V \end{bmatrix}\end{aligned}$$

With appropriate parameter values and operating point, as given in (Lie 1995), we have

$$A_{\text{analytic}} = \begin{bmatrix} -19.998 & -4.6209\cdot 10^{-2} \\ 3824.4 & 8.3018 \end{bmatrix}$$

Numeric linearization, with a perturbation according to 1% of the state values, gives the following A matrix

$$A_{\text{numeric}} = \begin{bmatrix} -19.998 & -5.1131\cdot 10^{-2} \\ 3824.4 & 9.2926 \end{bmatrix}$$

which is close to the analytic A matrix. The reactor is highly nonlinear and we try with a smaller perturbation corresponding to 0.1% of the state values. This gives

$$A_{\text{numeric},0.1\%} = \begin{bmatrix} -19.998 & -4.6675\cdot 10^{-2} \\ 3824.4 & 8.3956 \end{bmatrix}$$

which is seen to be even closer to the analytic solution.

This example has shown the possibilities of analytic and numeric linearization, as well as the difficulty of choosing a proper perturbation for numeric linearization.

4 Model predictive controller (MPC)

Commercial MPC algorithms often consists of two stages (Qin & Badgwell 1997), first steady state values (or target values) are calculated, and then these values are used as targets in the calculation of the optimal input values. The calculation of steady state values is a way of ensuring that the targets are feasible. In addition, by calculating steady state values one has the opportunity to add e.g. an economic type criterion if there are additional degrees of freedom. Finally, for the special case of an infinite horizon criterion with the possibility of changing future references and measured disturbances, we need the steady state values at the end of the horizon in order to shift the origin of the model. The model origin is shifted so that the variables in the criterion converges exponentially to a zero steady state, thus avoiding an infinite value of the criterion.

4.1 Steady state values

We assume that a linearized model (eq. 12) and the adjusted reference vector $\bar{\bar{r}}_{k+j}$ and disturbance vector $\bar{\bar{d}}_{k+j}$ are available (at time k):

$$\begin{aligned} \bar{\bar{r}}_{k+j} &= \bar{r}_{k+j} - \bar{y}_{k-1},\ j = 0, ..., N-1 \\ \bar{\bar{d}}_{k+j} &= \bar{d}_{k+j} - \bar{d}_{k-1},\ j = 0, ..., N-1, \end{aligned} \tag{14}$$

The future reference and disturbance vectors are provided by the process operators or simply taken as an extention of the present values into the future. N is a chosen control horizon where we allow input changes.

At each future time sample we calculate target values for the states and inputs of the process. The target values may be calculated using an economic criterion, or calculated as e.g.

$$\min_{\bar{x}^s_{k+j}, \bar{u}^s_{k+j}} \left(\bar{y}^s_{k+j} - \bar{\bar{r}}_{k+j}\right)^T Q_s \left(\bar{y}^s_{k+j} - \bar{\bar{r}}_{k+j}\right),\ \forall j = 0, ..., N-1, \tag{15}$$

constrained by the steady state solution of the model

$$\begin{bmatrix} (I - A_k) & -B_k \\ -C_k & -D_k \end{bmatrix} \begin{bmatrix} \bar{x}^s_{k+j} \\ \bar{u}^s_{k+j} \end{bmatrix} = \begin{bmatrix} E_k \bar{\bar{d}}_{k+j} \\ F_k \bar{\bar{d}}_{k+j} - \bar{y}^s_{k+j} \end{bmatrix},\ \forall j = 0, ..., N-1, \tag{16}$$

with bounds

$$\begin{aligned} \bar{u}^{\min}_{k+j} \le \bar{u}^s_{k+j} + \bar{u}_{k-1} \le \bar{u}^{\max}_{k+j},\ \forall j = 0, ..., N-1 \\ \bar{y}^{\min}_{k+j} \le \bar{y}^s_{k+j} + \bar{y}_{k-1} \le \bar{y}^{\max}_{k+j},\ \forall j = 0, ..., N-1, \end{aligned} \tag{17}$$

where $\bar{y}^{\min}_{k+j}$, $\bar{y}^{\max}_{k+j}$, $\bar{u}^{\min}_{k+j}$, and $\bar{u}^{\max}_{k+j}$ are minimum and maximum values corresponding to the original nonlinear model. If there are additional degrees of freedom we may specify an economic type criterion instead of eq. 15, and use eqs. 16-17 as constraints. In Appendix B it is shown how one may use the `lssol` algorithm (Gill, Hammarling, Murray, Saunders & Wright 1986) for calculating the steady state values.

The origin of the model is then shifted to the steady state values at time $N-1$

$$\begin{aligned} \left(\bar{\bar{x}}_{k+1} - \bar{x}^s_{k+N-1}\right) &= A_k \left(\bar{\bar{x}}_k - \bar{x}^s_{k+N-1}\right) + B_k \left(\bar{\bar{u}}_k - \bar{u}^s_{k+N-1}\right) + E_k \left(\bar{\bar{d}}_k - \bar{\bar{d}}_{k+N-1}\right) \\ \left(\bar{\bar{y}}_k - \bar{y}^s_{k+N-1}\right) &= C_k \left(\bar{\bar{x}}_k - \bar{x}^s_{k+N-1}\right) + D_k \left(\bar{\bar{u}}_k - \bar{u}^s_{k+N-1}\right) + F_k \left(\bar{\bar{d}}_k - \bar{\bar{d}}_{k+N-1}\right), \end{aligned} \tag{18}$$

which gives the shifted model

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + E_k d_k \\ y_k &= C_k x_k + D_k u_k + F_k d_k, \end{aligned} \tag{19}$$

where

$$\begin{aligned} x_{k+1} &= \bar{\bar{x}}_{k+1} - \bar{x}^s_{k+N-1} \\ x_k &= \bar{\bar{x}}_k - \bar{x}^s_{k+N-1} \\ u_k &= \bar{\bar{u}}_k - \bar{u}^s_{k+N-1} \\ d_k &= \bar{\bar{d}}_k - \bar{\bar{d}}_{k+N-1} \\ y_k &= \bar{\bar{y}}_k - \bar{y}^s_{k+N-1} \end{aligned} \tag{20}$$

The shifting of model origin to the steady state values at time $N-1$ makes the variables in the criterion converge exponentially to zero in steady state, thus ensuring a finite value of the criterion.

4.1.1 Shifting variables

The model origin is shifted twice, once during the linearization and once after the computation of steady state (target) values. The measured, estimated and calculated variables must be shifted along with the model as follows

$$\begin{aligned} \hat{x}_k &= \hat{\bar{\bar{x}}}_k - \hat{\bar{\bar{x}}}_{k-1} - \bar{x}^s_{k+N-1} \\ u_{k-1} &= \bar{u}_{k-1} - \bar{u}_{k-1} - \bar{u}^s_{k+N-1} = -\bar{u}^s_{k+N-1} \\ d_{k+i} &= \bar{d}_{k+i} - \bar{d}_{k-1} - \bar{\bar{d}}_{k+N-1},\ i = 0, ..., N-1 \\ u^s_{k+i} &= \bar{u}^s_{k+i} - \bar{u}^s_{k+N-1},\ i = 0, ..., N-1 \\ y^s_{k+i} &= \bar{y}^s_{k+i} - \bar{y}^s_{k+N-1},\ i = 0, ..., N-1 \end{aligned} \tag{21}$$

4.2 Optimal input values

This section is based on the algorithm presented in (Muske & Rawlings 1993), although several extensions are made, notably the inclusion of future reference and disturbance trajectories.

The infinite horizon criterion is

$$\min_{\mathcal{U}_k} J_k = \min_{\mathcal{U}_k} \sum_{j=0}^{\infty} \left[e^T_{k+j} Q e_{k+j} + \tilde{u}^T_{k+j} R \tilde{u}_{k+j} + \Delta u^T_{k+j} S \Delta u_{k+j} \right], \tag{22}$$

constrained by the model in eq. 19, i.e.

$$\begin{aligned} x_{k+1+j} &= A x_{k+j} + B u_{k+j} + E d_{k+j},\ \forall j = 0, 1, 2, ... \\ y_{k+j} &= C x_{k+j} + D u_{k+j} + F d_{k+j},\ \forall j = 0, 1, 2, ..., \end{aligned} \tag{23}$$

and the following inequality constraints

$$\begin{aligned} \bar{u}_{k+j}^{\min} \leq \bar{u}_{k+j} \leq \bar{u}_{k+j}^{\max},\ j = 0, 1, ..., N-1 \\ \bar{y}_{k+j}^{\min} \leq \bar{y}_{k+j} \leq \bar{y}_{k+j}^{\max},\ j = j_1, j_1+1, ..., j_2 \\ \Delta u_{k+j}^{\min} \leq \Delta u_{k+j} \leq \Delta u_{k+j}^{\max},\ j = 0, 1, ..., N, \end{aligned} \quad (24)$$

and where

$$\begin{aligned} e_{k+j} &= y_{k+j} - y_{k+j}^{s} \\ \tilde{u}_{k+j} &= u_{k+j} - u_{k+j}^{s} \\ \Delta u_{k+j} &= u_{k+j} - u_{k+j-1} \\ \mathcal{U}_k &= \left[u_k^T, u_{k+1}^T, ..., u_{k+N-1}^T\right]^T. \end{aligned} \quad (25)$$

The output constraints are active from sample $k+j_1$ to $k+j_2$. j_1 should be chosen so that feasibility is ensured from $k+j_1$, and j_2 should be chosen such that feasibility up to $k+j_2$ implies feasibility on the infinite horizon. Bounds for j_1 and j_2, so that feasibility is guaranteed, are developed in (Rawlings & Muske 1993).

Consider the Jordan decomposition of A_k

$$A_k = V_k J_k V_k^{-1} = \begin{bmatrix} V_k^u & V_k^s \end{bmatrix} \begin{bmatrix} J_k^u & 0 \\ 0 & J_k^s \end{bmatrix} \begin{bmatrix} \tilde{V}_k^u \\ \tilde{V}_k^s \end{bmatrix}, \quad (26)$$

where V_k^u and J_k^u are respectively the eigenvectors and Jordan blocks for the eigenvalues corresponding to the unstable modes of A_k, and V_k^s and J_k^s are respectively the eigenvectors and Jordan blocks for the eigenvalues corresponding to the stable modes of A_k. The following results can then be obtained.

Theorem 1 *Consider the model given by eq. 19, the criterion of eq. 22, and the definitions provided by eq. 25. Assume that*

$$u_{k+j} = 0, \forall j \in \{N, N+1, ...\}, \text{ which is equivalent to } \bar{\bar{u}}_{k+j} = \bar{u}_{k+N-1}^{s}, \forall j \in \{N, N+1, ...\} \quad (27)$$

$$\begin{aligned} d_{k+j} = 0, \forall j \in \{N, N+1, ...\}, &\text{ which is equivalent to } \bar{\bar{d}}_{k+j} = \bar{\bar{d}}_{k+N-1}, \forall j \in \{N, N+1, ...\} \\ u_{k+j}^{s} = 0, \forall j \in \{N, N+1, ...\}, &\text{ which is equivalent to } \bar{u}_{k+j}^{s} = \bar{u}_{k+N-1}^{s}, \forall j \in \{N, N+1, ...\} \\ y_{k+j}^{s} = 0, \forall j \in \{N, N+1, ...\}, &\text{ which is equivalent to } \bar{y}_{k+j}^{s} = \bar{y}_{k+N-1}^{s}, \forall j \in \{N, N+1, ...\}, \end{aligned}$$

thus there are no changes in the inputs, measured disturbances, or steady state inputs and outputs, from sample N and forward. If in addition we add the equality constraint (ref. eq. 26)

$$\tilde{V}_k^u x_{k+N} = 0, \quad (28)$$

which corresponds to bringing the unstable modes to zero at time $k+N$, then the *infinite horizon criterion can be written as the following finite horizon criterion*

$$\min_{\mathcal{U}_k} J_k = \min_{\mathcal{U}_k} \left(x_{k+N}^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s x_{k+N} + \Delta u_{k+N}^T S \Delta u_{k+N} \right. \\ \left. + \sum_{j=0}^{N-1} \left[e_{k+j}^T Q e_{k+j} + \tilde{u}_{k+j}^T R \tilde{u}_{k+j} + \Delta u_{k+j}^T S \Delta u_{k+j} \right] \right), \tag{29}$$

with $\bar{Q}_k$ *given by the discrete Lyapunov equation*

$$\bar{Q}_k = \left(V_k^s\right)^T C_k^T Q C_k V_k^s + \left(J_k^s\right)^T \bar{Q}_k J_k^s, \tag{30}$$

Proof. See Appendix C.1. ∎

Proposition 2 *Consider the model given by eq. 19, the criterion of eq. 22, the inequality constraints given by eq. 24, the equality constraint given in eq. 28, and the definitions and assumptions provided by eqs. 25 and 27. This minimization problem can be formulated as the following standard constrained QP (Quadratic Programming) problem*

$$\min_{\mathcal{U}_k} J_k = \min_{\mathcal{U}_k} \left(\frac{1}{2} \mathcal{U}_k^T H_k \mathcal{U}_k + c_k^T \mathcal{U}_k \right), \tag{31}$$

subject to

$$b_{L,k} \leq \begin{bmatrix} \mathcal{U}_k \\ \bar{A}_k \mathcal{U}_k \end{bmatrix} \leq b_{U,k} \tag{32}$$

where

$$H_k = 2\left((P^u)^T (C_k^u)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s C_k^u P^u + \mathcal{P}^T (I_{N+1} \otimes S) \mathcal{P} + (\mathcal{H}_k^u)^T (I_N \otimes Q) \mathcal{H}_k^u + (I_N \otimes R)\right) \quad (33)$$

$$c_k^T = 2\left(x_k^T \left(A_k^N\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s C_k^u P^u + \mathcal{D}_k^T \left(P^d\right)^T \left(C_k^d\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s C_k^u P^u + u_{k-1}^T \mathcal{L}^T (I_{N+1} \otimes S) \mathcal{P} + x_k^T \mathcal{O}_k^T (I_N \otimes Q) \mathcal{H}_k^u + \mathcal{D}_k^T \left(\mathcal{H}_k^d\right)^T (I_N \otimes Q) \mathcal{H}_k^u - (Y_k^s)^T (I_N \otimes Q) \mathcal{H}_k^u - (\mathcal{U}_k^s)^T (I_N \otimes R)\right)$$

$$\bar{A}_k = \begin{bmatrix} \mathcal{H}_k^u \\ \mathcal{P} \\ \tilde{V}_k^u C_k^u P^u \end{bmatrix}$$

$$b_{L,k} = \begin{bmatrix} \mathcal{U}_k^{min} - \bar{u}_{k-1} \otimes \mathbf{1}_N - \bar{u}_{k+N-1}^s \otimes \mathbf{1}_N \\ \mathcal{Y}_k^{min} - \bar{y}_{k-1} \otimes \mathbf{1}_N - \bar{y}_{k+N-1}^s \otimes \mathbf{1}_N - \mathcal{O}_k x_k - \mathcal{H}_k^d \mathcal{D}_k \\ \Delta\mathcal{U}_k^{min} - \mathcal{L} u_{k-1} \\ -\tilde{V}_k^u \left(A_k^N x_k + C_k^d P^d \mathcal{D}_k\right) \end{bmatrix}$$

$$b_{U,k} = \begin{bmatrix} \mathcal{U}_k^{max} - \bar{u}_{k-1} \otimes \mathbf{1}_N - \bar{u}_{k+N-1}^s \otimes \mathbf{1}_N \\ \mathcal{Y}_k^{max} - \bar{y}_{k-1} \otimes \mathbf{1}_N - \bar{y}_{k+N-1}^s \otimes \mathbf{1}_N - \mathcal{O}_k x_k - \mathcal{H}_k^d \mathcal{D}_k \\ \Delta\mathcal{U}_k^{max} - \mathcal{L} u_{k-1} \\ -\tilde{V}_k^u \left(A_k^N x_k + C_k^d P^d \mathcal{D}_k\right) \end{bmatrix}$$

and definitions of $\mathcal{P}$ *,*$\mathcal{H}_k^u$*,* $\mathcal{D}_k$*,* $\mathcal{H}_k^d$*,* $\mathcal{O}_k$*, etc. are provided in Appendix C.2.* $\otimes$ *denotes the Kronecker product.*

Proof. See Appendices C.1-C.2. ■

A possible choice of QP solver is `sqopt` which solves problems in the form of eqs. 31–32. The `sqopt` algorithm for solving constrained linear and quadratic problems (Gill, Murray & Saunders 1997), is available with a Matlab interface in the Tomlab environment (Holmström 2001). Other formulations, choice of QP solvers, and variables are investigated more thoroughly in (Lie, Dueñas Díez & Hauge 2002).

5 Estimating the states and parameters

Using a state space model in an MPC application, as in the previous section, requires estimation of the states unless all states are measured. A Kalman filter is used at PM6 for estimating the states in the paper machine model. The Kalman filter equations for a linear time variant process are derived in Appendix D.1. The paper machine model is nonlinear and thus an extended Kalman filter is used for estimating the states in the model. An algorithm for estimating the states in a nonlinear model is reviewed in Appendix D.2.

Due to disturbances and model errors, the MPC presented in previous sections is likely to exhibit steady state offset from the setpoints. The most common way to handle this problem is to assume a step disturbance at the model outputs and estimate the size of the step in a deadbeat fashion (Qin & Badgwell 1997), (Muske & Rawlings 1993). Other methods also exist, such as assuming the disturbance to originate from the process inputs (Muske & Rawlings 1993). In (Muske & Badgwell 2002) various disturbance models which provide offset-free control are discussed, and conditions for offset-free control are developed. In Appendix D.3 we have shown how the MPC and Kalman filter can be redesigned to prevent steady state offset by estimating the bias and adding this to the model outputs. Although this is the most commonly used method for removing steady state offset, it is often a poor method for solving the problem, notably if the disturbances enters the inputs or states (Muske & Rawlings 1993), (Muske & Badgwell 2002). The main point is that offset-free control can be obtained with many different disturbance models, however to obtain best possible performance the disturbance should be included in the model where it enters in the real process.

The question of where the disturbances enter in a real process is easy to answer: everywhere. As pointed out in (Muske & Badgwell 2002), only a limited number of biases or parameters can be estimated on-line, thus the choice of which parameters or biases to estimate must be based on experience with the process and model. Three biases have been selected for on-line estimation in the paper machine model. The first two are biases on the estimated total- and filler thick stock consistencies (see eq. 3). These disturbances are estimated using a ballistic estimator (Slora 2001), and thus they are assumed to be good candidates for having time-varying biases. The third bias estimated on-line is for the total wire tray concentration, i.e. a bias in one of the outputs. In Appendix D.4 we have shown how arbitrary parameters and biases in the model can be estimated on-line by an augmented Kalman filter. It is also shown how the linearization, calculation of steady state values, and optimization may be carried out on the augmented system.

5.1 Tuning and validation

In theory, and in the true Kalman filter, the noise characteristics of the process should be found and used in the Kalman filter equations. However, these characteristics are hard, if not impossible, to find. Thus, one often aims for a suboptimal Kalman filter, where the noise characteristics are used as tuning parameters until satisfactory Kalman filter performance is obtained. Specifically, the tuning parameters are the process noise covariance matrix Q_k and the measurement noise covariance matrix R_k. In the augmented Kalman filter (as described in section D.4), the augmented process noise covariance matrix $\tilde{Q}_k$ is used. Often, it is assumed that only diagonal elements are non-zero. Thus, for the paper machine model there are three diagonal elements in R_k (three outputs), and six diagonal elements in $\tilde{Q}_k$ (three states plus three estimated parameters). The first element (upper left corner) in R_k corresponds to the variance of the basis weight measurement, the second element (the element on the second row and second column) in R_k corresponds to the variance of the paper ash measurement, etc. Similarly for the diagonal elements in $\tilde{Q}_k$, the first diagonal element corresponds to the variance of the first state variable (the concentration of filler in the reject tank), and e.g. the last element on the diagonal corresponds to the variance of the last parameter to be estimated (bias in the wire tray total concentration).

When tuning and validating the (suboptimal) Kalman filter, we have used a few facts and rules of thumb, e.g.:

- The tuning and validation (with different data sets) should aim at good tracking properties (i.e. the estimated outputs should follow the measured outputs to some extent), good filtering properties (i.e. the estimated outputs should not track measurement noise), and a sound balance between the updating of states and updating of parameters (e.g. the parameters should not vary a lot while the states are more or less resting).

- It can be shown that it is the ratio of the various variances that determines the performance of the Kalman filter. Thus, one need not be careful about finding realistic variance values.

- It is possible to estimate the variances, using a parameter estimation method. This is done for a constant gain Kalman filter (i.e. the individual variances are not estimated, but the Kalman filter gain matrix is estimated) in (Hauge & Lie 2002). The drawback with this method is that the Kalman filter will be rather aggressive, and some de-tuning procedure is needed (but it may give a good starting point).

- Start the tuning by finding approximate values for the various variances. The measurement variances can be approximately found by visually studying the random variations in the measurements. It is harder to find suitable starting values for the process noise variances and the parameter estimate variances. However, the expected state and parameter values will give good indications of reasonable starting values. Consider e.g. a concentration that is expected to have a value around 0.05 (5%). If one assumes that the noise entering this state is approximately 1% of the state value, we see that the variance will be a very small number. In the Kalman filter used at PM6, the measurement variances are much larger than the process and parameter variances (around 10^8 larger).

- In general, increasing the measurement variances leads to a slower updating of state estimates. The same result is obtained by decreasing the process variance. Thus, decreasing the process variance leads to a slower updating of state estimates.

- Since the parameters are augmented states, changing the parameter variances has much of the same effect as changing the state variances. Increasing the parameter variances leads to a faster updating of parameter estimates, thus also leading to a faster elimination of estimation error (the difference between estimated outputs and measured outputs).

Validation results for the augmented Kalman filter are shown in Figures 2-4 .

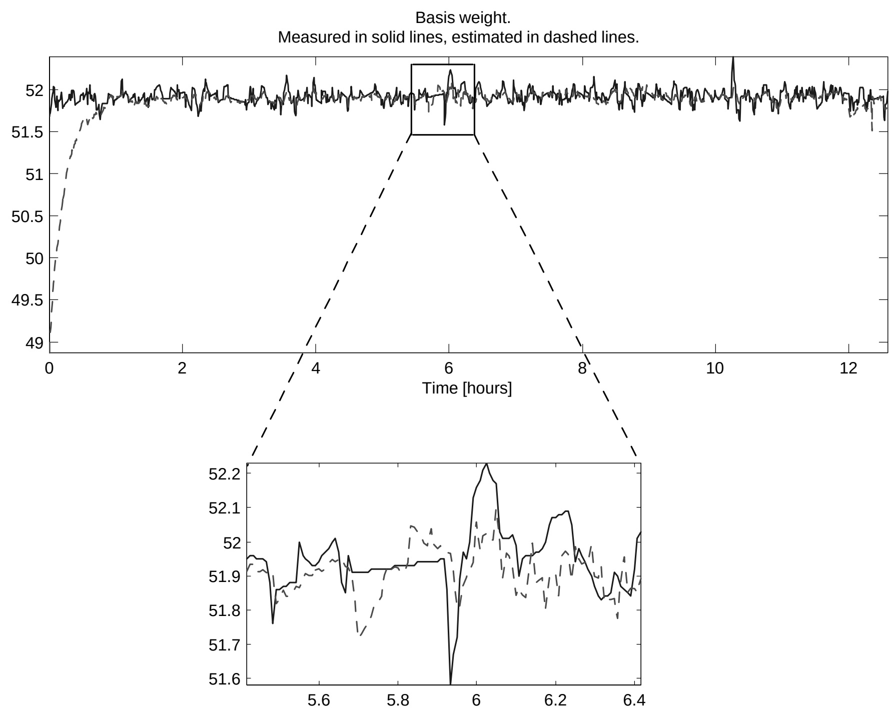

Figure 2: Validation of Kalman filter performance. The measured basis weight is shown in solid line and the estimated is shown in dashed line.

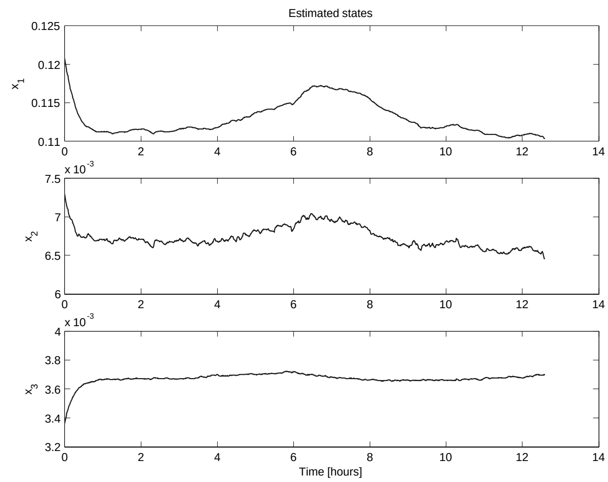

Figure 3: The estimated states for the validation shown in Figure 2.

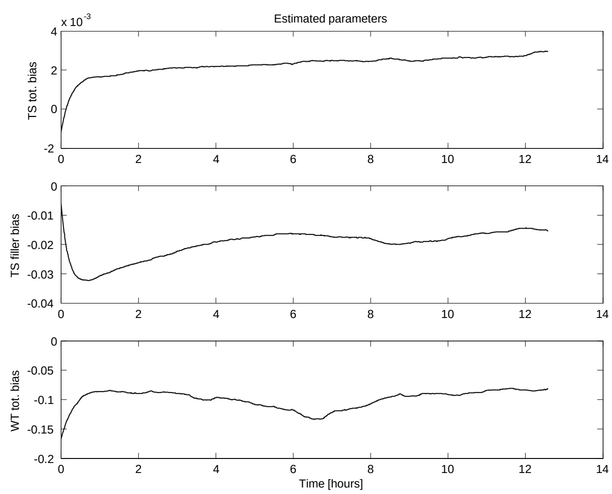

Figure 4: The estimated parameters for the validation shown in Figure 2.

6 Results

6.1 Implementation and interface

The MPC was installed at PM6 in March 2002. During the first two months, the MPC, the Kalman filter and the model were continuously tuned, retuned, and validated in open and closed loop. Some structural changes were also made during these months. From May 2002, the MPC has been in operation more or less continuously. The process operators still have the original "pre-MPC era" control configuration available, but the MPC has been the preferred choice from the beginning. Furthermore, the operators have been very active in making suggestions for improvements and new features in the system. Some of these suggestions are implemented, and others are being considered for implementation.

In addition to discussing and involving the operators in the project from the beginning, it is our opinion that the MPC interface has been very important for the positive operator attitude. Figure 5 shows part of the MPC interface at PM6. The upper row in the figure shows the basis weight, setpoint for basis weight, and the flow of thick stock. The middle row shows the paper ash, setpoint for paper ash, and the flow of filler added to the short circulation. The lower row shows the total concentration in the wire tray, the corresponding setpoint, and the flow of retention aid added to the short circulation. The interface and pairing of inputs and outputs are based on the pre-MPC control configuration, basically because this is how the operators and engineers at PM6 are used to see it. The vertical dashed line in the middle of each row is the current time. When Figure 5 was captured, the paper machine was in the middle of a grade change, and studying the figure carefully, one may see the setpoints change at the current time. The setpoints for the new grade were submitted to the MPC some time before the grade change, so at the time of the grade change the outputs are actually half way to the new setpoints. In terms of gaining operator acceptance for the MPC, this feature of previewing the action taken by the controller has been very helpful. The operators can specify a grade change e.g. half an hour into the future, and see how the MPC will achieve the change: how the inputs will be manipulated to reach the new setpoints.

6.2 Reduction of variation

An important objective with the MPC was to reduce variation in consistencies, basis weigh, paper ash, paper moisture, and more. Figure 6 shows and example with the wire tray concentration and the paper ash. The bottom line indicates whether the MPC is on (at 1) or off (at 0). When the controller is off, the original control configuration is used. The MPC provides a distinct effect of reduced variation in these two outputs.

The main objective of the project "Stabilization of the wet end at PM6" was to increase the total efficiency by 0.47%. This is an objective that is unmeasurable, due to many factors affecting the total efficiency. Thus, several sub-goals were defined which were assumed easier to measure and validate. The sub-goals, and results,

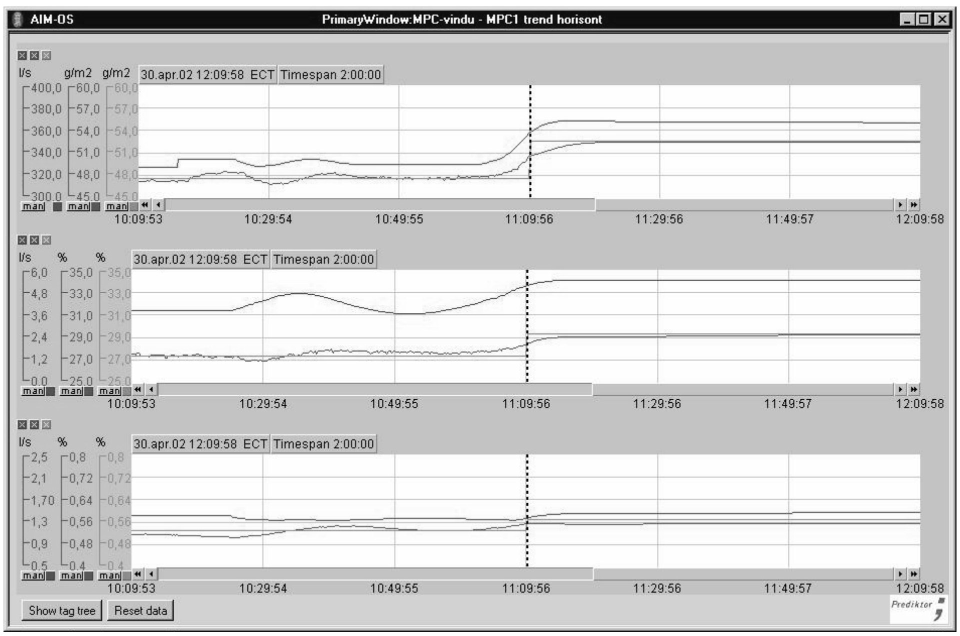

Figure 5: Part of the MPC interface at PM6.

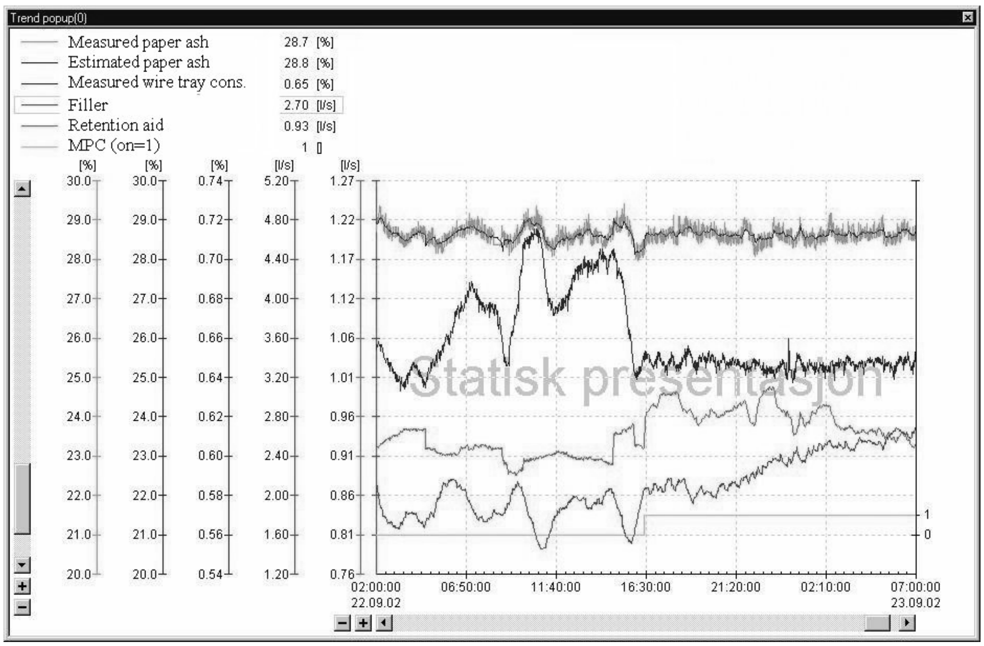

Figure 6: Wire tray concentration and paper ash, with (bottom line is 1) and without (bottom line is 0) MPC. From top to bottom the following variables are shown: Measured and estimated paper ash (overlapping), wire tray total concentration, retention aid, filler, and MPC on/off indication.

concerning reduced variability are:

Variable	Sub-goal (red. std. dev.)	Result
Total cons. in the wire tray	60%	OK
Filler cons. in the wire tray	50%	OK
Total cons. in the headbox	50%	OK
Filler cons. in the headbox	35%	OK
Basis weight	20%	No change achieved
Paper ash	20%	OK
Paper moisture	20%	OK

These sub-goals were defined in 1999 when the project was initiated. In 2001 a new scanning device for measuring e.g. basis weight and paper ash was installed at PM6. This significantly improved the control of the basis weight using the "old" controllers. The results in the table above are calculated with the measurement devices as of 2002, comparing the old control configuration with the MPC control configuration. Exact numbers for the reduction in standard deviation are not given, as they vary from day to day, and from operator to operator.

6.3 Other benefits of MPC

In addition to reducing the variation in key paper machine variables, several other benefits are obtained using MPC. Some of these benefits arise from utilizing the developed model, not only for control purposes, but also as a replacement for measurements when these are not available or not trustworthy.

Grade changes in automatic mode Previously, grade changes were carried out manually or partly manually (the setpoints were changed a number of times before they were equal to the new grade) by the operators. With a mechanistic model, applicable over a wide range of operating conditions, the grade changes are carried out using the MPC (see Figure 5). This has resulted in faster grade changes and operator independent grade changes. During larger grade changes, the use of MPC results in less off-spec paper being produced during the change. Using one mechanistic model, the grade change is handled in a straight forward fashion, as there is no need to switch between various local models.

Control during sheet breaks The basis weight and paper ash outputs can not be measured during sheet breaks. Previously, during sheet breaks the flow of thick stock and filler were frozen at the value they had immediately prior to the break. Usually the sheet breaks last less than half an hour, and the output variables are not far from target values when the paper is back on the reel. However, occasionally the sheet breaks last longer periods and there may be e.g. velocity changes during the break, leading to off-spec paper being produced for a period following the break. Another frequently experienced problem are large measurement errors immediately

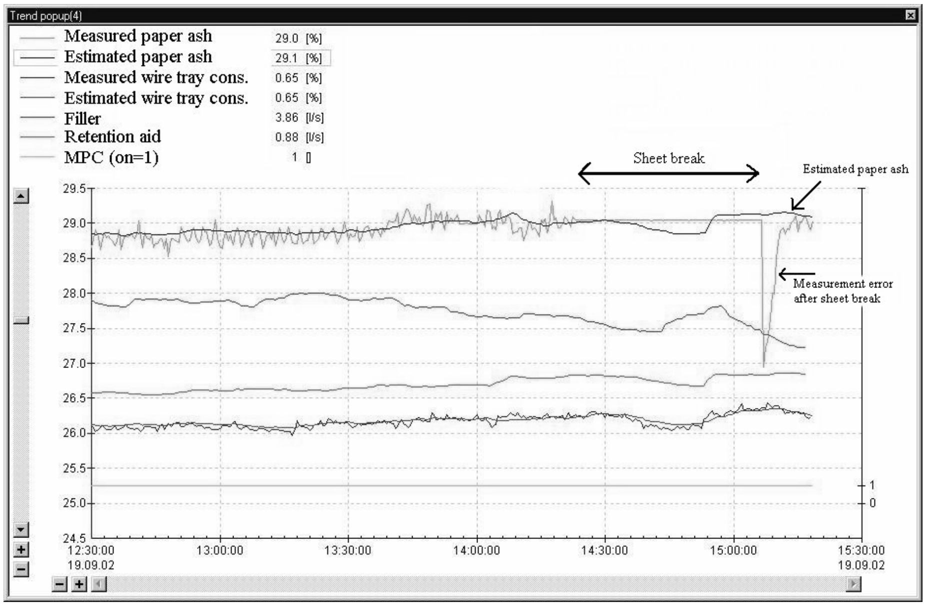

Figure 7: Sheet break. From top to bottom the following variables are shown: Measured and estimated paper ash (overlapping), filler, retention aid, measured and estimated wire tray total concentration (overlapping),and MPC on/off indication.

after a sheet break. With the MPC, the Kalman filter estimates the basis weight and paper ash during sheet breaks, and these estimates are used in the MPC as if no break had taken place. Thus, when the paper is back on the reel, the outputs are close to their setpoints.

In Figure 7 a sheet break, followed by a large measurement error, is shown. The two lines at the top are the measured and estimated paper ash (the lines are overlapping to some extent). During the sheet break the measured value is lost, and thus frozen at the value immediately prior to the break. When the paper is back on the reel, a large measurement error occurs giving a difference between measured and estimated value above 2%. The measured value converges to the estimated value before the estimate is updated in the Kalman filter. The MPC use the estimated values and is thus unaffected by the erroneous measurement. Studying the inputs, it is obvious that it is the measurement that is erroneous, and not the estimate. The rise in measured paper ash from approximately 27% to 29% in less than 10 minutes is too fast to be realistic by itself, and the fact that this happens during a period when the dosage of retention aid is constant and the filler is decreasing is very hard to explain.

Control during start ups Previously, the controllers were not set to automatic mode before the outputs were close to the setpoints, following a start up. With a model based controller using a mechanistic model with a wide operating range, the MPC is set to automatic mode early during start ups. This results in faster start ups, and less off-spec paper being produced.

Control during periods with poor measurements Occasionally a special filler is added to the stock, to increase the brightness of the paper. During these periods the consistency measurements are not trustworthy as they are based on optical measurement methods. This problem is solved within the MPC / Kalman filter framework by neglecting the updates of the consistency estimate, relying on the estimate alone. For each output, there is an option within the MPC to neglect the updating of states based on this output. This is done based on experience with periods of poor measurements, even when only standard filler is used. Figures 8–9 show an example of a period of poor wire tray consistency measurement. There are large variations in all outputs in the first half of the period shown in the figures. When the MPC was switched on, the updating of states from the wire tray consistency measurement was switched off. The effect is pronounced, as the paper ash, basis weight, moisture, and also all inputs vary considerably less in this latter half. Note that the measurement of wire tray consistency is the only variable that varies equally much in the first and second halves.

Filtering of measurements The Kalman filter estimates are used in the MPC instead of the measurements. This leads to smoother controller action, and eliminates the need for additional filtering.

Updating of model parameters The model is augmented so that some key parameters/biases are updated automatically. This reduces the need for model maintenance off-line. However, should there be larger changes in the process, such as if the white water tank is removed, or a new retention aid is used, then it will probably be necessary to re-tune the model and controller.

7 Conclusions

A mechanistic nonlinear model of the wet end of PM6 at Norske Skog Saugbrugs has been developed and used in an MPC application. The MPC uses an infinite horizon criterion, successive linearization of the model, and estimation of states and parameters by an augmented Kalman filter.

Variation in important quality variables and consistencies in the wet end have been reduced substantially, compared to the variation prior to the MPC implementation. The MPC also provides better efficiency through faster grade changes, control during sheet breaks and start ups, and better control during periods of poor measurements.

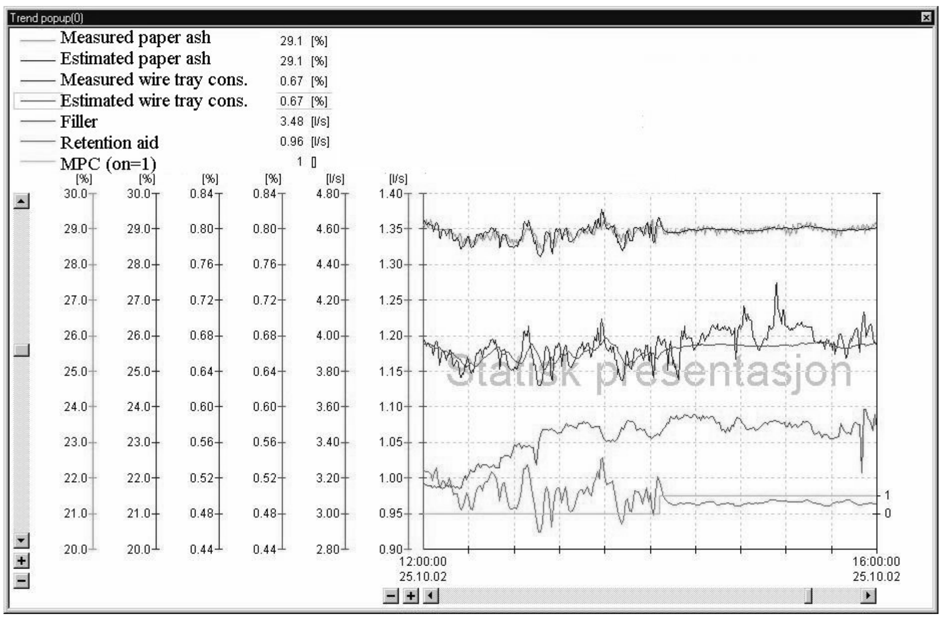

Figure 8: Period of poor wire tray consistency measurement. During the second half, the controller relies on the estimated consistency, rather than the measured. From top to bottom the following variables are shown: Measured and estimated paper ash (overlapping), measured and estimated wire tray total concentration (overlapping), filler, retention aid, and MPC on/off indication.

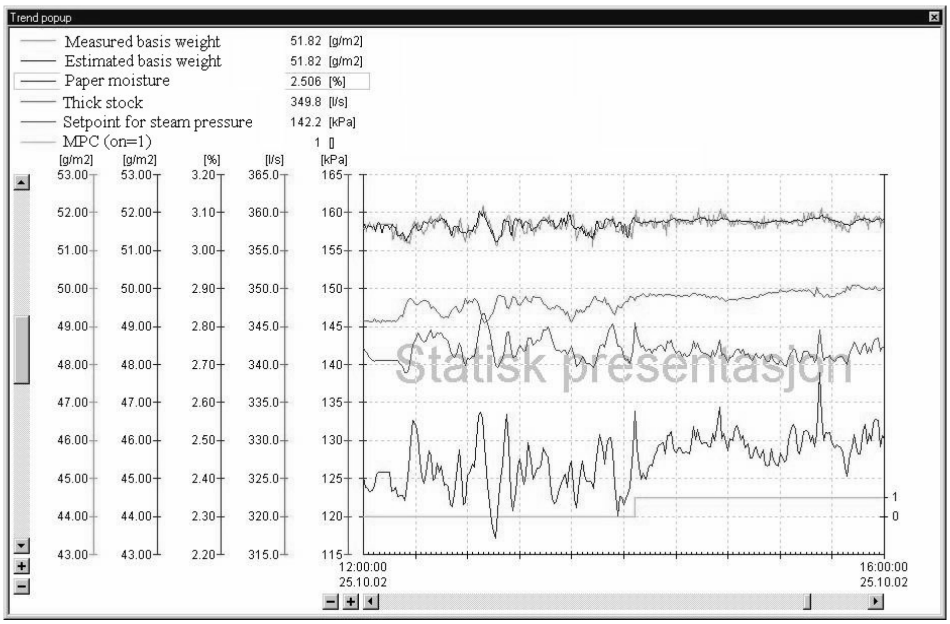

Figure 9: Period of poor wire tray consistency measurement. During the second half, the controller relies on the estimated consistency, rather than the measured. From top to bottom the following variables are shown: Measured and estimated basis weight (overlapping), flow of thick stock, setpoint for steam pressure, paper moisture, and MPC on/off indication.

Acknowledgements The authors would like to thank the employees at PM6 for their cooperation in providing information and data for this paper, and for their general helpfulness. Steinar Sælid and Anders Veberg from Prediktor AS are acknowledged for information about Apis and Apis MPC. The work of Tor Anders Hauge is financially supported by the Research Council of Norway (project number 134557/432), with additional financial support by Norske Skog Saugbrugs.

A Notes about notation

Variables in original units, i.e. unscaled and unshifted, are denoted by a bar above the variable, e.g. $\bar{x}$ and $\bar{u}$. Variables in the linearized model, i.e. variables that have origin corresponding to the center of linearization are denoted by a double bar above the variable, e.g. $\bar{\bar{x}}$ and $\bar{\bar{u}}$. Finally, variables shifted first by linearization and then by the steady state values at time $k+N-1$ are shown as e.g. x and u.

B Example: Finding the steady state values with `lssol`

The solution to eqs. 15-17 can be found by using standard optimization software. The `lssol` algorithm (Gill et al. 1986) for solving constrained linear least squares problems, is available with a Matlab interface in the Tomlab environment (Holmström 2001). The algorithm solves the following problems (amongst others)

$$\min_{z^s_{k+j}} J^s_{k+j} = \min_{z^s} \frac{1}{2}(b^s_{k+j} - A^s_{k+j} z^s_{k+j})^T (b^s_{k+j} - A^s_{k+j} z^s_{k+j}), \tag{34}$$

with constraints

$$b_{L,k+j} \leq \begin{bmatrix} z^s_{k+j} \\ \bar{A}_k z^s_{k+j} \end{bmatrix} \leq b_{U,k+j}, \tag{35}$$

where $j = 0, ..., N-1$ so that the optimization must in principle be carried out at each sample in the future horizon. In practice one would only calculate the steady state values once for every change in the future reference values, measured disturbance values, or max/min values.

The matrix A^s_{k+j} is then found in eq. 16 as

$$A^s_{k+j} = \begin{bmatrix} (I - A_k) & -B_k \\ -C_k & -D_k \end{bmatrix}, \tag{36}$$

while b^s_{k+j} is

$$b^s_{k+j} = \begin{bmatrix} E_k \bar{\bar{d}}_{k+j} \\ F_k \bar{\bar{d}}_{k+j} - \bar{y}^s_{k+j} \end{bmatrix}. \tag{37}$$

The constraints on $\bar{u}^s_{k+j}$ is

$$\bar{u}^{\min}_{k+j} - \bar{u}_{k-1} \leq \bar{u}^s_{k+j} \leq \bar{u}^{\max}_{k+j} - \bar{u}_{k-1},\ j = 0, ..., N-1 \tag{38}$$

and the constraints on $\bar{y}^s_{k+j}$ can be formulated as

$$\begin{gathered}
\bar{y}^{\min}_{k+j} - \bar{y}_{k-1} \leq \bar{y}^s_{k+j} \leq \bar{y}^{\max}_{k+j} - \bar{y}_{k-1},\ j = 0, ..., N-1 \\
\Downarrow \\
\bar{y}^{\min}_{k+j} - \bar{y}_{k-1} \leq C_k \bar{x}^s_{k+j} + D_k \bar{u}^s_{k+j} \leq \bar{y}^{\max}_{k+j} - \bar{y}_{k-1},\ j = 0, ..., N-1 \\
\Downarrow \\
\bar{y}^{\min}_{k+j} - \bar{y}_{k-1} - F_k \bar{\bar{d}}_{k+j} \leq \begin{bmatrix} C_k & D_k \end{bmatrix} \begin{bmatrix} \bar{x}^s_{k+j} \\ \bar{u}^s_{k+j} \end{bmatrix} \leq \bar{y}^{\max}_{k+j} - \bar{y}_{k-1} - F_k \bar{\bar{d}}_{k+j} \\
,\ j = 0, ..., N-1.
\end{gathered} \tag{39}$$

The constraints can then be written, as in eq. 35, as

$$\begin{aligned}
b_{L,k+j} &= \begin{bmatrix} -\infty \\ \bar{u}^{\min}_{k+j} - \bar{u}_{k-1} \\ \bar{y}^{\min}_{k+j} - \bar{y}_{k-1} - F_k \bar{\bar{d}}_{k+j} \end{bmatrix},\ j = 0, ..., N-1 \\
b_{U,k+j} &= \begin{bmatrix} \infty \\ \bar{u}^{\max}_{k+j} - \bar{u}_{k-1} \\ \bar{y}^{\max}_{k+j} - \bar{y}_{k-1} - F_k \bar{\bar{d}}_{k+j} \end{bmatrix},\ j = 0, ..., N-1 \\
\bar{A}_k &= \begin{bmatrix} C_k & D_k \end{bmatrix}.
\end{aligned} \tag{40}$$

C Proofs for finite horizon criterion

C.1 Reduction to finite horizon criterion

We split the infinite horizon criterion in two sums

$$\begin{aligned}
\min_{\mathcal{U}_k} J_k &= \min_{\mathcal{U}_k} \sum_{j=0}^{\infty} \left[e^T_{k+j} Q e_{k+j} + \tilde{u}^T_{k+j} R \tilde{u}_{k+j} + \Delta u^T_{k+j} S \Delta u_{k+j} \right] \\
&= \min_{\mathcal{U}_k} \left(\sum_{j=N}^{\infty} \left[e^T_{k+j} Q e_{k+j} + \tilde{u}^T_{k+j} R \tilde{u}_{k+j} + \Delta u^T_{k+j} S \Delta u_{k+j} \right] \right. \\
&\quad \left. + \sum_{j=0}^{N-1} \left[e^T_{k+j} Q e_{k+j} + \tilde{u}^T_{k+j} R \tilde{u}_{k+j} + \Delta u^T_{k+j} S \Delta u_{k+j} \right] \right),
\end{aligned} \tag{41}$$

where the sum from zero to $N-1$ can be easily calculated, while the infinite sum from N to infinity must be studied carefully. We take a closer look at each of the three terms in that part of the criterion where the sum is infinite. First we study the $\tilde{u}$ term:

$$\sum_{j=N}^{\infty} \tilde{u}_{k+j}^T R \tilde{u}_{k+j} = \sum_{j=N}^{\infty} \left(u_{k+j} - u_{k+j}^s\right)^T R \left(u_{k+j} - u_{k+j}^s\right) = 0 \tag{42}$$

which is a direct consequence of the definition of $\tilde{u}_{k+j}$ in eq. 25, and the assumptions in eq. 27. We then study the Δu term:

$$\sum_{j=N}^{\infty} \Delta u_{k+j}^T S \Delta u_{k+j} = \Delta u_{k+N}^T S \Delta u_{k+N}, \tag{43}$$

because of assumptions made in eq. 27. Finally we study the e term:

$$\begin{aligned}
\sum_{j=N}^{\infty} e_{k+j}^T Q e_{k+j} &= \sum_{j=N}^{\infty} \left(y_{k+j} - \overbrace{y_{k+j}^s}^{0}\right)^T Q \left(y_{k+j} - \overbrace{y_{k+j}^s}^{0}\right) \\
&= \sum_{j=N}^{\infty} \left(C_k x_{k+j} + D_k \overbrace{u_{k+j}}^{0} + F_k \overbrace{d_{k+j}}^{0}\right)^T Q \left(C_k x_{k+j} + D_k \overbrace{u_{k+j}}^{0} + F_k \overbrace{d_{k+j}}^{0}\right) \\
&= \sum_{j=N}^{\infty} x_{k+j}^T C_k^T Q C_k x_{k+j},
\end{aligned} \tag{44}$$

where we have used eqs. 23, 25 and 27. Eq. 44 needs to be studied further, but first we will establish some facts needed. Consider a decomposition of the A_k matrix into its Jordan form

$$A_k = V_k J_k V_k^{-1} = \begin{bmatrix} V_k^u & V_k^s \end{bmatrix} \begin{bmatrix} J_k^u & 0 \\ 0 & J_k^s \end{bmatrix} \begin{bmatrix} \tilde{V}_k^u \\ \tilde{V}_k^s \end{bmatrix}, \tag{45}$$

where V_k^u and J_k^u are respectively the eigenvectors and Jordan blocks for the eigenvalues corresponding to the unstable modes of A_k, and V_k^s and J_k^s are respectively the eigenvectors and Jordan blocks for the eigenvalues corresponding to the stable modes of A_k.Consider x_{k+j}, $\forall j = N, ...\infty$

$$\begin{aligned}
x_{k+N} &= x_{k+N} \\
x_{k+N+1} &= A_k x_{k+N} + B_k \overbrace{u_{k+N}}^{0} + E_k \overbrace{d_{k+N}}^{0} \\
x_{k+N+2} &= A_k^2 x_{k+N} + B_k \overbrace{u_{k+N+1}}^{0} + E_k \overbrace{d_{k+N+1}}^{0} \\
&\cdots
\end{aligned}$$

which gives

$$x_{k+j} = A_k^{j-N} x_{k+N}, \forall j = N, N+1, N+2, ... \tag{46}$$

Inserting eq. 45 into 46 gives

$$\begin{aligned} x_{k+j} &= A_k^{j-N} x_{k+N} = \left(V_k J_k V_k^{-1}\right)^{j-N} x_{k+N} \\ &= V_k J_k^{j-N} V_k^{-1} x_{k+N} \\ &= \left[\begin{array}{cc} V_k^u & V_k^s \end{array}\right] \left[\begin{array}{cc} (J_k^u)^{j-N} & 0 \\ 0 & (J_k^s)^{j-N} \end{array}\right] \left[\begin{array}{c} \tilde{V}_k^u \\ \tilde{V}_k^s \end{array}\right] x_{k+N} \\ &= V_k^u (J_k^u)^{j-N} \tilde{V}_k^u x_{k+N} + V_k^s (J_k^s)^{j-N} \tilde{V}_k^s x_{k+N}, \ \forall j = N, N+1, N+2, ... \end{aligned} \tag{47}$$

At this point we introduce a new constraint, i.e. we force the unstable modes to zero at time $k+N$

$$\tilde{V}_k^u x_{k+N} = 0, \tag{48}$$

and this gives then

$$x_{k+j} = V_k^s (J_k^s)^{j-N} \tilde{V}_k^s x_{k+N}, \ \forall j = N, N+1, N+2, ...$$

This result is inserted into eq. 44:

$$\begin{aligned} &\sum_{j=N}^{\infty} x_{k+j}^T C_k^T Q C_k x_{k+j} \\ &= \sum_{j=N}^{\infty} \left(V_k^s (J_k^s)^{j-N} \tilde{V}_k^s x_{k+N}\right)^T C_k^T Q C_k \left(V_k^s (J_k^s)^{j-N} \tilde{V}_k^s x_{k+N}\right) \\ &= x_{k+N}^T \left(\tilde{V}_k^s\right)^T \left[\sum_{j=N}^{\infty} \left((J_k^s)^{j-N}\right)^T (V_k^s)^T C_k^T Q C_k V_k^s (J_k^s)^{j-N}\right] \tilde{V}_k^s x_{k+N} \\ &= x_{k+N}^T \left(\tilde{V}_k^s\right)^T \left[\sum_{j=0}^{\infty} \left((J_k^s)^{j}\right)^T (V_k^s)^T C_k^T Q C_k V_k^s (J_k^s)^{j}\right] \tilde{V}_k^s x_{k+N} \\ &= x_{k+N}^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s x_{k+N}, \end{aligned} \tag{49}$$

where $\bar{Q}_k$ is given by the discrete Lyapunov equation

$$\bar{Q}_k = (V_k^s)^T C_k^T Q C_k V_k^s + (J_k^s)^T \bar{Q}_k J_k^s, \tag{50}$$

because

$$\begin{aligned} &\overbrace{\sum_{j=0}^{\infty} \left((J_k^s)^{j}\right)^T (V_k^s)^T C_k^T Q C_k V_k^s (J_k^s)^{j}}^{\bar{Q}_k} \\ &= (V_k^s)^T C_k^T Q C_k V_k^s + (J_k^s)^T \overbrace{\sum_{j=0}^{\infty} \left((J_k^s)^{j}\right)^T (V_k^s)^T C_k^T Q C_k V_k^s (J_k^s)^{j}}^{\bar{Q}_k} (J_k^s). \end{aligned}$$

From eq. 41, 42, 43, and 49, we have the criterion

$$\begin{aligned}\min_{\mathcal{U}_k} J_k = \min_{\mathcal{U}_k} \Bigg(& \sum_{j=N}^{\infty} \left[e_{k+j}^T Q e_{k+j} + \tilde{u}_{k+j}^T R \tilde{u}_{k+j} + \Delta u_{k+j}^T S \Delta u_{k+j}\right] \\ & + \sum_{j=0}^{N-1} \left[e_{k+j}^T Q e_{k+j} + \tilde{u}_{k+j}^T R \tilde{u}_{k+j} + \Delta u_{k+j}^T S \Delta u_{k+j}\right] \Bigg) \\ = \min_{\mathcal{U}_k} \Bigg(& x_{k+N}^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s x_{k+N} + \Delta u_{k+N}^T S \Delta u_{k+N} \\ & + \sum_{j=0}^{N-1} \left[e_{k+j}^T Q e_{k+j} + \tilde{u}_{k+j}^T R \tilde{u}_{k+j} + \Delta u_{k+j}^T S \Delta u_{k+j}\right] \Bigg),\end{aligned} \tag{51}$$

with $\bar{Q}_k$ given by eq. 50, $\tilde{V}_k^s$ found from the Jordan decomposition of A_k (see eq. 45), and with the additional equality constraint given by eq. 48.

C.2 Formulation as standard QP problem

Define

$$\mathcal{U}_k = \begin{bmatrix} u_k \\ \vdots \\ u_{k+N-1} \end{bmatrix}, \mathcal{U}_k^s = \begin{bmatrix} u_k^s \\ \vdots \\ u_{k+N-1}^s \end{bmatrix}, \mathcal{Y}_k = \begin{bmatrix} y_k \\ \vdots \\ y_{k+N-1} \end{bmatrix}, \tag{52}$$
$$\mathcal{Y}_k^s = \begin{bmatrix} y_k^s \\ \vdots \\ y_{k+N-1}^s \end{bmatrix}, \mathcal{D}_k = \begin{bmatrix} d_k \\ \vdots \\ d_{k+N-1} \end{bmatrix}, \Delta\mathcal{U}_k = \begin{bmatrix} \Delta u_k \\ \vdots \\ \Delta u_{k+N} \end{bmatrix}.$$

Using the model in eq. 19 to predict future state and output values, we have the following results

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + E_k d_k \\ x_{k+2} &= A_k x_{k+1} + B_k u_{k+1} + E_k d_{k+1} \\ &= A_k^2 x_k + A_k B_k u_k + A_k E_k d_k + B_k u_{k+1} + E_k d_{k+1} \\ x_{k+3} &= A_k x_{k+2} + B_k u_{k+2} + E_k d_{k+2} \\ &= A_k^3 x_k + A_k^2 B_k u_k + A_k^2 E_k d_k + A_k B_k u_{k+1} + A_k E_k d_{k+1} + B_k u_{k+2} + E_k d_{k+2} \\ &\vdots \\ x_{k+i} &= A_k^i x_k + \sum_{j=1}^{i} A_k^{i-j} B_k u_{k+j-1} + \sum_{j=1}^{i} A_k^{i-j} E_k d_{k+j-1}, \end{aligned} \tag{53}$$

and

$$\begin{aligned} y_{k+i} &= C_k x_{k+i} + D_k u_{k+i} + F_k d_{k+i} \\ &= C_k A_k^i x_k + \sum_{j=1}^{i} C_k A_k^{i-j} B_k u_{k+j-1} + \sum_{j=1}^{i} C_k A_k^{i-j} E_k d_{k+j-1} + D_k u_{k+i} + F_k d_{k+i}. \end{aligned} \tag{54}$$

We then derive equations for $\mathcal{Y}_k$ and $\Delta\mathcal{U}_k$ in terms of $\mathcal{U}_k$, $\mathcal{D}_k$, x_k and u_{k-1}

$$\begin{aligned} \mathcal{Y}_k &= \mathcal{O}_k x_k + \mathcal{H}_k^u \mathcal{U}_k + \mathcal{H}_k^d \mathcal{D}_k \\ \Delta\mathcal{U}_k &= \mathcal{P}\mathcal{U}_k + \mathcal{L} u_{k-1}, \end{aligned} \tag{55}$$

where

$$\mathcal{O}_k = \begin{bmatrix} C_k \\ C_k A_k \\ C_k A_k^2 \\ \vdots \\ C_k A_k^{N-1} \end{bmatrix} \tag{56}$$

$$\mathcal{H}_k^u = \begin{bmatrix} D_k & 0 & \cdots & 0 & 0 \\ C_k B_k & D_k & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_k A_k^{N-3} B_k & C_k A_k^{N-4} B_k & \cdots & D_k & 0 \\ C_k A_k^{N-2} B_k & C_k A_k^{N-3} B_k & \cdots & C_k B_k & D_k \end{bmatrix}$$

$$\mathcal{H}_k^d = \begin{bmatrix} F_k & 0 & \cdots & 0 & 0 \\ C_k E_k & F_k & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_k A_k^{N-3} E_k & C_k A_k^{N-4} E_k & \cdots & F_k & 0 \\ C_k A_k^{N-2} E_k & C_k A_k^{N-3} E_k & \cdots & C_k E_k & F_k \end{bmatrix}$$

$$\mathcal{P} = \begin{bmatrix} I & 0 & \cdots & 0 \\ -I & I & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -I & I \\ 0 & \cdots & 0 & -I \end{bmatrix}$$

$$\mathcal{L} = \begin{bmatrix} -I \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

Each term in the criterion in eq. 51, and in the constraints in eqs. 24 and 48 is now written in terms of the unknown variable $\mathcal{U}_k$, and the known variables $\mathcal{U}_k^s$, $\mathcal{Y}_k^s$, $\mathcal{D}_k$, $\mathcal{U}_k^{\min}$, $\mathcal{U}_k^{\max}$, $\mathcal{Y}_k^{\min}$, $\mathcal{Y}_k^{\max}$, $\Delta\mathcal{U}_k^{\min}$, $\Delta\mathcal{U}_k^{\max}$, x_k and u_{k-1}.

C.2.1 The criterion

We start with the first term in the criterion in eq. 51, using the result in eq. 53

$$
\begin{aligned}
& x_{k+N}^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s x_{k+N} \qquad (57) \\
&= \left(A_k^N x_k + \sum_{j=1}^{N} A_k^{N-j} B_k u_{k+j-1} + \sum_{j=1}^{N} A_k^{N-j} E_k d_{k+j-1}\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s \left(\cdot\right) \\
&= \left(A_k^N x_k + C_k^u P^u \mathcal{U}_k + C_k^d P^d \mathcal{D}_k\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s \left(A_k^N x_k + C_k^u P^u \mathcal{U}_k + C_k^d P^d \mathcal{D}_k\right) \\
&= \left(x_k^T \left(A_k^N\right)^T + \mathcal{U}_k^T \left(P^u\right)^T \left(C_k^u\right)^T + \mathcal{D}_k^T \left(P^d\right)^T \left(C_k^d\right)^T\right) \\
&\quad \cdot \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s \left(A_k^N x_k + C_k^u P^u \mathcal{U}_k + C_k^d P^d \mathcal{D}_k\right) \\
&= \overbrace{\mathcal{U}_k^T \left(P^u\right)^T \left(C_k^u\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s C_k^u P^u \mathcal{U}_k}^{\text{quadratic}} \\
&+ \overbrace{2 x_k^T \left(A_k^N\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s C_k^u P^u \mathcal{U}_k + 2 \mathcal{D}_k^T \left(P^d\right)^T \left(C_k^d\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s C_k^u P^u \mathcal{U}_k}^{\text{linear}} \\
&+ \overbrace{\left(A_k^N x_k + C_k^d P^d \mathcal{D}_k\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s \left(A_k^N x_k + C_k^d P^d \mathcal{D}_k\right)}^{\text{constant}}
\end{aligned}
$$

where

$$
\begin{aligned}
C_k^u &= \begin{bmatrix} B_k & A_k B_k & \cdots & A_k^{N-1} B_k \end{bmatrix} \qquad (58) \\
C_k^d &= \begin{bmatrix} E_k & A_k E_k & \cdots & A_k^{N-1} E_k \end{bmatrix} \\
P^u &= I_N^{\text{rot}90} \otimes I_r \\
P^d &= I_N^{\text{rot}90} \otimes I_g
\end{aligned}
$$

and $I_N^{\text{rot}90}$ is an $N \times N$ identity matrix rotated 90 degrees, I_m and I_g are identity matrices of size m and g respectively, and $\otimes$ is the Kronecker product.

We then study the two Δu terms in the criterion (eq. 51), using the result obtained in eq. 55

$$
\begin{aligned}
&\sum_{j=0}^{N} \Delta u_{k+j}^T S \Delta u_{k+j} \\
&= \Delta \mathcal{U}_k^T \left(I_{N+1} \otimes S\right) \Delta \mathcal{U}_k \\
&= \left(\mathcal{P}\mathcal{U}_k + \mathcal{L}u_{k-1}\right)^T \left(I_{N+1} \otimes S\right) \left(\mathcal{P}\mathcal{U}_k + \mathcal{L}u_{k-1}\right) \\
&= \left(\mathcal{U}_k^T \mathcal{P}^T + u_{k-1}^T \mathcal{L}^T\right) \left(I_{N+1} \otimes S\right) \left(\mathcal{P}\mathcal{U}_k + \mathcal{L}u_{k-1}\right) \\
&= \overbrace{\mathcal{U}_k^T \mathcal{P}^T \left(I_{N+1} \otimes S\right) \mathcal{P}\mathcal{U}_k}^{\text{quadratic}} + \overbrace{2u_{k-1}^T \mathcal{L}^T \left(I_{N+1} \otimes S\right) \mathcal{P}\mathcal{U}_k}^{\text{linear}} + \overbrace{u_{k-1}^T \mathcal{L}^T \left(I_{N+1} \otimes S\right) \mathcal{L}u_{k-1}}^{\text{constant}}.
\end{aligned}
\tag{59}
$$

Next, the control error term is studied, using the definitions in eqs. 25 and 52, and the results from eq. 55

$$
\begin{aligned}
&e_{k+j}^T Q e_{k+j} \\
&= \left(\mathcal{Y}_k - \mathcal{Y}_k^s\right)^T \left(I_N \otimes Q\right) \left(\mathcal{Y}_k - \mathcal{Y}_k^s\right) \\
&= \left(\mathcal{O}_k x_k + \mathcal{H}_k^u \mathcal{U}_k + \mathcal{H}_k^d \mathcal{D}_k - Y_k^s\right)^T \left(I_N \otimes Q\right) \left(\mathcal{O}_k x_k + \mathcal{H}_k^u \mathcal{U}_k + \mathcal{H}_k^d \mathcal{D}_k - Y_k^s\right) \\
&= \overbrace{\mathcal{U}_k^T \left(\mathcal{H}_k^u\right)^T \left(I_N \otimes Q\right) \mathcal{H}_k^u \mathcal{U}_k}^{\text{quadratic}} \\
&+ \overbrace{2x_k^T \mathcal{O}_k^T \left(I_N \otimes Q\right) \mathcal{H}_k^u \mathcal{U}_k + 2\mathcal{D}_k^T \left(\mathcal{H}_k^d\right)^T \left(I_N \otimes Q\right) \mathcal{H}_k^u \mathcal{U}_k - 2\left(Y_k^s\right)^T \left(I_N \otimes Q\right) \mathcal{H}_k^u \mathcal{U}_k}^{\text{linear}} \\
&+ \overbrace{\left(\mathcal{O}_k x_k + \mathcal{H}_k^d \mathcal{D}_k - Y_k^s\right)^T \left(I_N \otimes Q\right) \left(\mathcal{O}_k x_k + \mathcal{H}_k^d \mathcal{D}_k - Y_k^s\right)}^{\text{constant}}
\end{aligned}
\tag{60}
$$

The last term in the criterion to be studied, using the definitions in eqs. 25 and 52, is

$$
\begin{aligned}
&\sum_{j=0}^{N-1} \tilde{u}_{k+j}^T R \tilde{u}_{k+j} \\
&= \left(\mathcal{U}_k - \mathcal{U}_k^s\right)^T \left(I_N \otimes R\right) \left(\mathcal{U}_k - \mathcal{U}_k^s\right) \\
&= \left(\mathcal{U}_k^T - \left(\mathcal{U}_k^s\right)^T\right) \left(I_N \otimes R\right) \left(\mathcal{U}_k - \mathcal{U}_k^s\right) \\
&= \overbrace{\mathcal{U}_k^T \left(I_N \otimes R\right) \mathcal{U}_k}^{\text{quadratic}} - \overbrace{2\left(\mathcal{U}_k^s\right)^T \left(I_N \otimes R\right) \mathcal{U}_k}^{\text{linear}} + \overbrace{\left(\mathcal{U}_k^s\right)^T \left(I_N \otimes R\right) \mathcal{U}_k^s}^{\text{constant}}
\end{aligned}
\tag{61}
$$

C.2.2 The constraints

First we define the following

$$\mathcal{U}_k^{\min} = \begin{bmatrix} \bar{u}_k^{\min} \\ \vdots \\ \bar{u}_{k+N-1}^{\min} \end{bmatrix}, \mathcal{U}_k^{\max} = \begin{bmatrix} \bar{u}_k^{\max} \\ \vdots \\ \bar{u}_{k+N-1}^{\max} \end{bmatrix} \tag{62}$$
$$\mathcal{Y}_k^{\min} = \begin{bmatrix} -\infty \\ \vdots \\ -\infty \\ \bar{y}_{k+j_1}^{\min} \\ \vdots \\ \bar{y}_{k+j_2}^{\min} \\ -\infty \\ \vdots \\ -\infty \end{bmatrix}, \mathcal{Y}_k^{\max} = \begin{bmatrix} \infty \\ \vdots \\ \infty \\ \bar{y}_{k+j_1}^{\max} \\ \vdots \\ \bar{y}_{k+j_2}^{\max} \\ \infty \\ \vdots \\ \infty \end{bmatrix}$$
$$\Delta\mathcal{U}_k^{\min} = \begin{bmatrix} \Delta\bar{u}_k^{\min} \\ \vdots \\ \Delta\bar{u}_{k+N}^{\min} \end{bmatrix}, \Delta\mathcal{U}_k^{\max} = \begin{bmatrix} \Delta\bar{u}_k^{\max} \\ \vdots \\ \Delta\bar{u}_{k+N}^{\max} \end{bmatrix}$$

The constraints are given in eqs. 24 and 48. These are now written in terms of the unknown variable $\mathcal{U}_k$, and the known variables $\mathcal{U}_k^s$, $\mathcal{Y}_k^s$, $\mathcal{D}_k$, $\mathcal{U}_k^{\min}$, $\mathcal{U}_k^{\max}$, $\mathcal{Y}_k^{\min}$, $\mathcal{Y}_k^{\max}$, $\Delta\mathcal{U}_k^{\min}$, $\Delta\mathcal{U}_k^{\max}$, x_k and u_{k-1}. The constraints on the inputs are reformulated using eqs. 8, 20, 52, and 62

$$\begin{aligned} &\bar{u}_{k+j}^{\min} \leq \bar{u}_{k+j} \leq \bar{u}_{k+j}^{\max},\ j = 0, 1, ..., N-1 \\ &\Downarrow \\ &\bar{u}_{k+j}^{\min} \leq u_{k+j} + \bar{u}_{k-1} + \bar{u}_{k+N-1}^s \leq \bar{u}_{k+j}^{\max},\ j = 0, 1, ..., N-1 \\ &\Downarrow \\ &\mathcal{U}_k^{\min} \leq \mathcal{U}_k + \bar{u}_{k-1} \otimes \mathbf{1}_N + \bar{u}_{k+N-1}^s \otimes \mathbf{1}_N \leq \mathcal{U}_k^{\max} \\ &\Downarrow \end{aligned}$$
$$\mathcal{U}_k^{\min} - \bar{u}_{k-1} \otimes \mathbf{1}_N - \bar{u}_{k+N-1}^s \otimes \mathbf{1}_N \leq \mathcal{U}_k \leq \mathcal{U}_k^{\max} - \bar{u}_{k-1} \otimes \mathbf{1}_N - \bar{u}_{k+N-1}^s \otimes \mathbf{1}_N, \tag{63}$$

where $\mathbf{1}_N$ is an N dimensional vector with 1 in all elements. Next the constraints on the outputs are reformulated using eqs. 11, 20, 52, 55, and 62

$$\begin{aligned}
&\bar{y}_{k+j}^{\min} \le \bar{y}_{k+j} \le \bar{y}_{k+j}^{\max},\ j = j_1, j_1+1, ..., j_2 \\
&\quad\Downarrow \\
&\bar{y}_{k+j}^{\min} \le y_{k+j} + \bar{y}_{k-1} + \bar{y}_{k+N-1}^{s} \le \bar{y}_{k+j}^{\max},\ j = j_1, j_1+1, ..., j_2 \\
&\quad\Downarrow \\
&\mathcal{Y}_k^{\min} \le \mathcal{Y}_k + \bar{y}_{k-1} \otimes \mathbf{1}_N + \bar{y}_{k+N-1}^{s} \otimes \mathbf{1}_N \le \mathcal{Y}_k^{\max} \\
&\quad\Downarrow \\
&\quad\mathcal{Y}_k^{\min} - \bar{y}_{k-1} \otimes \mathbf{1}_N - \bar{y}_{k+N-1}^{s} \otimes \mathbf{1}_N \\
&\quad\le \mathcal{O}_k x_k + \mathcal{H}_k^u \mathcal{U}_k + \mathcal{H}_k^d \mathcal{D}_k \\
&\quad\le \mathcal{Y}_k^{\max} - \bar{y}_{k-1} \otimes \mathbf{1}_N - \bar{y}_{k+N-1}^{s} \otimes \mathbf{1}_N \\
&\quad\Downarrow \\
&\quad\mathcal{Y}_k^{\min} - \bar{y}_{k-1} \otimes \mathbf{1}_N - \bar{y}_{k+N-1}^{s} \otimes \mathbf{1}_N - \mathcal{O}_k x_k - \mathcal{H}_k^d \mathcal{D}_k \\
&\quad\le \mathcal{H}_k^u \mathcal{U}_k \\
&\quad\le \mathcal{Y}_k^{\max} - \bar{y}_{k-1} \otimes \mathbf{1}_N - \bar{y}_{k+N-1}^{s} \otimes \mathbf{1}_N - \mathcal{O}_k x_k - \mathcal{H}_k^d \mathcal{D}_k
\end{aligned} \tag{64}$$

The constraints on the input moves are reformulated using eqs. 25, 52, 55, and 62

$$\begin{aligned}
\Delta u_{k+j}^{\min} &\le \Delta u_{k+j} \le \Delta u_{k+j}^{\max},\ j = 0, 1, ..., N \\
&\Downarrow \\
\Delta \mathcal{U}_k^{\min} &\le \Delta \mathcal{U}_k \le \Delta \mathcal{U}_k^{\max} \\
&\Downarrow \\
\Delta \mathcal{U}_k^{\min} &\le \mathcal{P} \mathcal{U}_k + \mathcal{L} u_{k-1} \le \Delta \mathcal{U}_k^{\max} \\
&\Downarrow \\
\Delta \mathcal{U}_k^{\min} - \mathcal{L} u_{k-1} &\le \mathcal{P} \mathcal{U}_k \le \Delta \mathcal{U}_k^{\max} - \mathcal{L} u_{k-1}
\end{aligned} \tag{65}$$

Finally, we reformulate the constraint in eq. 48, using eqs. 57, and 58

$$\begin{aligned}
\tilde{V}_k^u x_{k+N} &= 0 \\
&\Downarrow \\
\tilde{V}_k^u \left(A_k^N x_k + C_k^u P^u \mathcal{U}_k + C_k^d P^d \mathcal{D}_k \right) &= 0 \\
&\Downarrow \\
\tilde{V}_k^u C_k^u P^u \mathcal{U}_k &= -\tilde{V}_k^u \left(A_k^N x_k + C_k^d P^d \mathcal{D}_k \right) \\
&\Downarrow \\
-\tilde{V}_k^u \left(A_k^N x_k + C_k^d P^d \mathcal{D}_k \right) \le \tilde{V}_k^u C_k^u P^u \mathcal{U}_k &\le -\tilde{V}_k^u \left(A_k^N x_k + C_k^d P^d \mathcal{D}_k \right)
\end{aligned} \tag{66}$$

C.2.3 Summary of standard QP problem

From eqs. 51, 57, 59, 60, and 61, we can write the criterion as

$$\min_{\mathcal{U}_k} J_k = \min_{\mathcal{U}_k} \left(\frac{1}{2} \mathcal{U}_k^T H_k \mathcal{U}_k + c_k^T \mathcal{U}_k + \kappa \right), \tag{67}$$

and using eqs. 24, 48, 63, 64, 65 and 66, we can write the constraints as

$$b_{L,k} \leq \begin{bmatrix} \mathcal{U}_k \\ \bar{A}_k \mathcal{U}_k \end{bmatrix} \leq b_{U,k}, \tag{68}$$

where

$$\begin{aligned} H_k &= 2 \Big((P^u)^T (C_k^u)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s C_k^u P^u + \mathcal{P}^T (I_{N+1} \otimes S) \mathcal{P} \\ &\quad + (\mathcal{H}_k^u)^T (I_N \otimes Q) \mathcal{H}_k^u + (I_N \otimes R) \Big) \\ c_k^T &= 2 \Big(x_k^T \left(A_k^N\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s C_k^u P^u + \mathcal{D}_k^T \left(P^d\right)^T \left(C_k^d\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s C_k^u P^u \\ &\quad + u_{k-1}^T \mathcal{L}^T (I_{N+1} \otimes S) \mathcal{P} + x_k^T \mathcal{O}_k^T (I_N \otimes Q) \mathcal{H}_k^u \\ &\quad + \mathcal{D}_k^T \left(\mathcal{H}_k^d\right)^T (I_N \otimes Q) \mathcal{H}_k^u - (Y_k^s)^T (I_N \otimes Q) \mathcal{H}_k^u - (\mathcal{U}_k^s)^T (I_N \otimes R) \Big) \\ \kappa &= \left(A_k^N x_k + C_k^d P^d \mathcal{D}_k\right)^T \left(\tilde{V}_k^s\right)^T \bar{Q}_k \tilde{V}_k^s \left(A_k^N x_k + C_k^d P^d \mathcal{D}_k\right) + u_{k-1}^T \mathcal{L}^T (I_{N+1} \otimes S) \mathcal{L} u_{k-1} \\ &\quad + \left(\mathcal{O}_k x_k + \mathcal{H}_k^d \mathcal{D}_k - Y_k^s\right)^T (I_N \otimes Q) \left(\mathcal{O}_k x_k + \mathcal{H}_k^d \mathcal{D}_k - Y_k^s\right) + (\mathcal{U}_k^s)^T (I_N \otimes R) \mathcal{U}_k^s \\ \bar{A}_k &= \begin{bmatrix} \mathcal{H}_k^u \\ \mathcal{P} \\ \tilde{V}_k^u C_k^u P^u \end{bmatrix} \\ b_{L,k} &= \begin{bmatrix} \mathcal{U}_k^{\min} - \bar{u}_{k-1} \otimes \mathbf{1}_N - \bar{u}_{k+N-1}^s \otimes \mathbf{1}_N \\ \mathcal{Y}_k^{\min} - \bar{y}_{k-1} \otimes \mathbf{1}_N - \bar{y}_{k+N-1}^s \otimes \mathbf{1}_N - \mathcal{O}_k x_k - \mathcal{H}_k^d \mathcal{D}_k \\ \Delta \mathcal{U}_k^{\min} - \mathcal{L} u_{k-1} \\ -\tilde{V}_k^u \left(A_k^N x_k + C_k^d P^d \mathcal{D}_k\right) \end{bmatrix} \\ b_{U,k} &= \begin{bmatrix} \mathcal{U}_k^{\max} - \bar{u}_{k-1} \otimes \mathbf{1}_N - \bar{u}_{k+N-1}^s \otimes \mathbf{1}_N \\ \mathcal{Y}_k^{\max} - \bar{y}_{k-1} \otimes \mathbf{1}_N - \bar{y}_{k+N-1}^s \otimes \mathbf{1}_N - \mathcal{O}_k x_k - \mathcal{H}_k^d \mathcal{D}_k \\ \Delta \mathcal{U}_k^{\max} - \mathcal{L} u_{k-1} \\ -\tilde{V}_k^u \left(A_k^N x_k + C_k^d P^d \mathcal{D}_k\right) \end{bmatrix} \end{aligned} \tag{69}$$

The criterion and constraints given by eqs. 67 and 68 are in the form used by e.g. the `sqopt` algorithm[3] (Gill et al. 1997). The `sqopt` algorithm is available with a Matlab interface in the Tomlab environment (Holmström 2001).

[3] The constant term κ in the criterion is not part of the `sqopt` algorithm (or any QP solver), and can be omitted without affecting the result of the optimization.

D State and parameter estimation

In this appendix we do not follow the notation used in other parts of this paper. For example $\bar{x}$ is here the a-priori state estimate, and not a state variable in original "global" units.

D.1 Kalman filter equations for linear time variant processes

In this section we derive the Kalman filter equations for a linear time variant system with colored process noise. The derivations are in particular based on (Ergon 2001) and to some extent on (Gelb 1974). We assume that the process is described by

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + E_k d_k + G_k w_k \\ y_k &= C_k x_k + D_k u_k + F_k d_k + v_k \end{aligned} \tag{70}$$

where $x_k \in \mathbb{R}^n$ is the state vector, $u_k \in \mathbb{R}^r$ is the (manipulated) input vector, $d_k \in \mathbb{R}^s$ is the measured disturbance vector, $y_k \in \mathbb{R}^m$ is the output vector, and k is the discrete time variable. w_k and v_k are zero mean uncorrelated white noise

$$\begin{aligned} E\left[w_k\right] = 0,\ E\left[v_k\right] &= 0, \\ E\left[w_k v_j^T\right] &= 0,\ \forall k, j, \\ E\left[w_k w_j^T\right] &= 0,\ E\left[v_k v_j^T\right] = 0,\ \forall k \neq j \\ E\left[w_k w_k^T\right] &= Q_k,\ E\left[v_k v_k^T\right] = R_k, \end{aligned} \tag{71}$$

where $E[\cdot]$ is the expectation operator, and $Q_k \in \mathbb{R}^{n\times n}$ and $R_k \in \mathbb{R}^{m\times m}$ are covariance matrices. In Figure 10 the Kalman filter structure is shown.

From Figure 10 we find the following equations for the Kalman filter

$$\begin{aligned} \bar{x}_{k+1} &= A_k \hat{x}_k + B_k u_k + E_k d_k \\ \hat{x}_k &= \bar{x}_k + K_k \cdot \varepsilon_k = \bar{x}_k + K_k\left(y_k - \bar{y}_k\right) \\ \bar{y}_k &= C_k \bar{x}_k + D_k u_k + F_k d_k. \end{aligned} \tag{72}$$

Define the covariance matrices

$$Z_k = E\left[\left(x_k - \hat{x}_k\right)\left(x_k - \hat{x}_k\right)^T\right] \text{ and } P_k = E\left[\left(x_k - \bar{x}_k\right)\left(x_k - \bar{x}_k\right)^T\right], \tag{73}$$

where

$$\begin{aligned} \hat{x}_k &= \bar{x}_k + K_k\left(y_k - \bar{y}_k\right) \\ &= \bar{x}_k + K_k\left(C_k x_k + D_k u_k + F_k d_k + v_k - C_k \bar{x}_k - D_k u_k - F_k d_k\right) \\ &= \bar{x}_k + K_k C_k x_k + K_k v_k - K_k C_k \bar{x}_k \\ &= (I - K_k C_k)\bar{x}_k + K_k C_k x_k + K_k v_k. \end{aligned}$$

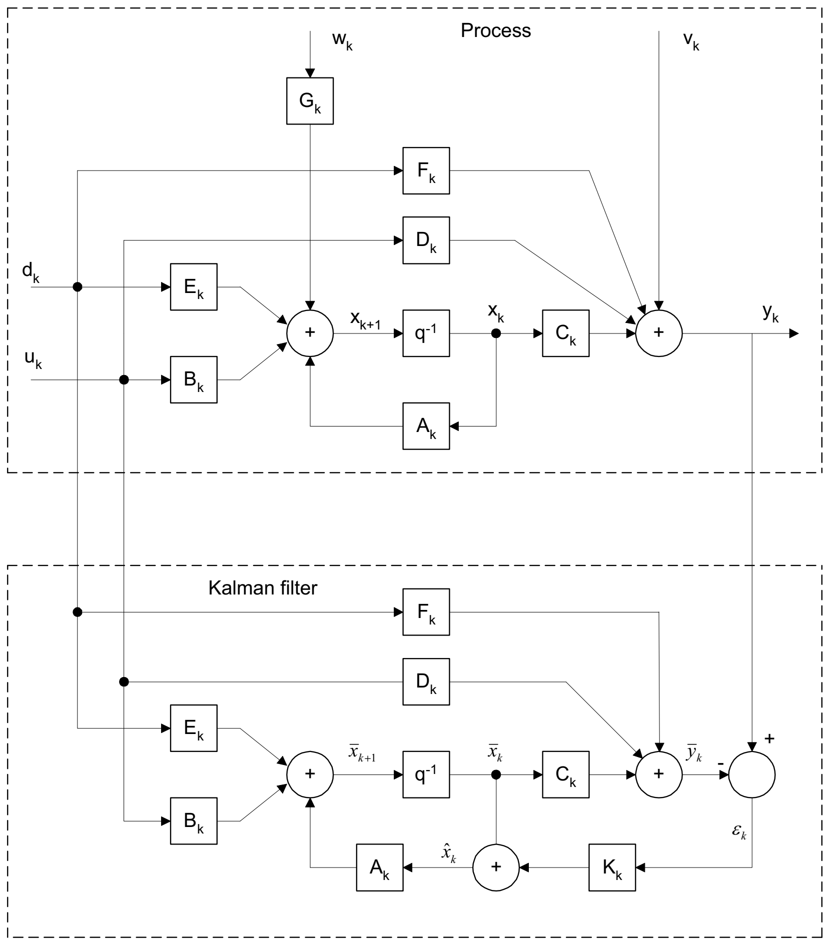

Figure 10: Structure of Kalman filter for linear time variant process with colored noise.

Then

$$
\begin{aligned}
Z_k &= E\left[(x_k - (I - K_kC_k)\bar{x}_k - K_kC_kx_k - K_kv_k)\,(\cdot)^T\right] \\
&= E\left[((I - K_kC_k)(x_k - \bar{x}_k) - K_kv_k)\,((I - K_kC_k)(x_k - \bar{x}_k) - K_kv_k)^T\right] \\
&= E\left[(I - K_kC_k)(x_k - \bar{x}_k)(x_k - \bar{x}_k)^T(I - K_kC_k)^T - (I - K_kC_k)(x_k - \bar{x}_k)v_k^TK_k^T\right. \\
&\left. -K_kv_k(x_k - \bar{x}_k)^T(I - K_kC_k)^T + K_kv_kv_k^TK_k^T\right] \\
&= (I - K_kC_k)E\left[(x_k - \bar{x}_k)(x_k - \bar{x}_k)^T\right](I - K_kC_k)^T - (I - K_kC_k)E\left[(x_k - \bar{x}_k)v_k^T\right]K_k^T \\
&- K_kE\left[v_k(x_k - \bar{x}_k)^T\right](I - K_kC_k)^T + K_kE\left[v_kv_k^T\right]K_k^T \\
&= (I - K_kC_k)P_k(I - K_kC_k)^T + K_kR_kK_k^T, \qquad (74)
\end{aligned}
$$

and where $E\left[(x_k - \bar{x}_k)v_k^T\right] = 0$ is a reasonable assumption since

$$
\begin{aligned}
x_k - \bar{x}_k &= A_{k-1}x_{k-1} + G_{k-1}w_{k-1} - A_{k-1}\bar{x}_{k-1} - K_{k-1}(y_{k-1} - \bar{y}_{k-1}) \\
&= A_{k-1}x_{k-1} + G_{k-1}w_{k-1} - A_{k-1}\bar{x}_{k-1} - K_{k-1}(C_{k-1}x_{k-1} + v_{k-1} - C_{k-1}\bar{x}_{k-1}) \\
&= (A_{k-1} - K_{k-1}C_{k-1})(x_{k-1} - \bar{x}_{k-1}) + G_{k-1}w_{k-1} - K_{k-1}v_{k-1},
\end{aligned}
$$

where we see that the state estimation error $x_k - \bar{x}_k$ only depends on past noise sequences.

We now seek to find the gain matrix K_k which minimizes the covariance Z_k, noting that $\min(Z_k)$ implies $\min(\mathrm{trace}(Z_k))$ and that for a symmetric matrix B we have the following rule

$$
\frac{\partial}{\partial A}\mathrm{trace}(ABA^T) = 2AB.
$$

Then the optimal gain matrix, or the Kalman filter gain matrix, is

$$
\begin{gathered}
\frac{\partial}{\partial K_k}(Z_k) = 0 \\
\Downarrow \\
K_k = P_kC_k^T\left(R_k + C_kP_kC_k^T\right)^{-1}. \qquad (75)
\end{gathered}
$$

Using the Kalman filter on-line we must find P_{k+1}

$$
\begin{aligned}
P_{k+1} &= E\left[(x_{k+1} - \bar{x}_{k+1})\,(x_{k+1} - \bar{x}_{k+1})^T\right] \\
&= E\left[(A_kx_k + G_kw_k - A_k\hat{x}_k)\,(A_kx_k + G_kw_k - A_k\hat{x}_k)^T\right] \\
&= E\left[(A_k(x_k - \hat{x}_k) + G_kw_k)\,(A_k(x_k - \hat{x}_k) + G_kw_k)^T\right] \\
&= E\left[A_k(x_k - \hat{x}_k)(x_k - \hat{x}_k)^TA_k^T + A_k(x_k - \hat{x}_k)w_k^TG_k^T\right. \\
&\quad\left. +G_kw_k(x_k - \hat{x}_k)^TA_k^T + G_kw_kw_k^TG_k^T\right] \\
&= A_kE\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T\right]A_k^T + A_kE\left[(x_k - \hat{x}_k)w_k^T\right]G_k^T \\
&\quad + G_kE\left[w_k(x_k - \hat{x}_k)^T\right]A_k^T + G_kE\left[w_kw_k^T\right]G_k^T \qquad (76) \\
&= A_kZ_kA_k^T + G_kQ_kG_k^T. \qquad (77)
\end{aligned}
$$

We see from this last equation that the K_k that minimized Z_k also minimizes P_{k+1}.

D.2 Extended Kalman filter for nonlinear processes

Assume the process is nonlinear

$$\begin{aligned} x_{k+1} &= f(x_k, u_k, d_k) + G_k w_k \\ y_k &= g(x_k, u_k, d_k) + v_k, \end{aligned} \tag{78}$$

with noise characteristics as given by Equation 71. Then the extended Kalman filter algorithm can be written as (Ergon 2001)

1. At time k, given y_k, u_k, $\bar{x}_k$,P_k, $A_k = \frac{\partial f}{\partial x_k}|_s$ and $C_k = \frac{\partial g}{\partial x_k}|_s$.
2. Compute the Kalman filter gain matrix as given by Equation 75
$$K_k = P_k C_k^T \left(R_k + C_k P_k C_k^T \right)^{-1}.$$
3. Compute updated state estimate
$$\hat{x}_k = \bar{x}_k + K_k \left(y_k - g(\bar{x}_k, u_k, d_k) \right).$$
4. Compute updated covariance matrix for state error, as given by Equation 74
$$Z_k = (I - K_k C_k) P_k (I - K_k C_k)^T + K_k R_k K_k^T.$$
5. Compute state estimate at time $k+1$
$$\bar{x}_{k+1} = f(\hat{x}_k, u_k, d_k).$$
6. Compute covariance matrix for state error, as given by Equation 77
$$P_{k+1} = A_k Z_k A_k^T + G_k Q_k G_k^T.$$
7. Set $k \to k+1$ and go to step 2.

Note that when there is a direct input to output term in the model, and the model is used in a control loop, we must know the input u_k before we can estimate $\hat{x}_k$ or $\bar{x}_{k+1}$. This means that the controller must rely on the estimate $\bar{x}_{k|k-1}$ when computing the inputs u_k.

D.3 Offset free control by bias estimation

We will in this section show how the MPC and Kalman filter developed in previous sections and appendices can be redesigned to prevent steady state offset by estimating the bias and adding this to the model outputs.

Assume the following augmented process model

$$\begin{bmatrix} x_{k+1}^{\text{nonlin}} \\ p_{k+1}^{\text{nonlin}} \end{bmatrix} = \begin{bmatrix} f(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}) \\ p_k^{\text{nonlin}} \end{bmatrix} \tag{79}$$
$$y_k^{\text{nonlin}} = g(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}) + p_k^{\text{nonlin}},$$

where the bias p_k^{nonlin} is added to the process outputs. We will now review some of the stages in the algorithm described in Section 2.

D.3.1 Linearization

The linearization can be carried out on the augmented system, the same way as was done for the original model in Section 3. However, by studying the structure of the augmented system we may carry out the linearization in a more efficient way

$$\begin{bmatrix} x_{k+1}^{\text{lin}} \\ p_{k+1}^{\text{lin}} \end{bmatrix} = \begin{bmatrix} A_k^x & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x_k^{\text{lin}} \\ p_k^{\text{lin}} \end{bmatrix} + \begin{bmatrix} B_k \\ 0 \end{bmatrix} u_k^{\text{lin}} + \begin{bmatrix} E_k \\ 0 \end{bmatrix} d_k^{\text{lin}} \tag{80}$$
$$y_k^{\text{lin}} = \begin{bmatrix} C_k^x & I \end{bmatrix} \begin{bmatrix} x_k^{\text{lin}} \\ p_k^{\text{lin}} \end{bmatrix} + D_k u_k^{\text{lin}} + F_k d_k^{\text{lin}}.$$

The linearization can thus be carried out on the original non-augmented system, and augmentation of the A and C matrices can be done after the linearization.

D.3.2 Steady state values, shifting the model and variables

Assuming the steady state value of p to be known and equal to p_k^{lin} we may calculate the steady state values as follows

$$\min_{\bar{x}_{k+j}^s, \bar{u}_{k+j}^s} \left(\bar{y}_{k+j}^s - r_{k+j}^{\text{lin}}\right)^T Q_s \left(\bar{y}_{k+j}^s - r_{k+j}^{\text{lin}}\right), \forall j = 0, ..., N-1, \tag{81}$$

constrained by the steady state solution to the model

$$\begin{bmatrix} (I - A_k^x) & -B_k \\ -C_k^x & -D_k \end{bmatrix} \begin{bmatrix} \bar{x}_{k+j}^s \\ \bar{u}_{k+j}^s \end{bmatrix} = \begin{bmatrix} E_k d_{k+j}^{\text{lin}} \\ F_k d_{k+j}^{\text{lin}} - \bar{y}_{k+j}^s + p_k^{\text{lin}} \end{bmatrix}, \forall j = 0, ..., N-1, \tag{82}$$

and

$$\bar{u}_{k+j}^{\min} \leq \bar{u}_{k+j}^s + u_{k-1}^{\text{nonlin}} \leq \bar{u}_{k+j}^{\max}, \forall j = 0, ..., N-1 \tag{83}$$
$$\bar{y}_{k+j}^{\min} \leq \bar{y}_{k+j}^s + y_{k-1}^{\text{nonlin}} \leq \bar{y}_{k+j}^{\max}, \forall j = 0, ..., N-1,$$

and

$$\bar{y}_{k+j}^s = C_k^x \bar{x}_{k+j}^s + p_{k+j}^{\text{lin}} + D_k \bar{u}_{k+j}^s + F_k d_{k+j}^{\text{lin}}, \forall j = 0, ..., N-1, \tag{84}$$

When shifting the augmented model to the steady state values at time $N-1$, all bias terms are set to zero since we assume the bias is constant into the future. Thus, there is no point in using the augmented model in the MPC.

D.3.3 Optimization

The optimization is carried out the same way as was done in Section 4, due to the fact that the augmented model reduces to the original model when assuming a constant bias term, and shifting it to the steady state values at time $N-1$.

D.3.4 Estimating the states

The process model is given by the augmented model of Equation 79, and we assume the real process is given by

$$\begin{aligned} \tilde{x}_{k+1}^{\text{nonlin}} &= \tilde{f}(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, p_k^{\text{nonlin}}) + \tilde{G}_k \tilde{w}_k \\ y_k^{\text{nonlin}} &= g(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}) + p_k^{\text{nonlin}} + v_k, \end{aligned} \tag{85}$$

where

$$\tilde{x}_{k+1}^{\text{nonlin}} = \begin{bmatrix} x_{k+1}^{\text{nonlin}} \\ p_{k+1}^{\text{nonlin}} \end{bmatrix} \tag{86}$$

$$\tilde{f}(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, p_k^{\text{nonlin}}) = \begin{bmatrix} f(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}) \\ p_k^{\text{nonlin}} \end{bmatrix}$$

$$\tilde{G}_k = \begin{bmatrix} G_k^x & 0 \\ 0 & G_k^p \end{bmatrix} \tag{87}$$

$$\tilde{w}_k = \begin{bmatrix} w_k^x \\ w_k^p \end{bmatrix}, \tag{88}$$

and where the noise characteristics are as given by Equation 71 (with tilde above appropriate elements).

We study the covariance matrices

$$\tilde{Z}_k = E\left[\left(\tilde{x}_k^{\text{nonlin}} - \tilde{\hat{x}}^{\text{nonlin}}\right)\left(\tilde{x}_k^{\text{nonlin}} - \tilde{\hat{x}}^{\text{nonlin}}\right)^T\right] \tag{89}$$

$$\tilde{P}_k = E\left[\left(\tilde{x}_k^{\text{nonlin}} - \tilde{\bar{x}}^{\text{nonlin}}\right)\left(\tilde{x}_k^{\text{nonlin}} - \tilde{\bar{x}}^{\text{nonlin}}\right)^T\right], \tag{90}$$

where

$$\begin{aligned} \tilde{x}_k^{\text{nonlin}} &= \begin{bmatrix} x_k^{\text{nonlin}} \\ p_k^{\text{nonlin}} \end{bmatrix} \\ \tilde{\hat{x}}^{\text{nonlin}} &= \begin{bmatrix} \hat{x}_k^{\text{nonlin}} \\ \hat{p}_k^{\text{nonlin}} \end{bmatrix} \\ \tilde{\bar{x}}^{\text{nonlin}} &= \begin{bmatrix} \bar{x}_k^{\text{nonlin}} \\ \bar{p}_k^{\text{nonlin}} \end{bmatrix} \end{aligned} \tag{91}$$

For a linearized model we then have

$$\begin{aligned}\hat{\tilde{x}}^{\text{lin}} &= \bar{\tilde{x}}^{\text{lin}} + \tilde{K}_k \left(y_k^{\text{lin}} - \bar{y}_k^{\text{lin}}\right) \\ &= \bar{\tilde{x}}^{\text{lin}} + \tilde{K}_k \left(\tilde{C}_k \tilde{x}_k^{\text{lin}} + D_k u_k^{\text{lin}} + F_k d_k^{\text{lin}} + v_k - \tilde{C}_k \bar{\tilde{x}}^{\text{lin}} - D_k u_k^{\text{lin}} - F_k d_k^{\text{lin}}\right) \\ &= \bar{\tilde{x}}^{\text{lin}} + \tilde{K}_k \left(\tilde{C}_k \tilde{x}_k^{\text{lin}} + v_k - \tilde{C}_k \bar{\tilde{x}}^{\text{lin}}\right) \\ &= \left(I - \tilde{K}_k \tilde{C}_k\right) \bar{\tilde{x}}^{\text{lin}} + \tilde{K}_k \tilde{C}_k \tilde{x}_k^{\text{lin}} + \tilde{K}_k v_k\end{aligned} \tag{92}$$

where

$$\tilde{K}_k = \begin{bmatrix} K_k^x \\ K_k^p \end{bmatrix}, \text{ and } \tilde{C}_k = \begin{bmatrix} C_k & I \end{bmatrix} \tag{93}$$

Then

$$\begin{aligned}\tilde{Z}_k &= E\left[\left(\tilde{x}_k^{\text{lin}} - \left(I - \tilde{K}_k \tilde{C}_k\right) \bar{\tilde{x}}^{\text{lin}} - \tilde{K}_k \tilde{C}_k \tilde{x}_k^{\text{lin}} - \tilde{K}_k v_k\right) (\cdot)^T\right] \\ &= E\left[\left(\left(I - \tilde{K}_k \tilde{C}_k\right) \left(\tilde{x}_k^{\text{lin}} - \bar{\tilde{x}}^{\text{lin}}\right) - \tilde{K}_k v_k\right) (\cdot)^T\right] \\ &= \left(I - \tilde{K}_k \tilde{C}_k\right) \tilde{P}_k \left(I - \tilde{K}_k \tilde{C}_k\right)^T + \tilde{K}_k R_k^T \tilde{K}_k^T,\end{aligned} \tag{94}$$

and

$$\tilde{K}_k = \tilde{P}_k^T \tilde{C}_k^T \left(R_k + \tilde{C}_k \tilde{P}_k^T \tilde{C}_k^T\right)^{-1}, \tag{95}$$

and finally

$$\tilde{P}_{k+1} = \tilde{A}_k \tilde{Z}_k \tilde{A}_k^T + \tilde{G}_k \tilde{Q}_k \tilde{G}_k^T, \tag{96}$$

where

$$\tilde{Q}_k = E\left[\tilde{w}_k \tilde{w}_k^T\right]. \tag{97}$$

Then the augmented Kalman filter algorithm can be written as

1. At time k, given y_k^{nonlin}, u_k^{nonlin},$\bar{\tilde{x}}_k^{\text{nonlin}}$, C_k, A_k and $\tilde{P}_k$. Augment model matrices
$$\tilde{A}_k = \begin{bmatrix} A_k & 0 \\ 0 & I \end{bmatrix}, \text{ and } \tilde{C}_k = \begin{bmatrix} C_k & I \end{bmatrix}$$
2. Compute the Kalman filter gain matrix
$$\tilde{K}_k = \tilde{P}_k^T \tilde{C}_k^T \left(R_k + \tilde{C}_k \tilde{P}_k^T \tilde{C}_k^T\right)^{-1}.$$
3. Compute updated state estimate
$$\hat{\tilde{x}}^{\text{nonlin}} = \bar{\tilde{x}}^{\text{nonlin}} + \tilde{K}_k \left(y_k^{\text{nonlin}} - g(\bar{x}_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}) - \bar{p}_k^{\text{nonlin}}\right)$$
4. Compute updated covariance matrix for state error
$$\tilde{Z}_k = \left(I - \tilde{K}_k \tilde{C}_k\right) \tilde{P}_k \left(I - \tilde{K}_k \tilde{C}_k\right)^T + \tilde{K}_k R_k^T \tilde{K}_k^T.$$

5. Compute state estimate at time $k+1$

$$\widetilde{\widetilde{x}}_{k+1}^{\text{nonlin}} = \widetilde{\widehat{f}}(\hat{x}_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, \hat{p}_k^{\text{nonlin}}) = \begin{bmatrix} f(\hat{x}_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}) \\ \hat{p}_k^{\text{nonlin}} \end{bmatrix}.$$

6. Compute covariance matrix for state error

$$\tilde{P}_{k+1} = \tilde{A}_k \tilde{Z}_k \tilde{A}_k^T + \tilde{G}_k \tilde{Q}_k \tilde{G}_k^T.$$

7. Set $k \to k+1$ and go to step 2.

D.4 Online parameter estimation by augmented Kalman filter

An alternative to bias estimation to obtain offset free control may be to estimate parameters and biases in the model on-line. The estimation can be done by augmenting the state vector by parameters and biases that we wish to estimate. The procedure is similar, but not equal, to the procedure in the previous section where only the (output) bias was estimated. We will in this section show how the MPC and Kalman filter developed in previous sections can be redesigned to prevent steady state offset by estimating parameters and/or biases in the model.

Assume the following augmented process model

$$\begin{aligned} \begin{bmatrix} x_{k+1}^{\text{nonlin}} \\ \theta_{k+1}^{\text{nonlin}} \end{bmatrix} &= \begin{bmatrix} f(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, \theta_k^{\text{nonlin}}) \\ \theta_k^{\text{nonlin}} \end{bmatrix} \\ y_k^{\text{nonlin}} &= g(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, \theta_k^{\text{nonlin}}), \end{aligned} \tag{98}$$

We will now review some of the stages in the algorithm described in Section 2.

D.4.1 Linearization

The linearization can be carried out on the augmented system, the same way as was done for the original model in Section 3. However, by studying the structure of the augmented system we may carry out the linearization in a more efficient way

$$\begin{aligned} \begin{bmatrix} x_{k+1}^{\text{lin}} \\ \theta_{k+1}^{\text{lin}} \end{bmatrix} &= \begin{bmatrix} A_k^x & A_k^\theta \\ 0 & I \end{bmatrix} \begin{bmatrix} x_k^{\text{lin}} \\ \theta_k^{\text{lin}} \end{bmatrix} + \begin{bmatrix} B_k \\ 0 \end{bmatrix} u_k^{\text{lin}} + \begin{bmatrix} E_k \\ 0 \end{bmatrix} d_k^{\text{lin}} \\ y_k^{\text{lin}} &= \begin{bmatrix} C_k^x & C_k^\theta \end{bmatrix} \begin{bmatrix} x_k^{\text{lin}} \\ \theta_k^{\text{lin}} \end{bmatrix} + D_k u_k^{\text{lin}} + F_k d_k^{\text{lin}}. \end{aligned} \tag{99}$$

D.4.2 Steady state values, shifting the model and variables

Assuming the steady state value of θ to be known and equal to θ_k^{lin} we may calculate the steady state values as follows

$$\min_{\bar{x}_{k+j}^s, \bar{u}_{k+j}^s} \left(\bar{y}_{k+j}^s - r_{k+j}^{\text{lin}}\right)^T Q_s \left(\bar{y}_{k+j}^s - r_{k+j}^{\text{lin}}\right), \forall j = 0, ..., N-1, \tag{100}$$

constrained by the steady state solution to the model

$$\begin{bmatrix} (I - A_k^x) & -B_k \\ -C_k^x & -D_k \end{bmatrix} \begin{bmatrix} \bar{x}_{k+j}^s \\ \bar{u}_{k+j}^s \end{bmatrix} = \begin{bmatrix} A_k^\theta \theta_k^{\text{lin}} + E_k d_{k+j}^{\text{lin}} \\ C_k^\theta \theta_k^{\text{lin}} + F_k d_{k+j}^{\text{lin}} - \bar{y}_{k+j}^s \end{bmatrix}, \forall j = 0, ..., N-1, \tag{101}$$

and

$$\begin{aligned} \bar{u}_{k+j}^{\min} &\leq \bar{u}_{k+j}^s + u_{k-1}^{\text{nonlin}} \leq \bar{u}_{k+j}^{\max}, \forall j = 0, ..., N-1 \\ \bar{y}_{k+j}^{\min} &\leq \bar{y}_{k+j}^s + y_{k-1}^{\text{nonlin}} \leq \bar{y}_{k+j}^{\max}, \forall j = 0, ..., N-1, \end{aligned} \tag{102}$$

and

$$\bar{y}_{k+j}^s = C_k^x \bar{x}_{k+j}^s + C_k^\theta \theta_{k+j}^{\text{lin}} + D_k \bar{u}_{k+j}^s + F_k d_{k+j}^{\text{lin}}, \forall j = 0, ..., N-1, \tag{103}$$

When shifting the augmented model to the steady state values at time $N-1$, all augmented states are set to zero since we assume the parameter values are constant into the future. Thus, there is no point in using the augmented model in the MPC.

D.4.3 Optimization

The optimization is carried out the same way as was done in Section 4, due to the fact that the augmented model reduces to the original model when assuming constant parameters and biases into the future, and when the model is shifted to the steady state values at time $N-1$.

D.4.4 Estimating the states

The process model is given by the augmented model of Equation 98, and we assume the real process is given by

$$\begin{aligned} \tilde{x}_{k+1}^{\text{nonlin}} &= \tilde{f}(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, \theta_k^{\text{nonlin}}) + \tilde{G}_k \tilde{w}_k \\ y_k^{\text{nonlin}} &= g(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, \theta_k^{\text{nonlin}}) + v_k, \end{aligned} \tag{104}$$

where

$$\begin{aligned} \tilde{x}_{k+1}^{\text{nonlin}} &= \begin{bmatrix} x_{k+1}^{\text{nonlin}} \\ \theta_{k+1}^{\text{nonlin}} \end{bmatrix} \\ \tilde{f}(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, \theta_k^{\text{nonlin}}) &= \begin{bmatrix} f(x_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, \theta_k^{\text{nonlin}}) \\ \theta_k^{\text{nonlin}} \end{bmatrix} \\ \tilde{G}_k &= \begin{bmatrix} G_k^x & 0 \\ 0 & G_k^\theta \end{bmatrix} \\ \tilde{w}_k &= \begin{bmatrix} w_k^x \\ w_k^\theta \end{bmatrix}, \end{aligned} \tag{105}$$

and where the noise characteristics are as given by Equation 71 (with tilde above appropriate elements).

We study the covariance matrices

$$\tilde{Z}_k = E\left[\left(\tilde{x}_k^{\text{nonlin}} - \widetilde{\hat{x}}^{\text{nonlin}}\right)\left(\tilde{x}_k^{\text{nonlin}} - \widetilde{\hat{x}}^{\text{nonlin}}\right)^T\right] \tag{106}$$

$$\tilde{P}_k = E\left[\left(\tilde{x}_k^{\text{nonlin}} - \widetilde{\bar{x}}^{\text{nonlin}}\right)\left(\tilde{x}_k^{\text{nonlin}} - \widetilde{\bar{x}}^{\text{nonlin}}\right)^T\right], \tag{107}$$

where

$$\tilde{x}_k^{\text{nonlin}} = \begin{bmatrix} x_k^{\text{nonlin}} \\ \theta_k^{\text{nonlin}} \end{bmatrix} \tag{108}$$

$$\widetilde{\hat{x}}^{\text{nonlin}} = \begin{bmatrix} \hat{x}_k^{\text{nonlin}} \\ \hat{\theta}_k^{\text{nonlin}} \end{bmatrix}$$

$$\widetilde{\bar{x}}^{\text{nonlin}} = \begin{bmatrix} \bar{x}_k^{\text{nonlin}} \\ \bar{\theta}_k^{\text{nonlin}} \end{bmatrix}$$

For a linearized model we then have

$$\begin{aligned} \widetilde{\hat{x}}^{\text{lin}} &= \widetilde{\bar{x}}^{\text{lin}} + \tilde{K}_k\left(y_k^{\text{lin}} - \bar{y}_k^{\text{lin}}\right) \\ &= \widetilde{\bar{x}}^{\text{lin}} + \tilde{K}_k\left(\tilde{C}_k\tilde{x}_k^{\text{lin}} + D_k u_k^{\text{lin}} + F_k d_k^{\text{lin}} + v_k - \tilde{C}_k\widetilde{\bar{x}}^{\text{lin}} - D_k u_k^{\text{lin}} - F_k d_k^{\text{lin}}\right) \\ &= \widetilde{\bar{x}}^{\text{lin}} + \tilde{K}_k\left(\tilde{C}_k\tilde{x}_k^{\text{lin}} + v_k - \tilde{C}_k\widetilde{\bar{x}}^{\text{lin}}\right) \\ &= \left(I - \tilde{K}_k\tilde{C}_k\right)\widetilde{\bar{x}}^{\text{lin}} + \tilde{K}_k\tilde{C}_k\tilde{x}_k^{\text{lin}} + \tilde{K}_k v_k \end{aligned} \tag{109}$$

where

$$\tilde{K}_k = \begin{bmatrix} K_k^x \\ K_k^\theta \end{bmatrix}, \text{ and } \tilde{C}_k = \begin{bmatrix} C_k^x & C_k^\theta \end{bmatrix} \tag{110}$$

Then

$$\begin{aligned} \tilde{Z}_k &= E\left[\left(\tilde{x}_k^{\text{lin}} - \left(I - \tilde{K}_k\tilde{C}_k\right)\widetilde{\bar{x}}^{\text{lin}} - \tilde{K}_k\tilde{C}_k\tilde{x}_k^{\text{lin}} - \tilde{K}_k v_k\right)(\cdot)^T\right] \\ &= E\left[\left(\left(I - \tilde{K}_k\tilde{C}_k\right)\left(\tilde{x}_k^{\text{lin}} - \widetilde{\bar{x}}^{\text{lin}}\right) - \tilde{K}_k v_k\right)(\cdot)^T\right] \\ &= \left(I - \tilde{K}_k\tilde{C}_k\right)\tilde{P}_k\left(I - \tilde{K}_k\tilde{C}_k\right)^T + \tilde{K}_k R_k^T \tilde{K}_k^T, \end{aligned} \tag{111}$$

and

$$\tilde{K}_k = \tilde{P}_k^T\tilde{C}_k^T\left(R_k + \tilde{C}_k\tilde{P}_k^T\tilde{C}_k^T\right)^{-1}, \tag{112}$$

and finally

$$\tilde{P}_{k+1} = \tilde{A}_k\tilde{Z}_k\tilde{A}_k^T + \tilde{G}_k\tilde{Q}_k\tilde{G}_k^T, \tag{113}$$

where

$$\tilde{A}_k = \begin{bmatrix} A_k^x & A_k^\theta \\ 0 & I \end{bmatrix} \tag{114}$$

$$\tilde{Q}_k = E\left[\tilde{w}_k\tilde{w}_k^T\right]. \tag{115}$$

Then the augmented Kalman filter algorithm can be written as

1. At time k, given y_k^{nonlin}, u_k^{nonlin}, $\tilde{\bar{x}}_k^{\text{nonlin}}$, C_k^x, C_k^θ, A_k^x, A_k^θ and $\tilde{P}_k$. Augment model matrices

$$\tilde{A}_k = \begin{bmatrix} A_k^x & A_k^\theta \\ 0 & I \end{bmatrix}, \text{ and } \tilde{C}_k = \begin{bmatrix} C_k^x & C_k^\theta \end{bmatrix}$$

2. Compute the Kalman filter gain matrix

$$\tilde{K}_k = \tilde{P}_k^T \tilde{C}_k^T \left(R_k + \tilde{C}_k \tilde{P}_k^T \tilde{C}_k^T \right)^{-1}.$$

3. Compute updated state estimate

$$\tilde{\hat{x}}^{\text{nonlin}} = \tilde{\bar{x}}^{\text{nonlin}} + \tilde{K}_k \left(y_k^{\text{nonlin}} - g(\bar{x}_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, \bar{\theta}_k^{\text{nonlin}}) \right)$$

4. Compute updated covariance matrix for state error

$$\tilde{Z}_k = \left(I - \tilde{K}_k \tilde{C}_k \right) \tilde{P}_k \left(I - \tilde{K}_k \tilde{C}_k \right)^T + \tilde{K}_k R_k^T \tilde{K}_k^T.$$

5. Compute state estimate at time $k+1$

$$\tilde{\bar{x}}_{k+1}^{\text{nonlin}} = \tilde{\hat{f}}(\hat{x}_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, \hat{\theta}_k^{\text{nonlin}}) = \begin{bmatrix} f(\hat{x}_k^{\text{nonlin}}, u_k^{\text{nonlin}}, d_k^{\text{nonlin}}, \hat{\theta}_k^{\text{nonlin}}) \\ \hat{\theta}_k^{\text{nonlin}} \end{bmatrix}.$$

6. Compute covariance matrix for state error

$$\tilde{P}_{k+1} = \tilde{A}_k \tilde{Z}_k \tilde{A}_k^T + \tilde{G}_k \tilde{Q}_k \tilde{G}_k^T.$$

7. Set $k \rightarrow k+1$ and go to step 2.

References

Ahn, S., Park, M. & Rhee, H. (1999), 'Extended Kalman filter-based nonlinear model predictive control for a continuous MMA polymerization reactor', *Ind. Eng. Chem. Res.* **38**(10), 3942–3949.

Amin, J., Mehra, R. K. & Arambel, P. (2001), Coordinated dynamic positioning of a multi-platform mobile offshore base using nonlinear model predictive control, *in* 'Eleventh (2001) International Offshore and Polar Engineering Conference'. Stavanger, Norway, June 17-22, 2001.

Austin, P., Mack, J., Lovett, D., Wright, M. & Terry, M. (2002), Improved wet end stability of a paper machine using model predictive control, *in* 'Control Systems 2002', STFi and SPCI, pp. 80–84. June 3-5, 2002, Stockholm, Sweden.

Badgwell, T. A. & Qin, S. J. (2001), Review of nonlinear model predictive control applications, *in* 'IEE Control Engineering Series', Vol. 61, pp. 3–32.

Bown, R. (1996), Physical and chemical aspects of the use of fillers in paper, *in* J. Roberts, ed., 'Paper Chemistry', 2 edn, Chapman and Hall, chapter 11.

Dennis, J. E. & Schnabel, R. B. (1996), *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Classics in Applied Mathematics, 16, SIAM, Philadelphia.

Ergon, R. (2001), *Introduction to Kalman Filtering with some Industrial Application Areas*, Lecture notes, Telemark University College, Norway.

Gattu, G. & Zafiriou, E. (1992), 'Nonlinear quadratic dynamic matrix control with state estimation', *Ind. Eng. Chem. Res.* **31**(4), 1096–1104.

Gelb, A., ed. (1974), *Applied Optimal Estimation*, The M.I.T. Press.

Gill, P. E., Hammarling, S. J., Murray, W., Saunders, M. A. & Wright, M. H. (1986), User's guide for lssol (version 1.0): A fortran package for constrained linear least-squares and convex quadratic programming, Technical Report SOL 86-1, Systems Optimization Laboratory (SOL), department of operations research, Stanford University.

Gill, P. E., Murray, W. & Saunders, M. A. (1997), 'User's guide for sqopt 5.3: A Fortran package for large-scale linear and quadratic programming'. (Draft, October 1997).

Gill, P. E., Murray, W. & Wright, M. H. (1981), *Practical Optimization*, Academic Press.

Glemmestad, B., Ertler, G. & Hillestad, M. (2002), Advanced process control in a Borstar PP plant, *in* 'ECOREPII, 2nd European Conference on the Reaction Engineering of Polyolefins. Lyon, France, 1-4 July 2002'.

Griewank, A. (2000), *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, SIAM, Philadelphia.

Griewank, A. & Corliss, G. F., eds (1991), *Automatic Differentiation of Algorithms: Theory, Implementation, and Application*, SIAM, Philadelphia. Proceedings of the first SIAM Workshop on automatic differentiation, held in Breckenridge, Colorado, January 6-8, 1991.

Hauge, T. A. & Lie, B. (2002), 'Paper machine modeling at Norske Skog Saugbrugs: A mechanistic approach', *Modeling, Identification and Control* **23**(1), 27–52.

Hillestad, M. & Andersen, K. S. (1994), Model predictive control for grade transitions of a polypropylene reactor, *in* 'ESCAPE4, 4th European Symposioum on Computer Aided Process Engineering, Dublin, March 1994'.

Holmström, K. (2001), The TOMLAB optimization environment v3.0 user's guide, Technical report, HKH MatrisAnalys AB.

Kosonen, M., Fu, C., Nuyan, S., Kuusisto, R. & Huhtelin, T. (2002), Narrowing the gap between theory and practice: Mill experiences with multivariable predictive control, *in* 'Control Systems 2002', STFi and SPCI, pp. 54–59. June 3-5, 2002, Stockholm, Sweden.

Lang, D., Tian, L., Kuusisto, R. & Rantala, T. (1998), Multivariable predictive control for the wet end, *in* 'Measurement and Control of Papermaking, Edinburgh, Scotland', Pira International.

Lee, J. H. & Ricker, N. L. (1994), 'Extended Kalman filter based nonlinear model predictive control', *Ind. Eng. Chem. Res.* **33**(6), 1530–1541.

Lee, K., Lee, J. H., Yang, D. R. & Mahoney, A. W. (2002), 'Integrated run-to-run and on-line model-based control of particle size distribution for a semi-batch precipitation reactor', *Computers and chemical engineering* **26**(7-8), 1117–1131.

Lie, B. (1995), Kompendium i prosessmodellering, Lecture notes, Telemark University College. (in norwegian).

Lie, B., Dueñas Díez, M. & Hauge, T. A. (2002), A comparison of implementation strategies for MPC, *in* 'International Symposium on Advanced Control of Industrial Processes'. June 10-11, 2002, Kumamoto, Japan.

Mack, J., Lovett, D., Austin, P., Wright, M. & Terry, M. (2001), Connoisseur model predictive control of a paper machine's wet-end., *in* 'ICHEME - Advances in Process Control'. York, UK, 24-25 September 2001.

McQuillin, D. L. & Huizinga, P. W. (1995), Reducing grade change time through the use of predictive multi-variable cotrol, *in* 'ECOPAPERTECH', pp. 73–82. Helsinki, Finland.

Muske, K. R. & Badgwell, T. A. (2002), 'Disturbance modeling for offset-free linear model predictive control', *Journal of Process Control* **12**, 617–632.

Muske, K. R. & Rawlings, J. B. (1993), 'Model predictive control with linear models', *AIChE Journal* **39**(2), 262–287.

Park, M. J., Hur, S. M. & Rhee, H. K. (2002), 'Online estimation and control of polymer quality in a copolymerization reactor', *AIChE Journal* **48**(5), 1013–1021.

Prasad, V., Schley, M., Russo, L. P. & Bequette, B. W. (2002), 'Product property and production rate control of styrene polymerization', *Journal of Process Control* **12**(3), 353–372.

Qin, S. J. & Badgwell, T. A. (1997), An overview of industrial model predictive control technology, *in* 'AIChE Symposium Series, No. 316', pp. 232–256.

Qin, S. J. & Badgwell, T. A. (1998), An overview of nonlinear model predictive control applications, *in* 'Nonlinear Model Predictive Control Workshop - Assessment and Future Directions'. Ascona, Switzerland, June 3-5, 1998.

Rawlings, J. B. & Muske, K. R. (1993), 'The stability of constrained receding horizon control', *IEEE Transactions on Automatic Control* **38**(10), 1512–1516.

Schei, T. S. & Singstad, P. (1998), Nonlinear model predictive control of a batch polymerization process, *in* 'American Control Conference'. Philadelphia, Pennsylvania, June 1998.

Slora, R. (2001), Stabilization of the wet end at PM6. part 1: Developing controllers for the thick stock, Technical Report A-rapport RSL20001, Norske Skog Saugbrugs. (confidential and in Norwegian).

Solberg, I. (1988), A Modular Implementation of the Extended Kalman Filter with Application to a Crushing and Screening Circuit, PhD thesis, Norwegian University of Science and Technology.

Van de Ven, T. G. M. (1984), 'Theoretical aspects of drainage and retention of small particles on the fourdrinier', *Journal of Pulp and Paper Science* **10**(3), 57–63.