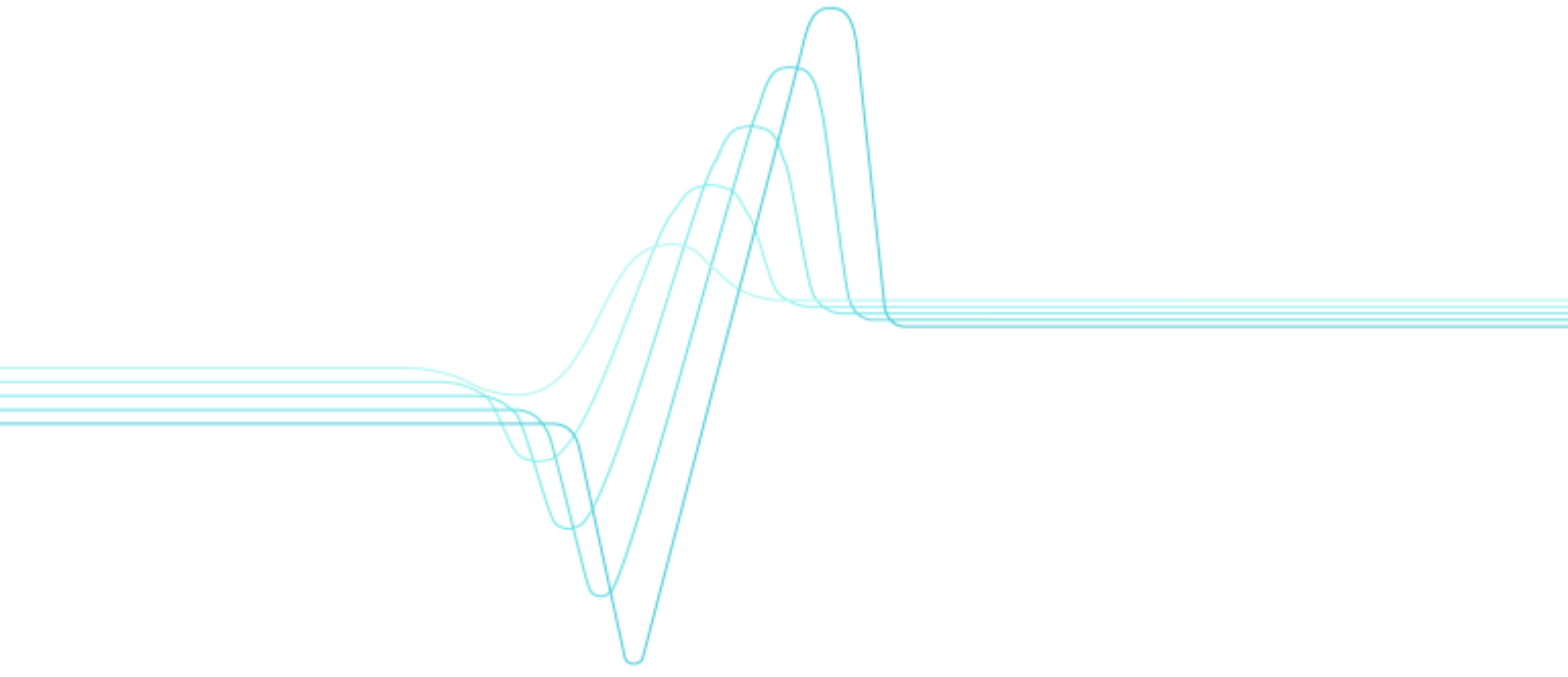
ESPOO 2005 VTT SYMPOSIUM 238

COST ACTION E36

Modelling and simulation in pulp and paper industry

Proceedings of Model Validation Workshop



VTT SYMPOSIUM 238

Keywords: COST Action E36, parameter identification, maintenance of models, dynamic model validation, sensitivity analysis, uncertainty analysis, pulp and paper industry, wet-end chemistry, runnability, emissions reduction

COST ACTION E36

Modelling and simulation in pulp and paper industry

Proceedings of Model Validation Workshop

Espoo, Finland, 6 October, 2005

Edited by

Johannes Kappen, PTS, Germany
Jussi Manninen, VTT, Finland
Risto Ritala, Tampere University of Technology, Finland

Organised by

VTT & Tampere University of Technology, Finland

ISBN 951–38–6300–X (URL:http://www.vtt.fi/inf/pdf/)
ISSN 1455–0873 (URL: http://www.vtt.fi/inf/pdf/)

JULKAISIJA – UTGIVARE – PUBLISHER

VTT, Vuorimiehentie 5, PL 2000, 02044 VTT
puh. vaihde 020 722 111, faksi 020 722 4374

VTT, Bergsmansvägen 5, PB 2000, 02044 VTT
tel. växel 020 722 111, fax 020 722 4374

VTT Technical Research Centre of Finland
Vuorimiehentie 5, P.O.Box 2000, FI–02044 VTT, Finland
phone internat. +358 20 722 111, fax + 358 20 722 4374

VTT Prosessit, Lämpömiehenkuja 3 A, PL 1604, 02044 VTT
puh. vaihde 020 722 111, faksi 020 722 5000

VTT Processer, Värmemansgränden 3 A, PB 1604, 02044 VTT
tel. växel 020 722 111, fax 020 722 5000

VTT Processes, Lämpömiehenkuja 3 A, P.O.Box 1604, FI–02044 VTT, Finland
phone internat. + 358 20 722 111, fax + 358 20 722 5000

Preface

COST E36 is a European Action on modelling and simulation in the pulp and paper industry. This instrument has been established in order to promote the exchanges of scientific knowledge within the European Community.

The main objective of the Action is to promote the development and application of modelling and simulation techniques in pulp and paper manufacturing processes.

The main benefit will be a better understanding of the process mechanisms and their control loops. This will help to find solutions for currently pending problems in the paper industry: improving paper quality, optimising wet end chemistry, enhancing runnability and reducing emissions by improving process design, process monitoring and decision support during operation. In the long run, this action should also contribute to designing superior or new product properties.

COST E36 Workshop on Model Validation, 6 October, 2005

Programme

Chairman:	R. Ritala, Tampere University of Technology, Vice Chairman of COST E36; Finland
9:00–9:15	Opening remarks J. Manninen, VTT, Chair of Working Group A, Finland
9:15–9:45	Generalization and validation of mathematical models for P&P applications E. Dahlquist, Mälardalen University, Chair of Working Group B, Sweden
9:45–10:15	Model uncertainty and prediction capabilities B. Lie, Telemark University College, Norway
10:15–10:45	Development of a tool to improve forecast accuracy of dynamicsimulation models for paper process G. Kamml, J. Kappen, PTS, Germany
10:45–11:15	Break
11:15–11:45	Validation as a crucial step for improving robustness of models: Application to paper quality predictions A. Alonso, A. Blanco, C. Negro, Complutense Univeristy of Madrid I. Sao Piao, Holmen Paper Madrid, Spain
11:45–12:15	A Simulation Study of the Validity of Multivariate Autoregressive Modeling O. Saarela, KCL, Finland
12:15–13:30	Lunch
Chairman:	J. Kappen, PTS, Chairman of Cost E36, Germany
13:30–14:00	Dynamic validation of multivariate linear soft sensors with reference laboratory measurements K. Latva-Käyrä, R. Ritala, Tampere University of Technology, Finland
14:00–14.30	Fiber classification – model development and validation K. Villforth, S. Schabel, Darmstadt University of Technology, Germany
14:30–15:00	Break
15:00–15:30	3D simulation of handsheets made of different pulps R. Vincent, M. Rueff, C. Voillot, EFPG, France
15:30–16:00	Industrial results of optimization of stickyseparation through fine screening systems J. Valkama, S. Schabel, Darmstadt University of Technology, Germany
16:00–16:30	Accuracy of water/COD, heat and stickies simulations in P&P applications J. Kappen, PTS, Chairman of Cost E36, Germany
16:30–	Discussion J. Manninen, VTT, Chair of Working Group A, Finland

Contents

Preface 3

Programme 4

Physical models and their validation for pulp and paper applications
E. Dahlquist
Malardalen University, Sweden 7

Model uncertainty and prediction capabilities
B. Lie
Telemark University College, Norway 19

Development of a tool to improve the forecast accuracy of dynamic simulation models for the paper process
G. Kamml[1], H.-M. Voigt[2]
[1]Papiertechnische Stiftung PTS, Germany
[2]Gesellschaft zur Förderung angewandter Informatik GFaI, Germany 31

Validation as a crucial step for improving robustness of models: Application to paper quality predictions
A. Alonso[1], A. Blanco[1], C. Negro[1], I. San Pío[2]
[1]Complutense University of Madrid, Spain
[2]HPM, Spain 39

A simulation study of the validity of multivariate autoregressive modeling
O. Saarela
KCL, Finland 49

Dynamic validation of multivariate linear soft sensors with reference laboratory measurements
K. Latva-Käyrä, R. Ritala
Tampere University of Technology, Finland 57

Fibre classification – model development and validation
K. Villforth, S. Schabel
Darmstadt University of Technology, Germany 65

3-D simulation of handsheets made of different pulps
R. Vincent[1, 2], M. Rueff[2], C. Voillot[2]
[1] TEMBEC (R&D), France
[2] EFPG, LGP2, France 71

An approach for modeling and optimizing of industrial fine screening processes
J. Valkama, K. Villforth, S. Schabel
Darmstadt University of Technology, Germany 79

Validation in mill practice: accuracy of water/COD, heat and stickies simulations in P&P applications
J. Kappen, W. Dietz
Papiertechnische Stiftung PTS, Germany 87

Physical models and their validation for pulp and paper applications

Erik Dahlquist, Malardalen University, Vasteras, Sweden

Abstract:
In pulp and paper industry there are a huge number of different types of process equipments. There are digesters, screens, filters, hydro cyclones, presses, dryers, boilers etc. In many cases the equipment suppliers want to consider single equipments as unique, and thus a special model is needed. This can give hundreds or thousands of different models to keep updated in a simulation package, and when it comes to model validation and testing, it becomes impossible to handle in reality. If we instead try to identify the basic physical principles of each unit, we can start from that point and then just add on special "extra features". E.g. screens, filters and presses all have similar basic principles. In this way it is possible to reduce the number of modules for a complete integrated mill with power supply to some 20 models. These are then tuned with existing data from literature on "real performance", and configured with rough geometric data. For a screen for instance there is normally data available on separation efficiency when a certain mixture has been operated under certain conditions, like flow rate, geometric dimensions of the screen with the screen plate etc, but very seldom including fiber size distribution, as this has not been measured. The same is normally the case for e.g. cleaners and other types of centrifugal separation devices. We start from first principle models and then tune these for different operational conditions, where once the size distribution was measured and some variables varied. In another case study other variables or conditions were investigated. A generalization of the model can be done by combining all this information, covering to at least to some extent all the different operational conditions, and all fiber sizes and concentration ranges. By then only fitting the model with the existing simple mass balance data for a specific equipment, you can get a reasonably accurate model for all kind of operations for this and similar types of equipments. In this paper a description is made of how a number of different equipments have been modeled in this way. Tuning and the result of model validation for different operational modes is also shown.

Introduction:
The reasons for using a dynamic simulator system may be many, but mainly fall into seven categories of use:
1) To train operators before start up of a new mill or to introduce new employees to the process before starting operating the real plant.[Ryan et al 2002]
2) To use the dynamic simulator for optimization of the process, in the design phase of a rebuild or expansion of an existing mill, or for a completely new green field mill.
3) To test the DCS functionality together with the process before start up of the real plant.
4) To optimize an existing process line, by testing different ways of operation for process improvements.
5) On-line prediction and control of a process line or part of a process line [Persson et al 2003]

6) Use in combination with an optimization algorithm for production planning or on-line optimization and control [Dhak et al 2004]
7) For diagnostics purposes[Karlsson et al 2003]
8) For decision support [Bell et al 2004]

It should also be noticed, that a simulator system can be anything from a small test model of a specific equipment, where the engineering and programming effort can be a couple of hours to get it into operation, to huge systems with thousands or even tens of thousands of DCS signals connected to a model of a whole factory. Here perhaps 10.000 engineering hours or more are needed for the project. Therefore you have to be sure to understand what you are really out for, before starting to discuss costs and time schedules for a simulator project!

Operator training:
One reason for using an operator training simulator can be to reduce the amount of e.g. paper breaks or down time of the process during the start up phase of a new paper machine or pulp mill. As often people with very little experience of paper machine or pulp mill operations are hired for new green field mills, it will be very risky to start the new process, if the operators do not get good training in advance. Here it can be interesting to refer to a study done in the US on how much we remember of information we are fed with:
10 % of what we see
30% of what we see and hear simultaneously
70 % of what we also train at simultaneously
and close to 100 % of what we repeatedly train at.

This is the reason for training at a dynamic process simulator, as the operators can acquire a very good knowledge even before the actual start up of the mill.

If we just make some rough estimates on what benefits a training simulator can mean, we may assume 10 % higher production the first month after starting up a new process line . For a paper mill this would mean approximately 400 USD/ton* 1000 tpd * 30 days* 10% = 1.2 MUSD in earnings. One paper break can be worth some 100,000 USD in lost production etc.
This is no guarantee on earnings, but a qualified estimate, showing that this is not just a game for fun.

At MNI, Malaysian Newsprint Industries, the start up to full production was achieved in a 20% shorter time period than “normal” although operators with no previous experience from pulp and paper industries were recruited. They had been trained during a eight week time period, half day in the simulator, while the other half being out in the mill looking at the real hardware.[Ryan et al 2002], [NOPS 1990].

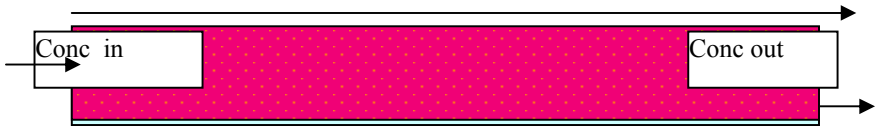
Simulator models
To build the simulator, we need a model for every single equipment in the plant. Some of these models can be very simple while others are very complex. In most cases we use a physical model as the basis and tune this with process data. This gives us a reasonably

good model over a large operational area, and it does not collapse even if we go significantly out of the normal operational area, which can be the case for a pure statistical model.

In reality we do not principally have that many different physical mechanisms in the major equipment. Most common is filtration or screening, where the mechanism is mechanical removal of particles on a mesh, wire or porous media.

Screens
The basic separation is done where fibres or particles are separated depending on mainly the ratio between the particle size and the pores they have to pass through. In filtering almost all fibres are separated, and then primarily the water is passing through the pores with a flow rate depending on the driving pressure as well as the pore size.

Often there is a concentration gradient, and thus it would be good to have the filter as in a thin vessel, and filtration/screening take place along the surface, with an ever increasing feed concentration.



The separation is also depending on the concentration, shear forces over the surface and the flow velocity through the filter (l/m2.h). In some cases also addition of chemicals (like in the stock prep) can increase flow rate or decrease "fines" passing through. This is also of interest to model.

The pressure drop over the filter or screen can be modelled as if it was a "fake valve", where the clean filter corresponds to the Admittance factor, which is the flow /h as a function of the driving pressure, when "the valve" is 100% open. Clogging of the screen due to different actions then results in a "valve opening" less than 100%.

In the pressure-flow network solver we use the relation $F_s = V*A*(\rho (P_1-P_2))^{\wedge 0.5}$, and the admittance A is calculated for normal flow (F_s) and the corresponding pressure drop (P_1-P_2). The valve opening V is the total open area of the screen and V= 100 for nominal pore area.. This is valid for pure water, which is used to calculate the admittance factor A (for max rotor speed as well).

The absolute **flow** through the screen plate is determined by the pressure in the feed.

The **concentration** of each fibre fraction in the **reject respectively accept** is determined from the ratio between the fibre and the pore opening, where a weighting is made between the hole pore area of the screen and the shape of the particle. A weighting can be done between length, width and thickness (three dimensions) , so that normally length

get a lower weighting factor than the other dimensions, especially in the slit type of screen.

By having maximum **rotor speed**, the actual screen holes area is almost as large as the nominal. The holes are clogged more and more, as the rotor velocity goes down, and a minimum value is set for zero rpm.

Concerning concentration, we have a function, that the **hole area** goes down as a function of the **concentration**, above a certain preset concentration. Above another conc, the screen is **totally clogged**.

The clogging of the screen is implemented by a ramp, where back flushing resets the open area of the openings to the original value, or to some lower values due to an irreversible clogging as one part of the total clogging.

When we configure a screen we first select holes or slots. The dimensions of these, the total hole area/m2 and the total screen area are inputs, and gives the total nominal hole area.

The ratio between the active hole area and the nominal is giving the average pore size and the pressure drop over the screen.

Active hole area=total nominal hole area*rpm_par*clog_par*conc_par

where

rpm_par=rpm_COF(6)+ ((1-COF(6)) * rpm/rpm_max) ; COF(6) = 0.3 as default.

This gives a realistic impact of the rotor for an average screen, reducing the separation efficiency from 90 to 71 % by increasing the rotor speed by 30 %.This is impact values reported from experiments for a typical screen. With COF(6) = 0.2, the impact will be going from 90 to 79 % separation efficiency, which is a bit more conservative.

Clog_par= short_clog_COF(3)/(short_clog_time+short-clog_COF(3)*
(long_clog_COF(4)/ (long_clog_time+ long_clog_COF(4))

conc_par= 1- (concentration in reject/COF(x))

where COF(3)=0.06 as default. The COF(3) is chosen so, that the maximum value of the reject before clogging the screen is used as COF(3). This may be 0.06 for a typical screen, giving the right effect on e.g. a screen going from 0.5 to 1.5 %,with an increase in separation efficiency from 58 to 81 %. COF(3)= 0.15 is the very maximum value of any screen.

Area_par= active_hole_area/ total_nominal_hole_area

Dh = hole diameter or slot width in mm.

Fibre/ particle: lengths ,diameters and heights are also given in mm; virtual radius is also in mm.

The separation efficiency (SepEff) for each particle size is calculated principally as the ratio between the weighted fibre size and the hole/slot diameter, compensated for the clogging of the pores by multiplying the hole diameter with the Area_par. To avoid division with 0, the hole diameter is added to 1.0, and the COF(7) is used for the tuning to different screen types

SepEff = (fiber_COF(20)*length + 10*fiber_COF(20)*(width + height))
SepEff = (virtual radius/accept flow rate)^COF(21)* SepEff* COF(7) /((1 + Dh* Area_par)

For slots Dh is calculated as Dh= slot width * 7.0. This has been seen as realistic from experimental data.

COF(7) has to be calculated (see below) while COF(20) =0. 0350 and COF(21) = 0.08 as default values. Default value for COF(7) = 3.44 for flow rates c 10-150 l/m2.s.

The mass balance between feed, reject and accept now is calculated as:

Mass_flow_reject =
SepEff * kg/h_each_fiber_fraction_in_feed*(m3/h_reject_flow/m3/h_feed_flow)^COF(22) (kg/h).
Mass_flow_accept= Mass_flow_feed - Mass_flow_reject

The concentration of each fiber fraction in the accept is the

Concentration_accept= Mass_flow_accept/ (m3_accept_flow_per_h).

Hydro cyclones, cleaners

The second most common type of equipment is the Hydrocyclone, or cleaner. This also includes a number of similar equipment like the deculator.

Cleaners are looked upon as a vessel that is either full (= separation working), or as not full (= separation is not working). Calculation of the liquid level in the cyclone or the common vessel for several cyclones is first done, and if positive the separation is calculated according to the following procedure:
Principally the deviation of particles from the stream lines during the rotational flow will be related to the volume of the particle divided by the friction of the particle surface relative the water, the density difference between particles and water, the rotational velocity, the cyclone diameter (giving the rotational velocity-higher for small diameter cyclones) and the viscosity of the water.

First we calculate a **shape factor** = (1+ fibre diameter)/ (1+ fibre length).Both lengths in mm. This compensates for the fact that an elliptical particle moves in a different way than a spherical particle.

We then calculate an **adjusted particle radius** : First we calculate the volume of the particle. For rectangular pieces Volume V = H*W*L (Height*Width*Length). For fibres

the volume is calculated from V= π*R^2 * L. Hereafter the radius R for the sphere with the same volume is calculated as

R = (V*3/ 4π) ^1/3.Length in meter.

The cyclone **volume** is calculated from the geometric inputs, to give the residence time = volume/feed_flow_rate.

An **adjusted cyclone radius** is calculated as the avaerage from swirl zone and bottom of cone.

The basic equation for gravitational and centrifugal separation matches lifting forces with buoyancy forces + forces due to liquid motion: $(4/3)*\pi R^3 *\rho_s*g=(4/3)*\pi*R^3 *\rho_{liq}*g+ 6*\pi*\mu*v(d)*R$ [Bird et al 2002] . Solving for deviation velocity due to gravitational forces becomes $v(g)= (2/9)*R^2 *(\rho_s -\rho_{liq})*g/\mu$

while the corresponding velocity for centrifugal forces becomes principally for a sphere:

$v(c)= (2/9)*R^2 *(\rho_s -\rho_{liq})*v(r)^2/(r*\mu)$.

The deviation v(c) is in m/s from the stream lines in radial direction due to the liquid turning around in the cyclone (radius r) with velocity v(r). It is calculated in our algorithm using an "adjusted radius" to the sphere for other particle shapes like fibres, and with a correction for the larger drag forces due to long, thin fibres compared to spheres, by multiplying with the shape factor:

v(d)= COF(11)* (adjusted particle radius)^2 *Shapefactor* (density difference water-particle)* [v(r)^2/r]*[1/ viscosity]

v(r) = Qin / Area of inlet pipe to cyclone

μ= viscosity, 10^-3 Ns/m2 for water.

Principally we can also include the effect of higher consistency and temperature effects on separation in the viscosity term. This means that μcorr=μ*0.02*(conc/2.0)^-2.2^for zero to 3 % consistency. The temperature effect on the viscosity is directly related by μ T1 = 1.002*10^-3 * (T1/20)^-0.737, where the temperature is given in ^oC, and compared to the viscosity at 20 ^oC, where it is 1.002 * 10^-3 .

The total distance for the particles will be given by multiplying with the residence time in the cyclone, and will depend on the liquid flow as well as the volume of the cyclone.

The shape factor takes into account freeness (surface roughness) as well as the shortest particle diameter. High surface area and long fibres will go preferably in the top or centre, compared to spheres and short fibres, assuming the same density. High density particles will go towards the wall, and downwards.

The absolute separation will also depend on the split rate between flow upwards (Qupaccept,center) resp downwards(Q down,wall,reject).If we assume 50 % volumetric flow in both, the separation will be 50 % of the fibres in each stream, if the density is the same of the fibres as for water, for average sized and shaped particles. To give the mass

separation we calculate the part of the incomming (inject) massflow that goes to the reject(wall) as Minj*Qrej/Qinj.

If the density is higher for particles, the relative distance compared to the distance of the radius of the cyclone, will give the extra separation efficiency of particles of a specific size towards the bottom compared to the top outlet. Where the flow rate is very low, also the gravity is considered, but then in relation to the liquid level in the vessel. This is giving the gravimetric separation efficiency as:

ηg= COF(11)* (adjusted particle radius)^2 *Shapefactor* (density difference water-particle)*g*(1/ viscosity)*(residence time/liquid level)

The separation factor for centrifugal forces will be calculated as :

ηc = v(c)* residence time in cyclone/ cyclone radius (upper part)

where v(c) = radial velocity and the particle residence time in the cyclone is Qin/volume of cyclone.

The massflow (M(I)) in the reject for each particle fraction (I) is calculated by:

Mrej(I) = Minj(I) * ((Qrej/Qinj) +ηc + ηg) for the reject and
Macc(I) = Minj(I) - Mrej(I) for the accept.

The mass separation efficiency of the cyclone then is ηs = Mrej/Minj= (Qrej/Qinj)+ ηc + ηg.

The concentration concerning particles of a certain size/shape going to the wall/bottom will be calculated according to:

Conc rej,bottom(I)= Mrej(I)/ Q bottom

For the top or centre we will get correspondingly:

Conc acc,top(I) = Macc(I)/ Q top

To get the negative effect of fast increases or decrease of the incoming flow Qinj on the separation, this is decreased according to adding a turbulence effect factor

$(\eta c+\eta g)(t) = (\eta c+\eta g)(t) - |\Delta v/v(t)|$

The mass balance is calculated, to give the concentration of fibres for each fraction (I) , in the top respectively bottom.

The pressure flow network makes use of the Bernoullie´s equation, which is principally:

$v_1^2/ 2g + p_1/\rho g + h_1 = v_2^2 / 2g + p_2 / \rho g + h_2$ + friction losses

where v_1 and v_2 is the velocities , p_1 and p_2 the pressures and h_1 and h_2 the liquid heads upstream resp downstream an entrainment like a valve, or in both ends of a pipe etc.

If we just look at a valve with the same liquid head on both sides, we can simplify this to principally

$v = \text{Constant} * SQRT((p_1-p_2)/\rho)$

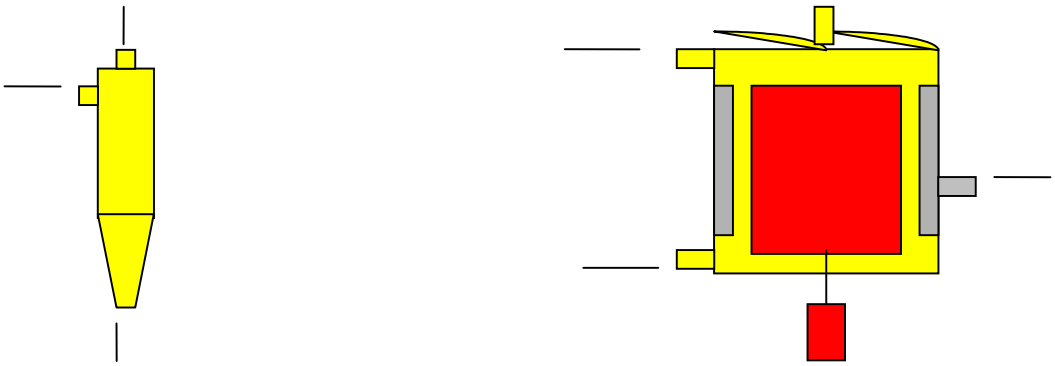
for a fully open valve, where the constant relates to the friction losses caused by different geometries. The velocity v multiplied with the open area of the valve, will give the flow through the valve, for a given pressure difference.

By using this technique and making a number of equations, one for each node, and then solve this set of equation simultaneously, we will determine the pressure and flow in the whole network of pipes and process equipments. For the different process equipments, pressure losses are determined due to the operating conditions, and thus are included in the calculation, as well.

Separation of fibres and other particles for the different process equipments is also calculated in each equipment algorithm. This gives the material balance over the equipments and for the whole network, for each time step, considering also dynamics. This is useful, when you want to test new advanced control algorithms , where you don´t have the DCS code for them, but can write them in Fortran, C++ or make use of Matlab instead. With a good process model, it is possible to test the control strategy before implementing it on the real process.

Other examples
There are many different models developed for both paper and pulp mills. For pulp mills the major focus has been on the digester, and several examples exist on good models for a number of different applications here, like [Bhartiya et al 2001],[Wisnewski et al 1997] and [Jansson et al 2004].

Model validation and tuning with process data
Examples of process equipments mentioned earlier can be the screen (to the right) and the hydro cyclone (= cleaner, to the left)



For the screen algorithm we assume a plug flow from the top and down wards, but with total mixing in radial direction. The rotor rotate with a relatively high velocity normally, giving shear forces at the screen surface. Fibres are mechanically separated at the screen if they are larger than a certain size in relation to the hole area or slot size, but also

depending on the concentration, shape, flow rate through the holes or slot, the rotor speed, temperature and volume reduction factor (reject rate).The model gives the pressure drop over the screen, as well as the mass and energy balance. At low reject flow, the concentration in the reject goes up, and if it becomes to high, the motor stops, the fibres accumulate and eventually the whole screen plugs up. The pressures around the screens are calculated, but also the fibre size distribution of the fibres in the different streams, as well as the amounts and concentrations.

In the table below results for the model algorithm calculations are compared to data from experiments done by technical institutes in Canada(Paprican) [Gooding et al 1992]and Sweden [STFI 1999]. The model used is the one described earlier.:

For a 1.4 mm hole screen, the following results can be seen:

Qrej/ Qfeed	**accept l/m2.s**	**Separation efficiency 0-0.5**		**0.5-1**		**2-5 mm fibers**	
		Exp	**Calc**	**Exp**	**Calc**	**Exp**	**Calc**
0.4	**99**	**74.8**	**77.1**	**80.5**	**80.6**	**92.0**	**91.8**
0.7	**16.5**	**89.3**	**89.0**	**92.3**	**93.0**	**97.6**	**100**

For a 0.4 mm slot screen, assuming 0.7 mm long , 0.025 mm wide fibers:

	Experimental			**Calculated Separation efficiency**		
Accept l/m2.s	**50**	**100**	**200**	**50**	**100**	**200**
Q_rej/Q_feed						
0.4	**0.61**	**0.52**		**0.63**	**0.52**	
0.29			**0.33**			**0.35**
0.09	**0.27**	**0.18**	**0.12**	**0.25**	**0.18**	**0.15**

For a cleaner(=hydrocyclone), the corresponding figures are shown below:

FlowQrej/Qinj		**0.25 mm**		**0.75 mm**		**1.75 mm**		**3.5 mm fibers**	
l/min		**Exp**	**Calc**	**Exp**	**Calc**	**Exp**	**Calc**	**Exp**	**Calc**
270	**0.10**	**0.24**	**0.25**	**0.29**	**0.31**	**0.32**	**0.32**	**0.32**	**0.33**
500	**0.26**	**0.56**	**0.54**	**0.67**	**0.66**	**0.74**	**0.70**	**0.77**	**0.71**

As can be seen, the prediction of the separation efficiency can be quite good, although not absolute. The reason is both the difficulty to make the experiments totally controlled (see e.g. the first row of the cleaner experiment, where the separation efficiency is not correlated to fibre size for high reject , low flow), and to catch all possible effects in one single model, based on first order principles. The tuning / configuration of the models are done to fit the actual equipment and the normal operating range, but will give reasonably good result also outside this area.

In reality data from many more papers were used to build the models, where different variable were varied, and different types of equipments tested. Unfortunately few of these

include both variation of flow rates, concentrations, reject/accept ratio, different fibre size distributions etc, but mostly only one of these, and normally not including looking at the effect on different fibre sizes. Still, together all these data bring new elements into the puzzle to build a good physical model. An example of starting with a physical model and tune it with statistical plant data was shown by [Pettersson 1998]

When we look at a new type of screen we use normally simple mass balance data to tune at some few operational conditions. These data will be used to tune some of the coefficients mentioned earlier, while the rest are used as default values until new data comes up that may be used to proceed with more detailed tuning for the specific equipment. Some of this has been collected from [Nilsson 1995] and [Jegeback 1990, 1993].

Simulators for other applications

After the start up, or for an existing mill directly, the dynamic simulator can be used for optimizing the process. This is done so, that different ways of running the process can be tested by the process engineers on the simulator, to see how e.g. water flows or dry solids contents etc are influenced, e.g. during a grade change.

By collecting data through the information management system, and sending them to the simulator, it may be possible to use the data for more advanced diagnostics [Karlsson et al 2003]. This will be for both the process and the sensors, by making use of the expert system functions residing in the simulator. With a model, that shows how different sensors and process parts are correlated to each other, predictions of performance can be made as well. By comparing to the real process signals, deviations and drifts can be diagnosed, to alarm the operator before it is possible to see the faults “manually”.

For the process optimization, higher fidelity models (the fidelity of the models can be selected for the most important equipments) may be needed[Bell et al 2004],[Dhak et al 2004][Hess 2000] and [Morari et al 1980]. The communication speed with the DCS systems does normally not need to be considered, as the process engineer can work on the simulator without the real time DCS system connected. With interaction between a simulator and an optimization algorithm the communication still may be the limiting parameter.

Conclusions

A model has been developed for major pulp and paper equipments. It has been shown by comparing experimental results to model predictions that a reasonably good prediction for separation efficiency for many fractions can be made for separation equipments operating under a wide operational range. The same model can be used for many different types of process equipments and models, where only a number of parameters (constants) have to be configured with data from relatively few experiments, or principally normal mass balance data that can be achieved from vendors or mills.

References:

NOPS Paper Machine Operator Training System, ABB Process Automation, 1990
Bird R.B. ,W.E.Stewart and E.N.Lightfoot: Transport Phenomena , by,John Wiley & sons, 2nd edition 2002.
Gooding Robert W. and Richard J.Kerekes: Consistency changes caused by pulp screening , Tappi Journal, Nov 1992, p 109-118.
Nilsson Anders: The simulation resource Extend , Licentiate Thesis Pulp&Paper Tech Dept ,Royal Inst of Technology in Stockholm, TRITA-PMT Report 1995:14.
Jegeback M. and B.Norman: Fextend- computer simulation of paper machine back water systems for process engineers, STFI report A 987, 1990.
Jegeback M.: Dynamic simulation of FEX back water system by Fextend, ,STFI report A 996, 1993
Data from experimental reports from STFI on screening and cleaning 1999.
Bell J., Dahlquist E.,Holmstrom K.,Ihalainen H.,Ritala R.,Ruis J.,Sujärvi M.,Tienari M: Operations decision support based on dynamic simulation and optimization. PulPaper2004 conference in Helsinki, 1-3 June 2004.Proceedings.
Dhak J.,Dahlquist E.,Holmstrom K.,Ruiz J.,Bell J.,Goedsch F: Developing a generic method for paper mill optimization.Control Systems 2004 in Quebec City, 14-17 June 2004. Proceedings.
Ulf Persson, Lars Ledung, Tomas Lindberg, Jens Pettersson, Per-Olof Sahlin and Åke Lindberg: "On-line Optimization of Pulp & Paper Production", in proceedings from TAPPI conference in Atlanta, 2003.
Wisnewski P.A, Doyle F.J and Kayihan F.: Fundamental continuous pulp digester model for simulation and control. AIChE Journal Vol 43, no 12, dec 1997, pp 3175-3192.
Pettersson J. (1998): On Model Based Estimation of Quality Variables for Paper Manufacturing. Tech LicThesis , KTH
Bhartiya , Dufour and Doyle (2001) : Thermal- Hydraulic modelling of a continuous pulp digester, in proceedings from Conference on Digester modelling in Annapolis, June, 2001.
Hess T. (2000): Process optimization with dynamic modeling offers big benefits, I&CS,August, p 43-48.
Karlsson C., Dahlquist E., "Process and sensor diagnostics - Data reconciliation for a flue gas channel", Värmeforsk Service AB, 2003, (in Swedish).
Morari M, Stephanopolous G and Arkun Y: Studies in the synthesis of control structures for chemical processes. Part 1: formulation of the problem.Process decomposition and the classification of the control task. Analysis of the optimizing control structures. American institute of Chemical Engineering Journal , 26(2), 220-232.1980.
Ryan K. and Dahlquist E.: MNI Experiences with process simulation. Proceedings Asia Paper, Singapore, 2002.
Jansson and Erik Dahlquist : Model based control and optimization in pulp industry, SIMS2004,Copenhagen Sept 23-24,2004.

Model Uncertainty and Prediction Capabilities

Bernt Lie
Telemark University College, P.O. Box 203, N-3901 Porsgrunn, Norway

September 19, 2005

Abstract

Deterministic and statistical descriptions of parametric model uncertainties are discussed, and illustrated with a case study from the paper industry. Prediction uncertainties under open loop and closed loop operation are then studied. The results illustrate the importance of a realistic description of parametric uncertainties, and also how closed loop operation can reduce the prediction sensitivity to parameter uncertainties.

1 Introduction

It is of interest to study how model uncertainties influence the prediction capabilities of models. Model uncertainties can be described in many ways. The study can be restricted to include parameter uncertainties under the assumption of structurally perfect models, or can be more realistic to include uncertainties in model structures. In general, it is difficult to describe uncertainties in the model structure.

Parametric uncertainties can either be deterministic in that they are based on a physical understanding of the system under study, or they can be based on statistics, e.g. from model fitting.

Models are used extensively in designing control solutions. One possibility is to design a control input entirely based on the model, and then inject the computed control input into the system. This is often denoted *open loop* operation. Alternatively, a feedback (*closed loop*) solution can be developed, where for each time step, a control mechanism checks whether the real system operates as expected from the model. If the real systems drifts away from the response predicted by the model, correction is introduced.

The paper is organized as follows: In Section 2, a case study model taken from paper machine 6 (PM6) at Norske Skog Saugbrugs, Norway is described. In Section 3, parameter estimation is described for the PM6 model, with resulting descriptions of uncertainties. In Section 4, prediction uncertainties are studied for the PM6 model, both in open loop operation and in closed loop control. In Section 5, some conclusions are drawn.

2 Case study: Paper machine model

A sketch of Paper Machine 6 (PM6) at Norske Skog Saugbrugs, Norway, is given in fig. 1. Hauge (2003) proposed the following linearized dynamic model of PM6:

$$x_{t+1} = Ax_t + Bu_t + Ed_t + Gw_t \tag{1}$$

$$y_t = Cx_t + Du_t + Fd_t + v_t \tag{2}$$

where[1]

$$A = \begin{pmatrix} 0.9702 & 0.3283 & 0 \\ 0.0018 & 0.9596 & -0.0197 \\ 0 & 0 & 0.8661 \end{pmatrix}, \quad C = \begin{pmatrix} 61 & 727 & 13109 \\ 83 & 986 & -1692 \\ 3 & 34 & -32 \end{pmatrix} \tag{3}$$

$$B = \begin{pmatrix} 1.3 & 160.1 & 0.2 \\ 0.1 & 10.1 & -33.4 \\ 1.3 & 0 & -0.7 \end{pmatrix} \times 10^{-6}, \quad D = \begin{pmatrix} 0.0029 & 0.3544 & 5.3831 \\ 0.0040 & 0.4815 & 7.1769 \\ 0.0001 & 0.0166 & -0.0554 \end{pmatrix} \tag{4}$$

$$E = \begin{pmatrix} 0.0247 & 0.0023 & 0 & 0 \\ 0.0016 & 0.0001 & 0 & 0 \\ 0.0134 & -0.0007 & 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 54.5613 & 5.1415 & -1.9777 & 51.0179 \\ 74.1090 & 6.9836 & 0 & -30.6923 \\ 2.5519 & 0.2405 & 0 & 0 \end{pmatrix}. \tag{5}$$

[1] There is a misprint in Hauge (2003): his B matrix should be multiplied by 10^{-6}.

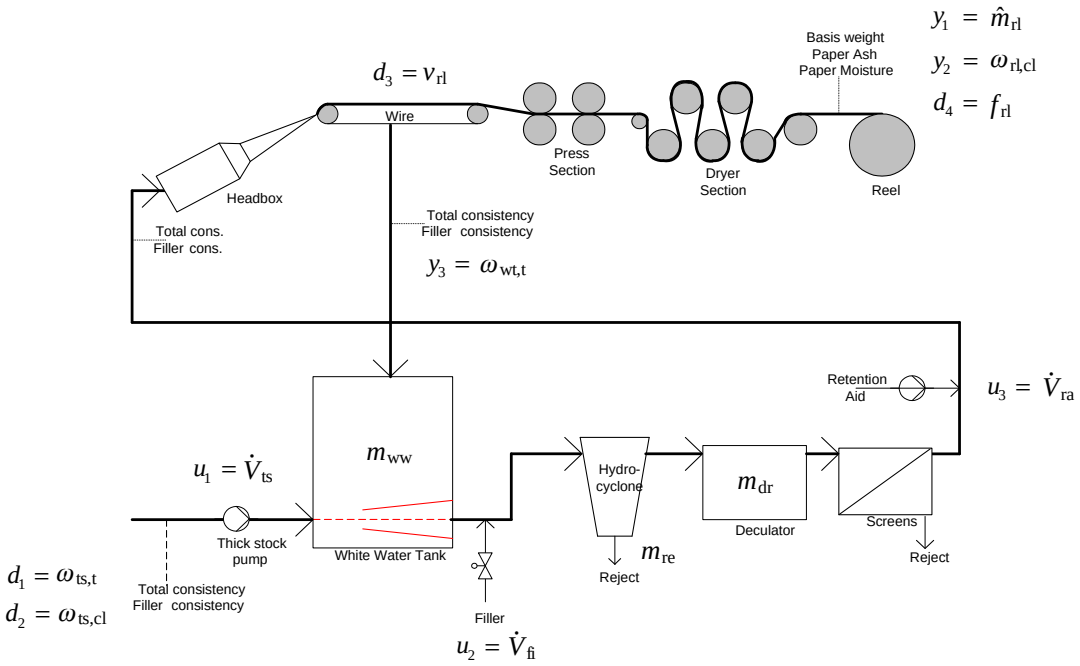


Figure 1: Functional sketch of PM6, with manipulated inputs u, disturbances d, and controlled outputs y. After (Hauge 2003)

Here t is a discrete time index with discretization time 0.5 min, manipulatable input u_t is the deviation from the operating values of $\left(\dot{V}_{ts}, \dot{V}_{fi}, \dot{V}_{ra}\right)$, measured disturbance d_t is the deviation from operating values of $(\omega_{ts,t}, \omega_{ts,cl}, f_{rl}, v_{rl})$, x_t is the state, controlled output y_t is the deviation from operating values of $(\hat{m}_{rl}, \omega_{rl,cl}, \omega_{wt,t})$, unmeasured disturbance w_t, and unmeasured output noise v_t.

In the model, $\dot{V}_{ts}$ is the volumetric feed flow of Thick Stock, $\dot{V}_{fi}$ is the volumetric feed flow of Filler (clay), $\dot{V}_{ra}$ is the volumetric feed flow of Retention Aid, $\omega_{ts,t}$ is the Thick Stock "Total" mass fraction (mass fraction of fibers + clay, also denoted total consistency), $\omega_{ts,cl}$ the Thick Stock mass fraction of Clay, f_{rl} is the moisture at the Reel (finished paper), v_{rl} is the linear velocity of paper at the Reel, $\hat{m}_{rl}$ is the mass per area of paper at the Reel, also denoted basis weight, $\omega_{rl,cl}$ is the mass fraction of clay at the reel, also denoted paper ash, while $\omega_{wt,t}$ is the Wire Tray "Total" mass fraction (fibers + clay). The state x_t has a physical interpretation related to accumulated masses m_j (see fig. 1), but is not of particular interest in this study.

3 Model uncertainty

3.1 True system

We will consider systems which are characterized by system order, system mapping, system states, and possibly by stochastic inputs. None of these are known in practice: we can only have models of them. To illustrate concepts, we will consider the model in eqs. 1 and 2 as the true system, and the matrices in eqs. 3–5 as the true model parameters. Furthermore, we will consider $x_1 = 0$ to be the true initial state, and the stochastic signals to be $w_t \equiv 0$ and $v_t \sim \mathcal{N}(0, S_v)$ where

$$S_v = \text{diag}\left(0.1^2, 0.1^2, 0.001^2\right)$$

is the constant covariance matrix of v_t.

3.2 Experimental data

Let $u_t \in \mathbb{R}^{n_u \times 1}$, $t \in \{1, \ldots, T\}$ denote the sequence of input data and $y_t \in \mathbb{R}^{n_y \times 1}$, $t \in \{1, \ldots, T\}$ the sequence of output data from the true system. A possible set of input sequences to the true system are

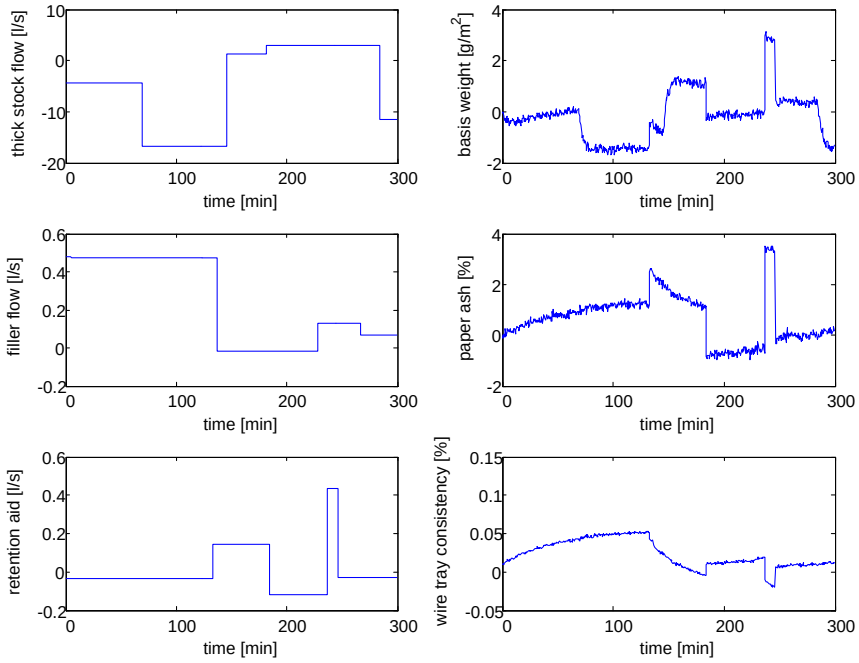


Figure 2: Experimental input and output data.

displayed in the left column of fig. 2, and the resulting output sequences from the true system are shown in the right column. Here, $T = 600$.

3.3 Initial information

Apart from the experimental data, the true system is unknown. Assume that based on prior knowledge of the system, we have developed a model with structure as in eqs. 1–2.[2] Furthermore, assume that we have a prior idea about the values of the matrices in the model: we happen to know B, C, and D perfectly[3], and for matrix A we have partial knowledge:

$$A = \begin{pmatrix} a_{11} & 0.3283 & 0 \\ a_{21} & 0.9596 & -0.0197 \\ 0 & 0 & 0.8661 \end{pmatrix};$$

A is known perfectly apart from parameters a_{11} and a_{21}, which are unknown. Finally, assume that we know that $x_1 = 0$, that $w_t \equiv 0$, but that we only know that v_t is *independently and identically distributed* (i.i.d.).

Based on this prior knowledge, we want to find estimates of parameters a_{11} and a_{21}; to simplify the notation, we introduce the parameter vector $\theta = (a_{11}, a_{21}) \in \mathbb{R}^{n_\theta \times 1}$. Before trying to find the estimates, assume that we also have a vague idea about the parameters θ:

- We are relatively sure that $\theta \in [0.5, 1] \times [0, 0.5]$, where $\times$ denotes Cartesian product.
- A first guess of θ is $\theta^0 = (0.75, 0.25)$.

In practice, there will be missing experimental data and outliers in the experimental data. We will not consider this added problem here.

3.4 Parameter estimation

Our model output will be denoted y_t^m, and y_t^m is a function of θ and $U_t = \{u_1, \ldots, u_t\}$, as well as the initial state x_1; $y_t^m(\theta; U_t, x_1)$, usually simplified to $y_t^m(\theta)$. Intuitively, we want to find θ such that the

[2] In reality, we will never be able to formulate a perfect model.
[3] Model parameters are never known perfectly in practice.

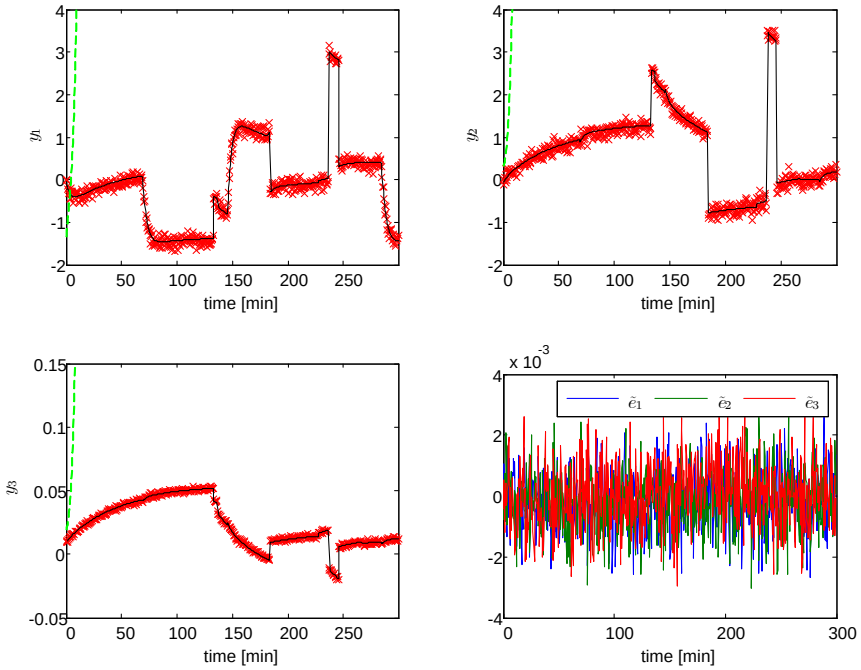


Figure 3: Comparison of real outputs y_t ($\times$, red color), model outputs $y_t^m\left(\hat{\theta}\right)$ (solid, black), initial model outputs $y_t^m\left(\theta^0\right)$ (dashed, green), as well as scaled model errors $\tilde{e}_t$ (errors for outputs $1, 2$ have been divided by 100).

model output y_t is as close as possible to the true output y_t. A common strategy for achieving this, is to introduce the model error $e_t \triangleq y_t - y_t^m(\theta)$ and thus choose our best estimate $\hat{\theta}$ such that

$$\hat{\theta} = \arg\min_{\theta} \sum_{t=1}^{T} e_t^T W_t e_t \tag{6}$$

is minimized; this estimate is the weighted least squares estimate $\hat{\theta}_{\text{WLS}}$, and W_t is the weight. If e_t is white noise with a normal distribution with zero mean and known covariance matrix S_e, and if we choose $W_t = S_e^{-1}$, then the weighted least squares estimate $\Delta\hat{\theta}_{\text{WLS}}$ is identical to the Maximum Likelihood estimate $\hat{\theta}_{\text{ML}}$.[4]

By using the initial information discussed in Section 3.3 in combination with the least squares criterion in eq. 6, and with $W_t = \text{diag}\left(\frac{1}{100^2}, \frac{1}{100^2}, 1\right)$,[5] the use of the `lsqnonlin` algorithm in Matlab leads to the least squares estimate

$$\hat{\theta} = (0.54141, 0.040948)\,.$$

Figure 3 shows outputs from the true system y_t ($\times$, red) and initial model outputs $y_t^m\left(\theta^0\right)$ (dashed, green), as well as model outputs $y_t^m\left(\hat{\theta}\right)$ (solid, black). As the figure indicates, the system is in fact unstable with parameter θ^0, so this initial guess doesn't make much sense: it is advisable to choose an initial guess θ^0 which gives gives a stable model.

Figure 3 also indicates the scaled output errors $\tilde{e}_t$, and the plot gives the impression that these errors are i.i.d. signals.

3.5 Parameter statistics

We have assumed that the output noise v_t observed in our experiments is a realization of a stochastic variable, and that the resulting model error sequence $e_t = y_t - y_t^m\left(\hat{\theta}\right)$ is i.i.d. Thus, if we had used other experimental data to find the parameter estimate $\hat{\theta}$, we would most likely have found another value $\hat{\theta}$

[4] In reality, S_e is almost always unknown.

[5] W_t is chosen so as to give each element of y_t the same weight in the criterion function J.

from this other sequence of modeling errors e_t. This indicates that the parameter estimate should be considered a realization of a stochastic variable Θ with a probability distribution $F_\Theta(\theta)$. If we had known this probability distribution, we could have computed e.g. the confidence interval for our estimate $\hat{\theta}$.

Unfortunately, the distribution $F_\Theta(\theta)$ is not known. A common strategy in parameter estimation is to assume that the scaled model error $\tilde{e}_t = e_t\sqrt{W_t}$ is both i.i.d., and that it is normally distributed, $\tilde{e}_t \sim \mathcal{N}(0, S_{\tilde{e}})$. If $S_{\tilde{e}} = \sigma^2 I$, σ^2 can be estimated from

$$\hat{\sigma}^2 = \frac{1}{T \cdot n_y - n_\theta} \sum_{t=1}^{T} e_t^T W_t e_t.$$

Next, if the output y_t^m is nonlinear in the parameter, then Θ is *approximately* normally distributed $\Theta \sim \mathcal{N}\left(\hat{\theta}, S_\Theta\right)$.[6] An estimate of S_Θ is

$$\hat{S}_\Theta = 2\hat{\sigma}^2 H^{-1}$$

where H is the Hessian of the criterion $\sum_{t=1}^{T} e_t^T W_t e_t$. Approximately[7], $H = 2\sum_{t=1}^{T} X_t^T X_t$ where X_t is the Jacobian of y_t^m wrt. θ.

For the experimental data used to find $\hat{\theta}$ in the previous section, we find that

$$\begin{aligned} \hat{\sigma}^2 &= 1.0126 \times 10^{-6}, \\ \hat{S}_\Theta &= \begin{pmatrix} 0.26594 & -0.024276 \\ -0.024276 & 0.0022161 \end{pmatrix}. \end{aligned}$$

With known distribution $F_\Theta(\theta)$ and suppose $n_\theta = 1$, we then seek the $\alpha \cdot 100\%$ *confidence region* defined by $[\theta_\ell, \theta_u]$ for Θ:

$$\Pr(\theta_\ell \leq \Theta \leq \theta_u) = \alpha.$$

The confidence region should be interpreted as follows: If we carry out infinitely many experiments on our system (e.g. PM6) and compute $\hat{\theta}$ and the $\alpha \cdot 100\%$ confidence region for each experiment, in $\alpha 100\%$ of the cases, the true parameter θ lies in the region. It is important to realize that this means that θ actually *may* lie outside of the computed confidence region.

Clearly, $\Pr(\theta_\ell \leq \Theta \leq \theta_u) = F_\Theta(\theta_u) - F_\Theta(\theta_\ell) = \alpha$ does not have a unique answer (θ_ℓ, θ_u). It is thus common to rephrase the formulation into

$$F_\Theta(\theta_\ell) = \Pr(\theta_\ell \leq \Theta) = \frac{1-\alpha}{2} \tag{7}$$

$$F_\Theta(\theta_u) = \Pr(\theta_u \leq \Theta) = 1 - \frac{1-\alpha}{2}, \tag{8}$$

where $\Theta \in [\theta_\ell, \theta_u]$.

For $n_\theta > 1$, it is common to introduce the new stochastic variable $\Delta = g\left(\Theta, \hat{\theta}\right) \in \mathbb{R}^1$ and consider $\Pr(\Delta \leq \delta) = \alpha$, or $F_\Delta(\delta) = \alpha$. Here we have assumed that we know the distribution $F_\Delta(\delta)$. By computing δ, we then find the confidence region of Θ from the inequality

$$g\left(\Theta, \hat{\theta}\right) \leq \delta.$$

Consider the stochastic variable Δ

$$\Delta = \frac{1}{2}\left(\Theta - \hat{\theta}\right)^T \hat{S}_\Theta^{-1}\left(\Theta - \hat{\theta}\right),$$

which can be shown to be $\mathcal{F}$-distributed with degrees of freedom n_θ and $n_y T - n_\theta$, $\mathcal{F}(\Delta; n_\theta, n_y T)$. Then δ is computed as, (Rawlings & Ekerdt 2002)

$$\delta = n_\theta \cdot \mathcal{F}^{-1}(\alpha; n_\theta, n_y T - n_\theta)$$

to give $\alpha \cdot 100\%$ confidence.

[6] This *may* be a poor approximation.

[7] Known as the Gauss-Newton approximation.

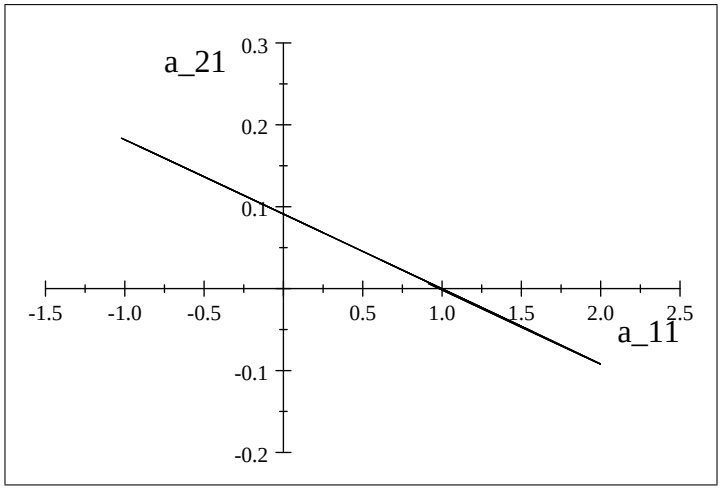


Figure 4: Confidence region based on normal distribution assumption, with 80% certainty that the true parameter is within the (degenerated) ellipsoid.

With the estimates for PM6 and desiring 80% confidence[8], we get[9]

$$\delta = 2 \cdot \mathrm{FInv}\left(0.80, 2, 3 \cdot 600 - 2\right) = 3.\,221\,8.$$

This gives

$$3.\,221\,8 = \frac{1}{2}\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} - \begin{pmatrix} 0.54141 \\ 0.040948 \end{pmatrix}\right)^T \begin{pmatrix} 0.26594 & -0.024276 \\ -0.024276 & 0.0022161 \end{pmatrix}^{-1} \left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} - \begin{pmatrix} 0.54141 \\ 0.040948 \end{pmatrix}\right),$$

which describes the ellipsoid in fig. 4. The figure illustrates that the confidence ellipsoid has virtually degenerated into a line, which is supported by the singular values of $\hat{S}_\Theta$ ($\sigma_1 = 0.27, \sigma_2 = 9.\,5 \times 10^{-8}$). The interpretation of this is that parameters a_{11} and a_{21} are linearly dependent, and that only $c_1 a_{11} + c_2 a_{21}$ can be identified. We will, however, not pursue this problem further, here.

The naïve interpretation of the confidence region is that with 80% confidence

$$a_{11} \in [-1, 2] \tag{9}$$
$$a_{21} \in [-0.1, 0.2]\,. \tag{10}$$

3.6 Bootstrap distribution

The confidence region found in the previous section is based on the assumption that the model deviation e_t is distributed normally, and that the output is a linear function of the parameters $\theta = (a_{11}, a_{21})$. If the distribution for e_t, $F_E(e)$, were known, we would not have to make this assumption. We could then compute $F_\Theta(\theta)$ by Monte Carlo simulation as follows:

1. First we would draw an outcome $\{e'_1, \ldots, e'_T\}$ from the random number generator for distribution $F_E(e)$.
2. Next, we would compute the alternative model response $y'_t = y_t^m\left(\hat{\theta}\right) + e'_t$.
3. Next, we could find the parameter $\hat{\theta}'$ from the data $(u_1, \ldots, u_T)$ and $(y'_1, \ldots, y'_T)$.
4. This procedure could then be repeated as many times as deemed necessary, yielding a distribution in $\hat{\theta}$.

The problem is, of course, that we do not have a random number generator for $F_E(e)$. In the Bootstrap method, (Efron & Tibshirani 1993), (Davison & Hinkley 1997), (Hastie, Tibshirani & Friedman

[8] It is more common to desire 90% confidence, 95% confidence, etc.

[9] $\mathrm{FInv}(\alpha, \nu_1, \nu_2)$ is a built in function in the word processor Scientific WorkPlace for computing $\mathcal{F}^{-1}(\alpha; \nu_1, \nu_2)$.

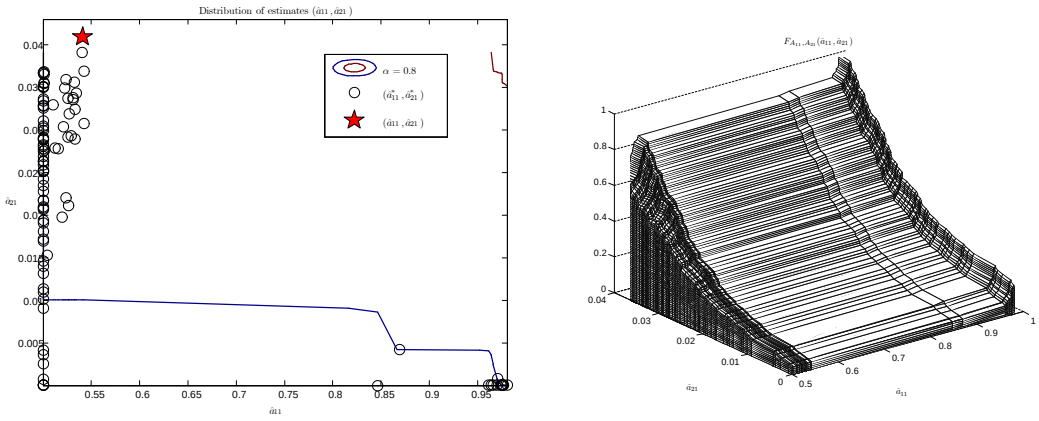


Table 1: Results from Bootstrapping with 100 resamples. The figures show the parameter distribution and 80% confidence region (left) and the cumulative distribution $\hat{F}_\Theta(\theta)$ (right),

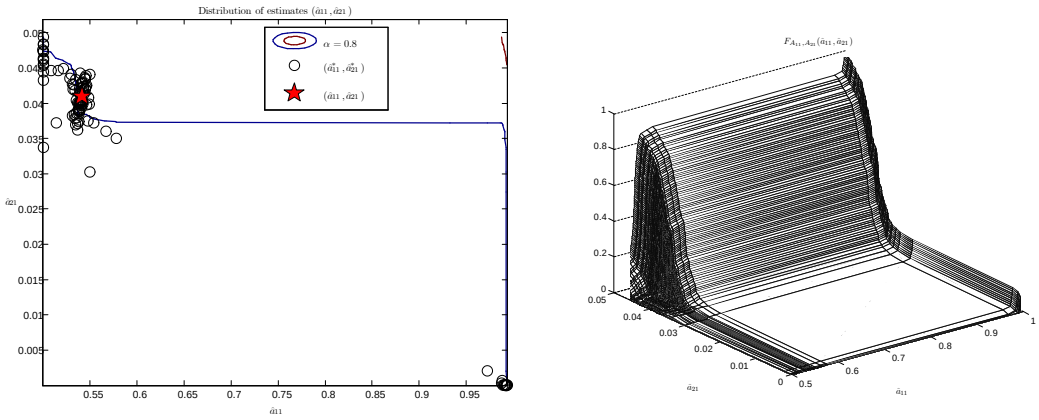


Table 2: Results from Bootstrapping with 100 resamples generated from $y_t^* = y_t^m\left(\hat{\theta}\right) + e_t^* n_t$. The figures show the parameter distribution and 80% confidence region (left) and the cumulative distribution $\hat{F}_\Theta(\theta)$ (right),

2001), we postulate that with $e_t = y_t - y_t^m\left(\hat{\theta}\right)$ i.i.d., the set $\{e_1, \ldots, e_T\}$ constitutes the entire population of possible outcomes from the random generator for $F_E(e)$. Naturally, this is not correct, but experience shows that it is a useful approximation. Since we thus have the entire population of outcomes, we can simply draw a set of outcomes $\{e_1^*, \ldots, e_T^*\}$ from $\{e_1, \ldots, e_T\}$ using a uniform random number generator. Then we find the bootstrap estimate $\hat{\theta}^*$ from the data $(u_1, \ldots, u_T)$ and $(y_1^*, \ldots, y_T^*)$ where $y_t^* = y_t^m\left(\hat{\theta}\right) + e_t^*$. In practice, this resampling of $\{e_1, \ldots, e_T\}$ is repeated a number B of times, leading to a distribution $\left\{\hat{\theta}_1^*, \ldots, \hat{\theta}_B^*\right\}$ which represents $F_\Theta(\theta)$.

Table 1 shows the resulting distribution in parameter estimates from 100 resamplings using the Bootstrap method. Since and estimate $\hat{F}_\Theta(\theta)$ of $F_\Theta(\theta)$ is found, we can thus compute an 80% confidence region for θ from eqs. 7–8. The result is indicated in the left figure of Table 1.

In practice, it is recommended to compute y_t^* as $y_t^* = y_t^m\left(\hat{\theta}\right) + e_t^* n_t$, where n_t is drawn from a random number generator for $\mathcal{N}(0,1)$. See e.g. (Ruiz, Casellano, González, Roca & Lema 2005). The results similar to those of Table 1, are shown in Table 2.

3.7 Uncertainty descriptions

We now consider a general parametric model with output $y_t^m(\theta; U_t, x_1)$ for describing the true output y_t of the real system. Uncertainty in the model can be described in various ways, e.g. by a deterministic description, or by a statistical description.

In a deterministic description, we may be able to say that based on physical understanding of the

system, the model parameters θ belong to a region $\theta \in \Omega$, typically described as a polyhedron (region bounded by plane faces). In the simplest case, the poyhedron is given by

$$\theta_j \in \left[\omega_j^\ell, \omega_j^u\right]. \tag{11}$$

If we have no preferences as to which value in Ω is the true parameter, we can also consider this as a statistical description with a uniform probability distribution within the polyhedron Ω.

If parameters are given with confidence intervals similar to in eq. 11, and with no information about the underlying distribution $F_\Theta(\theta)$, it is relatively common to assume that the true parameter is uniformly distributed within the interval. As described in Section 3.5, this is often a relatively unrealistic assumption.

In a statistical description, we have an estimate $\hat{F}_\Theta(\theta)$. This is e.g. the case when we find $\hat{F}_\Theta(\theta)$ from Bootstrapping. In that case, we in fact have available the entire population of Θ. We can thus easily construct a random number generator for Θ by drawing from the population using a uniform random number generator.

4 Prediction uncertainty

4.1 Optimal control problem

In the sequel, we will study the problem of Optimal Control under uncertainty. The following criterion is considered:

$$\min_u \frac{1}{2} \sum_{i=0}^{N-1} \left(e_{t+i}^T P e_{t+i} + u_{t+i}^T Q u_{t+i} + \Delta u_{t+i}^T R \Delta u_{t+i}\right)$$

where

$$\begin{aligned} e_t &\triangleq r_t - y_t^m \\ \Delta u_t &\triangleq u_t - u_{t-1}; \end{aligned}$$

r_t is the reference value for y_t, and we impose the constraints given by the model, i.e. eqs. 1–2. The following choices are made wrt. the criterion function:

$$\begin{aligned} P &= \operatorname{diag}\left(\frac{1}{4^2}, \frac{1}{2^2}, \frac{1}{0.1^2}\right), \\ Q &= 0, \\ R &= \operatorname{diag}\left(\frac{1}{20^2}, \frac{1}{2^2}, \frac{1}{0.4^2}\right). \end{aligned}$$

Furthermore, we choose the length $N = 60$ on the criterion horizon, which is equivalent to 30 min.

The following study is based on Monte Carlo simulations, using 10 drawings of the parameters.

4.2 Open loop

First we consider the uncertainty description given by eqs. 9–10, assuming a uniform distribution within this polyhedron: this is a description that would come out of a standard confidence interval analysis assuming normal distribution in e_t and a parameter linear model, if we in addition neglect the relationship given by fig. 4. The result of open loop operation, i.e. using the optimal control input computed from the nominal model without feedback, is displayed in Table 3.

Table 3 illustrates the problem of open loop operation: model uncertainty may lead to very uncertain predictions.

Next, we consider the uncertainty description by postulating that the bootstrap outcomes $\hat{\theta}^{*,j}$ constitute the population of possible parameters. We thus draw the parameters from this set in the Monte Carlos simulations. The result of open loop operation is displayed in Table 4.

Table 4 illustrates once again that open loop operation leads to uncertainty in the operation. However, comparison of the results in Tables 3 and 4 illustrates the significance of using a realistic model uncertainty description.

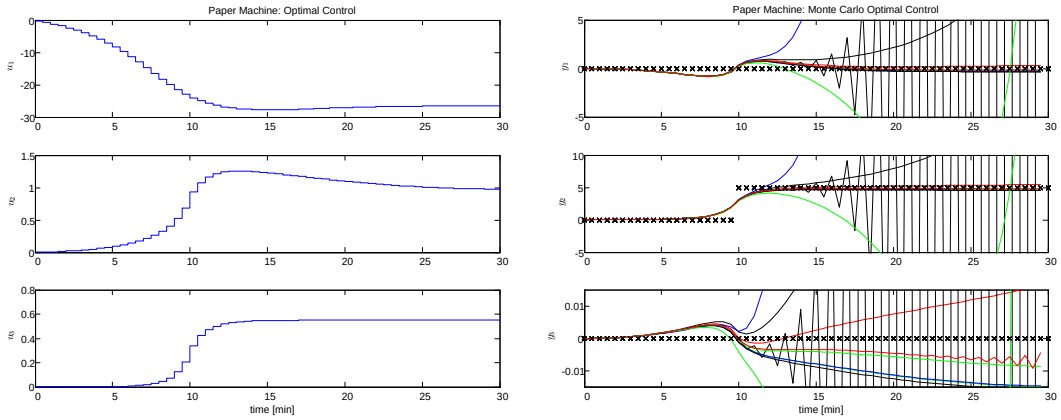


Table 3: Open loop operation with traditional uncertainty description. Control inputs u_t in left figure, and model outputs y_t (solid) and references r_t ($\times$) in right figure.

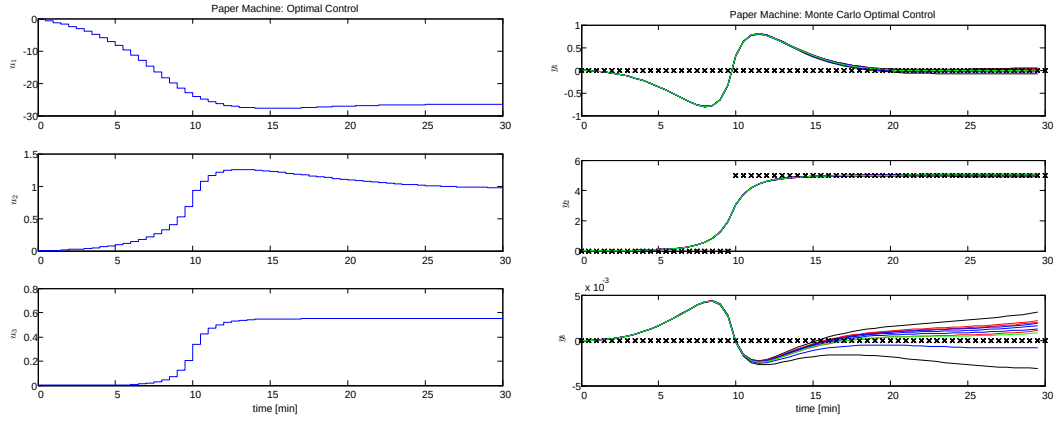


Table 4: Open loop operation with bootstrap based uncertainty description. Control inputs u_t in left figure, and model outputs y_t (solid) and references r_t ($\times$) in right figure.

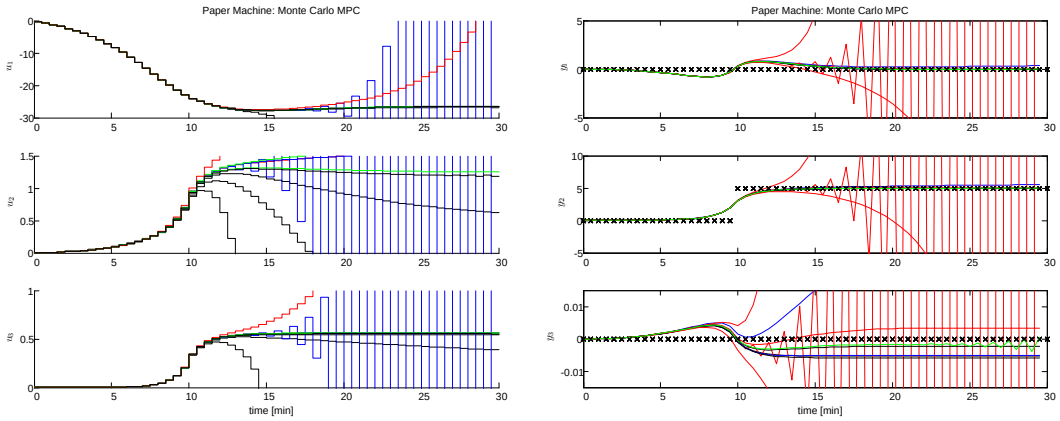


Table 5: Closed loop operation with traditional uncertainty description. Control inputs u_t in left figure, and model outputs y_t (solid) and references r_t ($\times$) in right figure.

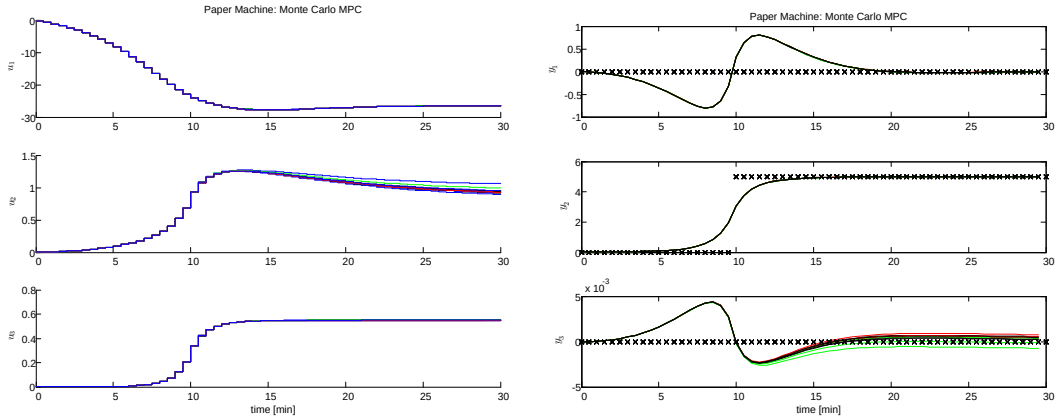


Table 6: Closed loop operation with bootstrap based uncertainty description. Control inputs u_t in left figure, and model outputs y_t (solid) and references r_t ($\times$) in right figure.

4.3 Closed loop

Next, we consider model predictive control MPC (sliding horizon optimal control, or receding horizon optimal control): here, the first control input in the optimal sequence is used. Then the response is measured, and used in a new computation of the optimal control. MPC leads to feedback.

The result of using feedback (MPC) and the "traditional" uncertainty description by drawing randomly from the polyhedron defined by eqs. 9–10, is displayed in Table 5. Since MPC is a feedback solution, the uncertainty is reflected in the control inputs. Even with feedback, the closed loop predictions are very uncertain with the somewhat naïve uncertainty description we have used.

When using the bootstrap uncertainty description, the closed loop uncertainty is illustrated by the figures in Table 6.

As these figures illustrate, the combination of a realistic uncertainty description with feedback drastically reduces the uncertainty in the system.

5 Conclusions

The main emphasis of this paper has been various methods for describing parametric uncertainty in structurally correct models. The methods have been illustrated by using a linear model of PM6 at Norske Skog Saugbrugs as a case study.

The importance of taking advantage of the full model uncertainty description is stressed. With a naïve use of confidence regions, underlying correlations in parameters are not taken advantage of. It is also illustrated how realistic parameter distributions can be found using the Bootstrap method.

The simulations of open loop and closed loop operation illustrate how a poor uncertainty description

leads to uncertain predictions: even closed loop operation cannot save a poor uncertainty description. The simulations also illustrate how the uncertainties can be reduced by using a proper uncertainty description, and how feedback (closed loop operation) will further reduce the system sensitivity to parameter uncertainties.

References

Davison, A. & Hinkley, D. (1997), *Bootstrap Methods and their Applications*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, UK. ISBN 0-251-57471-4.

Efron, B. & Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, Chapman & Hall, New York.

Hastie, T., Tibshirani, R. & Friedman, J. (2001), *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, Springer Series in Statistics, Springer, New York. ISBN 0-387-95284-5.

Hauge, T. A. (2003), Roll-out of Model Based Control with Application to Paper Machines, PhD thesis, Norwegian University of Science and Technology, and Telemark University College, Norway.

Rawlings, J. B. & Ekerdt, J. G. (2002), *Chemical Reactor Analysis and Design Fundamentals*, Nob Hill Publishing, Madison, Wisconsin.

Ruiz, G., Casellano, M., González, W., Roca, E. & Lema, J. (2005), Anaerobic digestion process parameter identification and marginal confidence intervals by multivariate steady state analysis and bootstrap, *in* L. Puigjaner & A. Espuña, eds, 'European Symposium on Computer Aided Process Engineering', Elsevier, Amsterdam, pp. 1327–1332. ISBN 0 444 51991 2.

Development of a tool to improve the forecast accuracy of dynamic simulation models for the paper process

Georg Kamml
Papiertechnische Stiftung PTS
Hess-Strasse 134, 80797 Munich, Germany
Hans-Michael Voigt, Karsten Neß
Gesellschaft zur Förderung angewandter Informatik GFaI
Rudower Chaussee 30, 12489 Berlin, Germany

Abstract

1. INTRODUCTION

Evaluating the accuracy for dynamic simulation models is a complex task. Even today this is still being done by comparing simulated data and recorded mill data "by hand". As a result, the accuracy of dynamic process models is unknown in practice. Nonetheless, major investments and important decisions are taken on the basis of these calculations. Dynamic models are built, bought and used without any quality assurance. In order to overcome this situation, a quantified measure of the accuracy has to be defined and applied in the course of modeling practice.
In a research project we compared different mathematical methods for time series evaluation. Simple statistical methods such as medium average, bandwidth or standard deviation, but also more complex calculations such as spectral energy, main frequency and correlation analysis were evaluated. These were used to calculate characteristic parameters ratings for every time series, together with a combined rating for the whole simulation model.
A software tool which is still under development handles different simulation scenarios and the corresponding parameter sets for calibration. Graphical visualization tools and an optimization algorithm are used to help the user find optimized calibration parameters and improve the accuracy of simulation models.
This paper gives an outline of the methodology used to evaluate complex dynamic models. An insight into the functioning of the software is given and an example of how this tool was used in a simulation project will be presented.

2. Dynamic simulation in the paper industry and current validation methods

Dynamic simulation is currently used in the paper industry mainly in the area of operator training and in the optimization of dynamic process changes such as grade changes or the handling of web breaks. Currently available dynamic simulation tools are based mainly on dynamic conservation equations of mass, energy, momentum and component mass fractions, following first principle thermodynamic calculations. The usual starting point in building a simulator is the process and instrumentation diagram (P&ID) of a process. Objects representing single unit operations are placed on a work sheet and connected according to the P&ID. The user has to parameterize each object by providing design and

operation parameters. As not all of the required parameters can be measured, some of these parameters must be estimated. In addition, the model is a simplified version of the original process. This is done either to reduce the work involved in creating the model or because of unknown physical, chemical or thermodynamic relationships. To improve the accuracy of a model, guessed or estimated parameters can be used as calibration parameters.
As long as dynamic simulation models are used for process optimization, data from the original process can be used to evaluate the behavior of the model. Time series of parameters from the original process can be compared with corresponding data from the simulation model. In addition, the response behavior of the model to process changes can be compared as well. Calibration parameters are then adjusted to improve the model. As early as 1992, simple statistical methods were used to evaluate a model within the scope of a simulation study performed for Procter and Gamble [1]. But even now, this is mainly done "by hand", i.e. plotting the time series for single parameters or for parameter sets in one diagram, one line on top of the other. The accuracy of the model then is evaluated by optical assessment.

3. Methods for evaluating time series

The target of the evaluation is on the one hand to develop criteria to compare two different simulation runs, but on the other hand to also provide feedback about problem areas within the model or within a sub-part of the model. Based on the evaluation criteria, a decision then has to be taken as to which simulation run is closer to the real behavior of the process.
A problem in evaluation is the very different accuracy of the individual time trends. Depending on the complexity of the model and the grade of abstraction, very accurate models or models which just show the general behavior of the original process have to be evaluated. Still, the criteria should give a robust evaluation.

3.1 Evaluation of simulation results

The results of simulation are evaluated in several steps:
The first step involves calculating several individual criteria for every measured value.
In a second step, these individual criteria for every measured value are conjunctively combined to form a quality evaluation for the respective measured value.
The last step involves the disjunctive combination of the quality criteria of all measured values to form a total criterion for simulation.
The following individual criteria are calculated for every measured value:

1. **Sub-criterion A**: The deviation of the average of the simulated values from the average of the measured values
2. **Sub-criterion B**: The divergence of the simulated values relative to the measured values
3. **Sub-criterion C**: The concurrence of the spectral signal energy of the simulated and measured values in different frequency ranges.
4. **Sub-criterion D**: If there is close agreement between the signal curves of the simulated and measured value, covariance analysis is an option that can be employed as an additional criterion.

There is a simple possibility of implementing additional criteria for special problems that could not be predicted within the framework of the development.
All sub-criteria were scaled to the interval [0...1], a value of 1.0 corresponding to ideal agreement for this criterion and values < 1.0 signifying a deviation for this criterion.

- Calculation of the deviation from the mean

The criterion f1 for the deviation from the mean is calculated in accordance with the following formula:

$$f_1 = \begin{cases} \dfrac{m_{PLS}}{m_{Simu}} & \text{for } m_{PLS} \le m_{Simu} \\ \dfrac{m_{Simu}}{m_{PLS}} & \text{for } m_{PLS} > m_{Simu} \end{cases}$$

- Calculation of divergence

When calculating the deviation of divergence, the measured and simulated values are first smoothed with a low-pass filter. The difference between the measured and the simulated value is then calculated for every measuring point. The increase α for this time series of difference values is estimated by linear regression. An increase of ZERO is ideal.
Sub-criterion f2 is calculated from the estimated increase in accordance with the following equation:

$$f_2 = \frac{1}{1 + |\alpha|}$$

- Calculation of spectral signal energy agreement

First of all, a Hamming window is applied to the measured and simulated values to calculate the deviations in spectral signal energy. The Fourier transformation is then calculated by FFT for both measurement series. The Fourier transformation of the measurement series serves as the basis for calculating the spectral signal energy.
The entire frequency spectrum is subdivided into several overlapping regions. The spectral power density for each of these regions is determined by the integration of the spectral signal energy across the subregion. The quotient of the spectral power density of the measured and simulated values is formed for all subregions. These quotients are weighted and combined conjunctively to form criterion f3.

- Covariance analysis

The correlation factor r of the measured and simulated values is calculated for the subcriterion f4 and is scaled to the interval [0...1] according to the following equation:

$$f_4 = \frac{r + 1}{2}$$

If the signal curves for the measured and simulated value deviate too far, it is not reasonable to include the covariance analysis in the evaluation. This subcriterion can be deactivated in this case. The value for f4 is then set to 1.0.

- Combining the subcriteria

Subcriteria f1…f4 are combined conjunctively to form a total criterion for the respective value in accordance with the following equation:

$$f_{MW} = f_1 * f_2 * f_3 * f_4$$

Conjunctive combination was chosen, since all individual criteria for one value are to be improved in order to achieve an improvement in the simulation quality for this value. The total criteria of the individual values are combined disjunctively to form a total criterion for simulation in accordance with the following equation:

$$f_{Simu} = \frac{1}{n}\sum_{i=1}^{n} f_{MW_i}$$

Disjunctive combination was chosen, since slight deteriorations in individual values are allowable in the overall evaluation of simulation quality, if these values are compensated for by improvements in the evaluation of other values.

4 Usage of the validation assistant

To demonstrate the usage of the validation assistant, a simplified model is used that includes only four tanks representing a pulper, a pulper chest, a chest for dilution water and one pulp chest. The model also includes the controllers for stock consistencies and filling levels. The validation assistant uses the above mentioned methods to evaluate the simulation model.
The validation assistant is available as a prototype in a stand-alone version and in a second version that is linked to the IDEAS simulation tool. The stand-alone version reads original and simulation data from files, the linked version uses a COM-interface to exchange data with the simulation tool. This enables the validation assistant to use optimization strategies to find optimized calibration parameters.
In a first step of the validation procedure, the user has to define starting parameters for all objects in the simulation model. Most of these parameters are quite well known and do not need any optimization. A first simulation run of the model is started using these standard parameters. In this phase, the validation assistant can be used to easily visualize trend lines from the original data and simulation data (fig. 1).
After modifying one calibration parameter, the impact of this parameter on the time trends can be visualized. This is done within a polar coordinate plot, which makes it easy to see the principal impact of the modified parameter on the trend lines without difficulty (fig. 2). The production rate was modified in this example. This parameter has an impact on the filling level of all chests, but no impact on the consistency after the pulper chest. Fig. 3 shows the result of optimizing the production rate. The evaluation of the default simulation is shown in the inner area (green), the evaluation with optimized production rate results in a bigger area (red). As expected, the evaluation is improved for all trends representing filling levels, but not for the trend representing the consistency after the pulper chest.

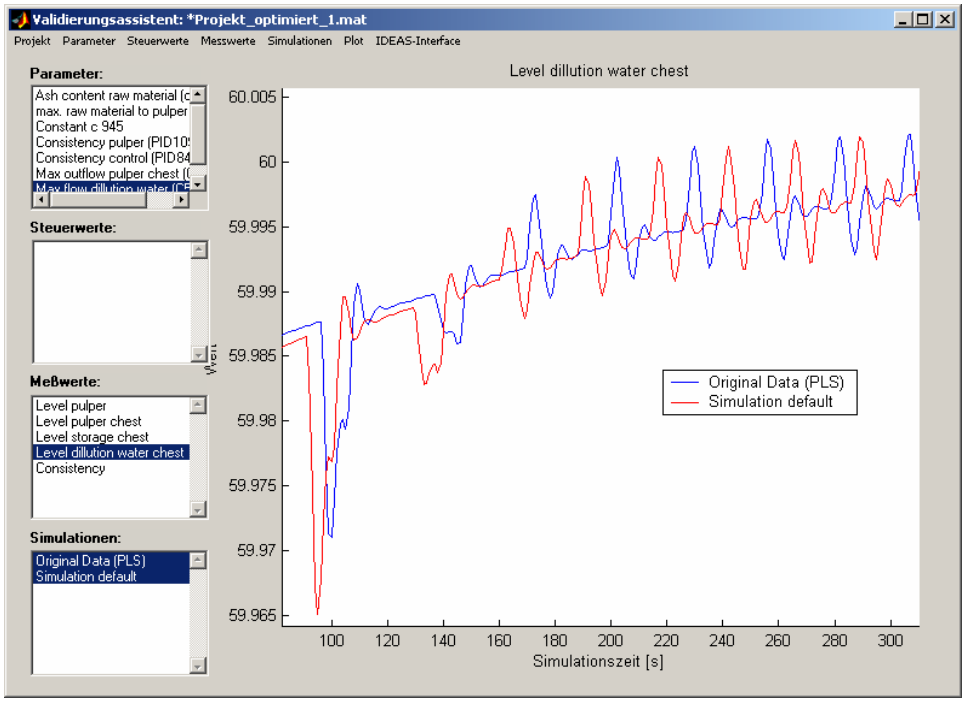


Fig. 1 The validation assistant can be used for easy visualization of trend lines

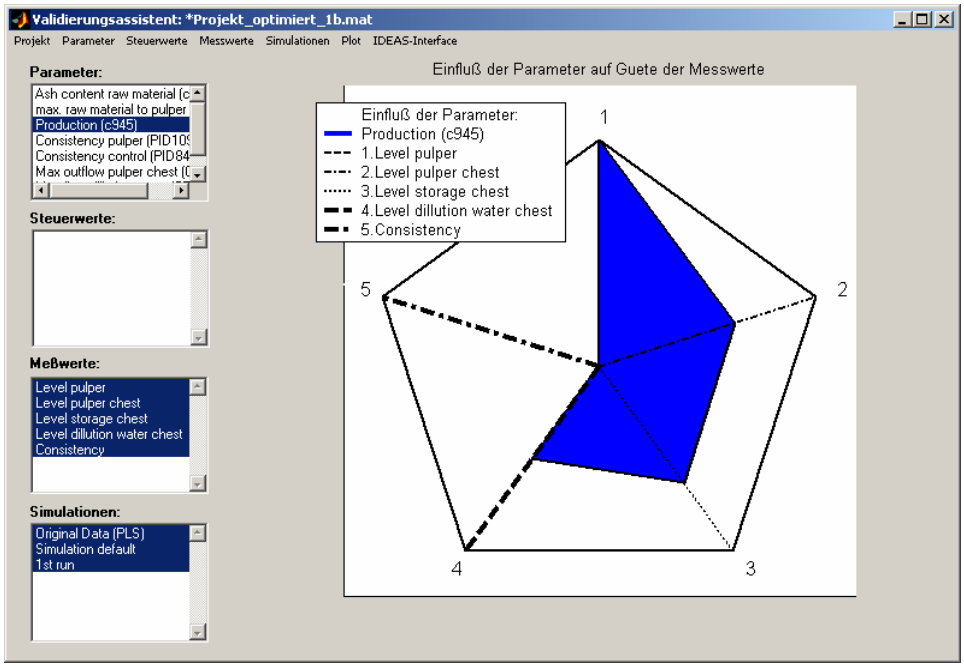


Fig. 2 Impact of one modified parameter on the single trend lines

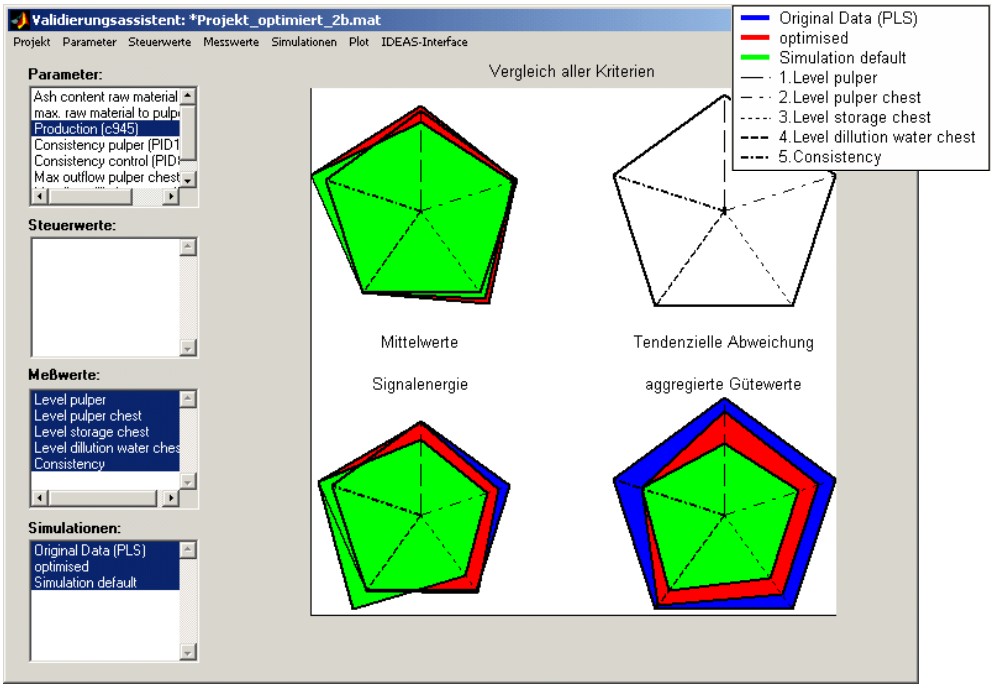


Fig. 3 Evaluation of all trend lines after optimization of one parameter

5. OPTIMIZER

From an optimization point of view, the simulation model constitutes a black box. The optimality of the system is measured by an objective as outlined in section 3. In general, information characterizing the functional properties of the objective like differentiability, smoothness etc. is missing. A special Evolution Strategy [2] as an incarnation of a stochastic search process is therefore used.
The optimization module uses an improved (1+1)-Evolution Strategy. The fitness function for the Evolution Strategy is given by the quality f_{Simu}. This objective is calculated after each simulation run of the model with any tested parameter set. During optimization, two parameter sets are compared by the objective f_{Simu}. The parameter set with higher fitness (greater value for f_{Simu}) will be selected as a parent for the next step. A child is generated by parameter variation (mutation) and its fitness value is calculated after simulation of the model.
The evolution strategy that was used is based on the Local Evolutionary Search Enhancement by Random Memorizing (LESRM) [3]. It uses a memory for former good solutions to decrease time consumption and to increase robustness of the search process. The memory is accessed at random to extract new, potentially good solutions.
The advantages of LESRM compared to other stochastic optimizers are robustness and efficiency. A disadvantage of the LESRM algorithm is space complexity. For an increasing number of parameters N, the space complexity of the randomly accessed memory increases by N^3. This is relevant for high-dimensional optimization problems. By using two concepts from nature, it was possible to reduce the space consumption of the memory to N without increasing time complexity. The two concepts used were:

1. Stochastic selection of incoming data
2. Forgetting stored data with low information gain

To avoid significant problems of elitist selection in Evolution Strategies, an aging of the parents was inserted. In nature, aging means a decrease in the fitness of an individual during its lifetime. We adopted this behavior for the (1+1)-Evolution Strategy with respect to the fitness function and the number of parameters to be optimized.
In addition, we extended the (1+1)-Evolution Strategy to constraint handling. Constraints are given by technological limits. The limits have to be tested in the model, and the results of all tests are aggregated to form one Boolean value which indicates the occurrence of a constraint violation. This value is used by the optimization process to generate technologically feasible solutions.

6. CONCLUSIONS

The validation assistant is now ready for use and has been tested on two different dynamic simulation models.
Using the validation assistant, it is possible to evaluate the accuracy of a dynamic process model versus original data from a process on an objective base. It is no longer merely a matter of visually comparing trend lines. This will help the user of a simulation model to have more confidence in the results obtained from the simulation model. This will also help to improve the acceptance of simulation results by the decision-makers in industry. Moreover, the validation assistant together with the included optimizer, will not just ensure the quality of simulation models, but will also speed up the process of model calibration and therefore make dynamic modeling cheaper. This will help to further increase the usability of dynamic simulation in process optimization, where up to now dynamic models were often not used due to the high costs of the models.

References

1 Fletcher A. Hinson, Ralph L. Kappelhoff:
”A validation procedure for first principle process dynamics simulation”;
TAPPI Proccedings: 1992 Engineering Conferrence p. 763 ff

2 Rechenberg, I:
“Evolutionsstrategie ’94”; Stuttgart,Bad Cannstatt: Frommann-Holzboog 1994

3 Voigt, H.-M., J.M. Lange:
„Local Evolutionary Search Enhancement by Random Memorizing”; Proceedings of the 1998 IEEE World Congress on Computational Intelligence –
The 1998 IEEE International Conference on Evolutionary Computation (ICEC'98), pp. 547 – 552,Piscataway: IEEE Press 1998

Validation as a crucial step for improving robustness of models: Application to paper quality predictions

A. Alonso, A. Blanco and C. Negro. Complutense University of Madrid. Spain.

I. San Pío. HPM. Spain.

1. INTRODUCTION

1.1. Objective

Papermaking industry has, in general, an important need of knowledge about how they can optimise the usefulness of their data by applying different modelling methods. Validation techniques are needed to check model performance and to improve the robustness of the model [1].

The objective of this study is to compare different types of validations, focusing on applications carried out to predict the paper properties based on wet-end performance at a paper mill.

Model building and validations have been carried out in a newsprint machine, involving more than 7,500 variables (process variables or product properties) and measurements from a focused beam reflectance measurement (FBRM) device [2]. The parameters selected for the prediction model were selected based on know-how and previous experiences at the mill using statistical analysis.

1.2. Methodology

From the initial set of variables, only 50 were pre-selected for modelling. Statistical analysis has allowed to reduce the number of model inputs to less than 10 in last Artificial Neural Network models. Predictions for different quality parameters have been very accurate (e.g. R^2 for paper formation predictions over 0.74). The next step was to optimise model usefulness and robustness through appropriated validation procedures.

Validation methodologies that have been carried out are the following:

- Cross validations for multiple regression models constructed for different periods of time.
- Validation with data that have been randomly extracted from model developing. The validation error has also been used as the stopping criteria in a first artificial neural network model. (Random validation)
- Mixed validation, that includes both data taken at the same period of time than data taken for model building and data from different periods of time. The objective of this procedure is to cover different operating conditions by including validation from the same periods of time than the used for model building, and assure robustness (in terms of generalization capabilities) by including validation data from different periods of time.

2. RESULTS AND DISCUSSION

2.1. Parameters selection

From the initial pre-selection of 50 process variables and product quality parameters (see table 1), several of them were chosen as the most important ones to track product properties, and thus the final set of parameters was established.

Table 1.- Studied parameters.

VARIABLE	UNITS	VARIABLE	UNITS
Headbox Consistency (DCS) PM61	g/l	Grammage (PaperLab) PM61	g/m²
Headbox Consistency (Lab) PM61	%	Thickness (PaperLab) PM61	µm
Headbox Ash content (DCS) PM61	g/l	Formation Index (PaperLab) PM61	
Headbox Ash content (Lab) PM61	%	Tensile Strength (PaperLab) PM61	kN/m
WW consistency (DCS) PM61	g/l	Tear Strength (PaperLab) PM61	Nm²/kg
WW Consistency (Lab) PM61	%	Elongation (PaperLab) PM61	
WW Ash content (DCS) PM61	g/l	Density (PaperLab) PM61	kg/m³
WW Ash content (Lab) PM61	%	Hand (PaperLab) PM61	cm³/g
Total Retention (DCS) PM61	%	Break Length (PaperLab) PM61	km
Total Retention (Lab) PM61	%	Long. TEA (PaperLab) PM61	J/m²
Ash Retention (DCS) PM61	%	Longitudinal TSI (PaperLab) PM61	kNm/g
Ash Retention (Lab) PM61	%	Transversal TSI (PaperLab) PM61	kNm/g
Flocculation Index PM61		Porosity (PaperLab) PM61	ml/min
DIP2 Pulp Flow	m³/n	Opacity (PaperLab) PM61	%
DIP3 Pulp Flow	m³/h	Brightness (PaperLab) PM61	%
Brokes Pulp Flow	m³/h	Moisture in Pope PM61	%
Flocculant dosage PM61	g/Ton	Base paper Ash content PM61	%
Coagulant Dosage PM61	g/Ton	Ash content in Pope PM61	%
Ultimer dosage PM61	g/Ton	Moisture in Size Press PM61	%
CaCO3 dosage PM61	l/h	Counts/s (1-5 µm)	Cts/s
Wire roll speed PM61	m/min	Counts/s (5-34 µm)	Cts/s
pH in PIT		Counts/s (34-1000 µm)	Cts/s
		Mean, 1/Lth Wt	µm
		Mean, Sqr Wt	µm
		Median, No Wt	µm
		%<5.41, No Wt	%
		%>34.15, No Wt	%

Final set of parameters:

- Input variables:
  - FBRM data:
    - Median chord length of CLD (Median, No Wt)
    - Counts per second 1-5 µm.
    - Counts per second 5-34 µm.
    - Counts per second 34-1000 µm.
  - Machine speed, given as head roll speed.
  - Consistency of the headbox pulp.
  - Ash content in headbox pulp.
  - Deinked pulp (DIP2) flow.
  - Deinked pulp (DIP3) flow.
  - Broke pulp flow.
  - Flocculant dosage.
  - Coagulant dosage.
  - Ultimer dosage.
  - Carbonate dosage.

- Output variables:
  - Formation index.
  - Tensile strength.
  - Elongation.

- Tear strength.
- Paper ash content in pope.

These output variables were proposed by the mill personnel as the more representative and critical ones for their products.

2.2. Data filtering

A data filtering step is needed to consider only those data that are really representative of the phenomena happening in the mill. Several filtering steps are needed:

1. Machine stops and starting ups.
2. Product classification. In order to build accurate models, each product should be considered separately, since some non considered variables can change between them. This is more important when changes are made between different paper qualities. In Table 2 the product distribution by months can be observed, with the number of data that have been acquired during this period.
3. Breaks. In order to improve data accuracy and to determine how representative they are of the paper machine phenomena, data corresponding to breaks have been discarded.
4. Product out of specifications. Since some of the reels produced can be out of specifications, data corresponding to those products should be removed from the model, since they do not correspond to normal behaviour of the paper machine.
5. Abnormal data. This last step in data filtration corresponds to discarding all of those data that moves away from common values for a given parameter. The general criteria for this data removing has been selected considering the values as a normal distribution and removing all the data that has values lower than the mean minus three times the standard deviation or higher than the mean plus three times the standard deviation.

2.3. Model building

Statistics and advanced modelling methods give different alternatives for modelling [3,4,5]. From all the alternatives, three different approaches for finding data correlations and model building have been considered:

- Monthly multiple regression.
- Multiple regression with all data since July.
- Neural network modelling.

2.3.1. Monthly multiple regression

Two different regression models are used for each case: one considering a constant factor (K in tables), which can reflect all the situations that are not considered in the model and another without this constant. Since we are using 14 input variables, the first model will give 15 parameter values and the second one 14.

The correlation coefficients and model parameters, that are calculated by random validation with data from the same period of time than those used for modelling, change from one month to the others and, therefore, the model will need to be checked for each specific case. Quite good correlations are found in different periods, as it can be observed in table 2 for each month and the different principal products.

Table 2. Correlation coefficients by months and product type.

WITH K	JULY				AUGUST			
	NP A	NP B	NP C	FCO B	NP A	NP B	NP C	FCO B
FORMATION INDEX	-	0.76	0.77	0.82	0.87	0.44	0.68	-
TENSILE STRENGTH	-	0.63	0.63	0.52	0.46	0.39	0.54	-
QCS POPE ASH CONTENT	-	0.39	0.69	0.7	0.42	0.43	0.8	-
ELONGATION	-	0.34	0.72	-	0.88	0.6	0.89	-
TEAR STRENGTH	-	0.39	0.78	0.62	0.51	0.4	0.64	-

NO K	JULY				AUGUST			
	NP A	NP B	NP C	FCO B	NP A	NP B	NP C	FCO B
FORMATION INDEX	-	0.75	0.67	0.82	0.87	0.37	0.68	-
TENSILE STRENGTH	-	0.62	0.61	0.52	0.45	0.29	0.54	-
QCS POPE ASH CONTENT	-	0.35	0.69	0.68	0.36	0.43	0.78	-
ELONGATION	-	0.34	0.72	-	0.88	0.58	0.88	-
TEAR STRENGTH	-	0.39	0.77	0.59	0.5	0.4	0.63	-

WITH K	SEPTEMBER				OCTOBER			
	NP A	NP B	NP C	FCO B	NP A	NP B	NP C	FCO B
FORMATION INDEX	0.94	0.69	-	-	0.58	0.42	-	0.76
TENSILE STRENGTH	0.53	0.58	-	-	0.52	0.54	-	0.69
QCS POPE ASH CONTENT	0.55	0.56	-	-	0.55	0.3	-	0.9
ELONGATION	0.57	0.39	-	-	0.83	0.62	-	0.3
TEAR STRENGTH	0.81	0.33	-	-	0.52	0.26	-	0.77

NO K	SEPTEMBER				OCTOBER			
	NP A	NP B	NP C	FCO B	NP A	NP B	NP C	FCO B
FORMATION INDEX	0.94	0.69	-	-	0.55	0.4	-	0.75
TENSILE STRENGTH	0.45	0.58	-	-	0.49	0.5	-	0.66
QCS POPE ASH CONTENT	0.55	0.53	-	-	0.51	0.3	-	0.89
ELONGATION	0.57	0.39	-	-	0.81	0.62	-	0.28
TEAR STRENGTH	0.81	0.32	-	-	0.51	0.25	-	0.75

It is observed that:

- Correlation coefficient for one specific grade and one specific model can change drastically when the analysis time period is different (For instance in table 3 see correlation for the model with K for Newsprint B in July and August).
- Correlation coefficients are always better when using models with K, as it was expected.
- FCO B production gives in general good correlations with all the analysed process variables except for elongation.

It is observed that there are big differences when comparing the parameter data and the values obtained in the models for the different months, which shows the difficulty of obtaining a general model. However it was decided to carry out a trial applying the model from one month to predict the data from another month. As it was expected, the results are not good enough as it is shown in figure 1.

2.3.2. Full data multiple regression

Since the monthly multiple regression approach was not good enough, a new trial has been made considering as input data those corresponding to the period comprising July to end of October, in order to see if the correlation is still good enough or some of the obtained data for each month were casual.

The multiple regression model was developed only for the two most common products, which are Newsprint A and Newsprint B. The obtained correlation coefficients are summarized in table 3. The mean values of the coefficients obtained for each month separately are also shown to compare the model accuracy.

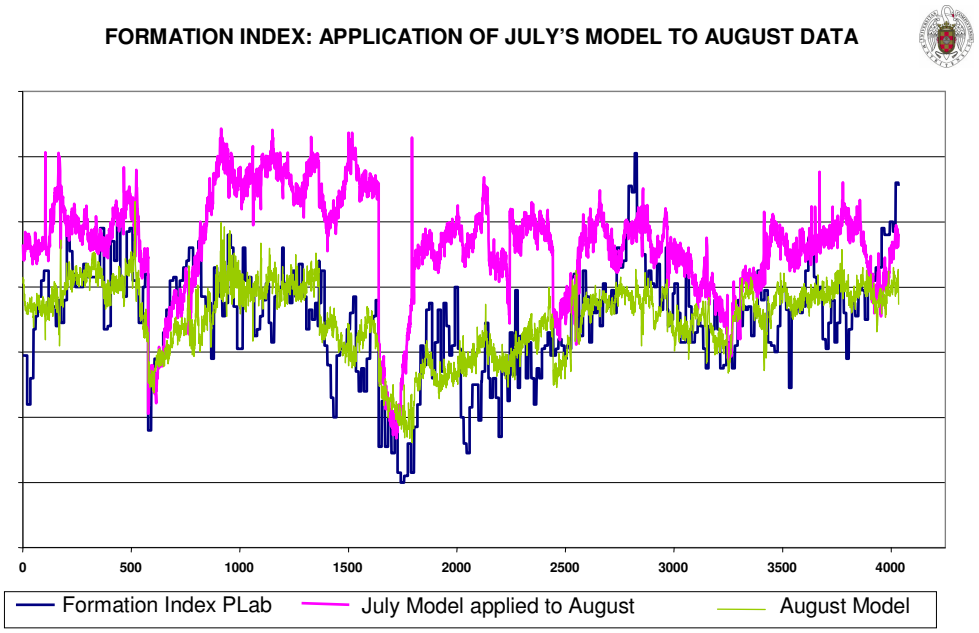


Figure 1. Formation index model developed with data from July applied to data from August.

Table 3. Correlation coefficients obtained with the full period, compared to the mean values of the correlations for each month.

	NPA		MONTHLY MEAN VAL	
	WITH K	NO K	WITH K	NO K
FORMATION INDEX	0.84	0.84	0.8	0.79
TENSILE STRENGTH	0.41	0.41	0.5	0.46
QCS POPE ASH CONTENT	0.36	0.33	0.51	0.47
ELONGATION	0.55	0.52	0.76	0.75
TEAR STRENGTH	0.5	0.49	0.61	0.61

	NPB		MONTHLY MEAN VAL	
	WITH K	NO K	WITH K	NO K
FORMATION INDEX	0.59	0.59	0.58	0.55
TENSILE STRENGTH	0.47	0.47	0.53	0.5
QCS POPE ASH CONTENT	0.23	0.14	0.42	0.4
ELONGATION	0.34	0.33	0.49	0.48
TEAR STRENGTH	0.4	0.36	0.34	0.34

As it can be observed, only for the formation index a good model is really obtained, mostly with the data from Newsprint A productions. This model is considered only adequate as a prediction tool for the formation index value.

As an example, figures 2 and 3 show the model predicted data for the formation index for the two type of products studied.

2.3.3. Artificial Neural Networks (ANNs)

An artificial neural network (ANN) can be defined as a data processing system consisting of a large number of simple, highly interconnected processing elements (artificial neurons) in an architecture inspired by the structure of the cerebral cortex of the brain [6].

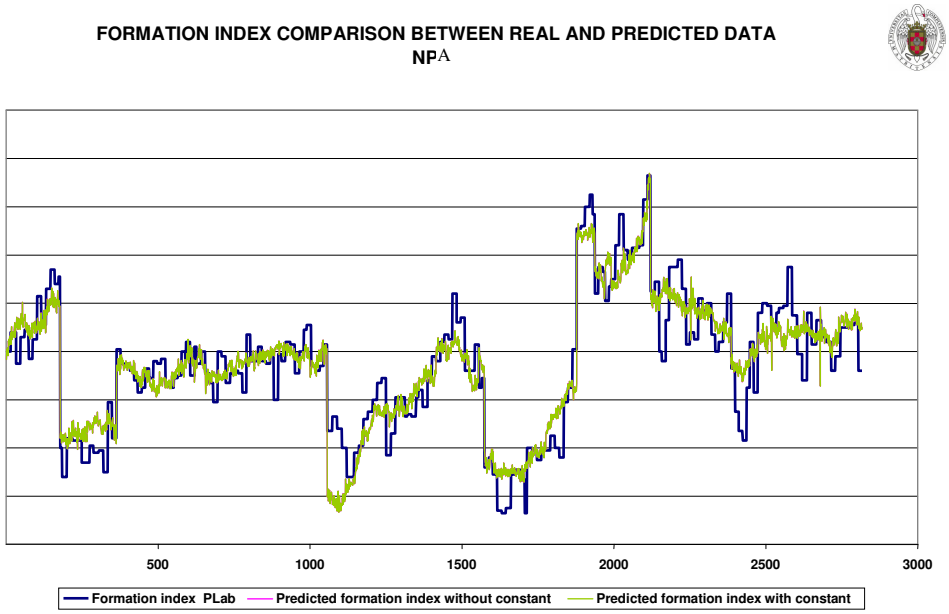


Figure 2. Formation index comparison between real and predicted data for Newsprint A (Random validation). Magenta and green lines are almost overlapped.

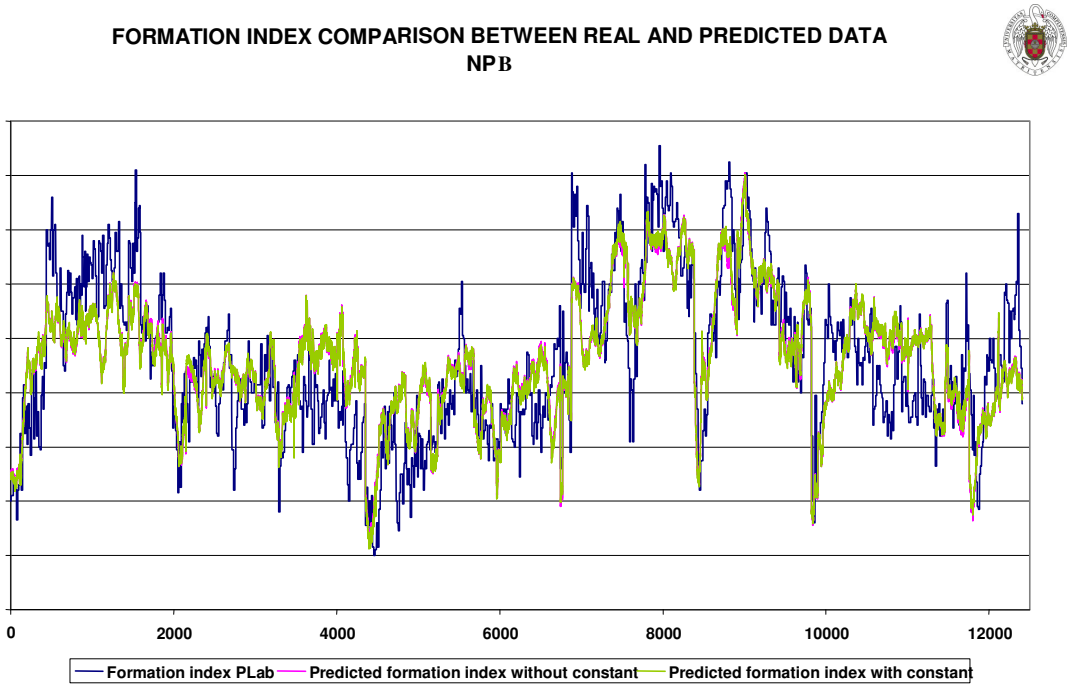


Figure 3. Formation index comparison between real and predicted data for Newsprint B (Random validation). Magenta and green lines are almost overlapped.

A deeper approach to data modelling is being performed by means of neural network modelling. Very promising results have been obtained. For example the R^2 for the correlation between the predicted and the real formation index for Newsprint B is 0.73 (figure 4).

In order to compare different types of validation, two different ANNs have been developed with data from July to September 2004 and Newsprint B grade. Minimum validation error has been

the used criterion to select optimum number of epochs. The difference between these ANNs is the validation dataset.

- In ANN-1, 8593 mill data have been used for training and 2858 mill data have been used only for random validation (25% of total data from the same period of time).
- In ANN-2, 8593 mill data have been used again for training and 4067 mill data have been used for mixed validation (2858 data from July-September 2004 period and 1209 data from October 2004).

ANN-1

Figure 4 shows an example of comparison between laboratory data and predicted data for the period July-September 2004. Results at this period are logically the best, because validation data are quite closed to data used for model training.

For all the outputs, a correlation index 'R' is calculated to evaluate each prediction. Figure 5 shows correlation coefficients for the analysed product quality parameters obtained with the artificial neural network, as well as average errors for each parameter. These errors are calculated as a percentage of the difference between the minimum and maximum values of each quality parameter in the training dataset.

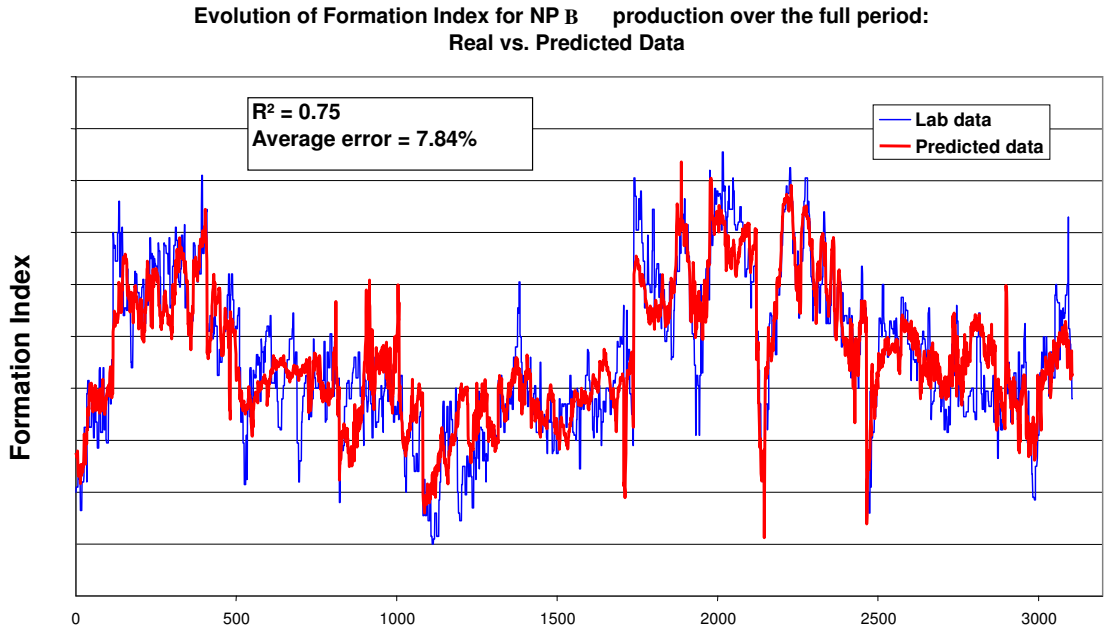


Figure 4. Neural network model comparison to real formation index data.

This model has been simulated with data from October 2004 and November 2004, obtaining the errors summarised in figure 5. This ANN does not take into account data from October or November 2004 at validation, so differences in accuracy for each month are probably due to how closer are operating conditions respect to training period.

ANN-2

This ANN has been trained with mixed validation. Results from validation and simulation for October and November 2004 are also summarised in figure 5.

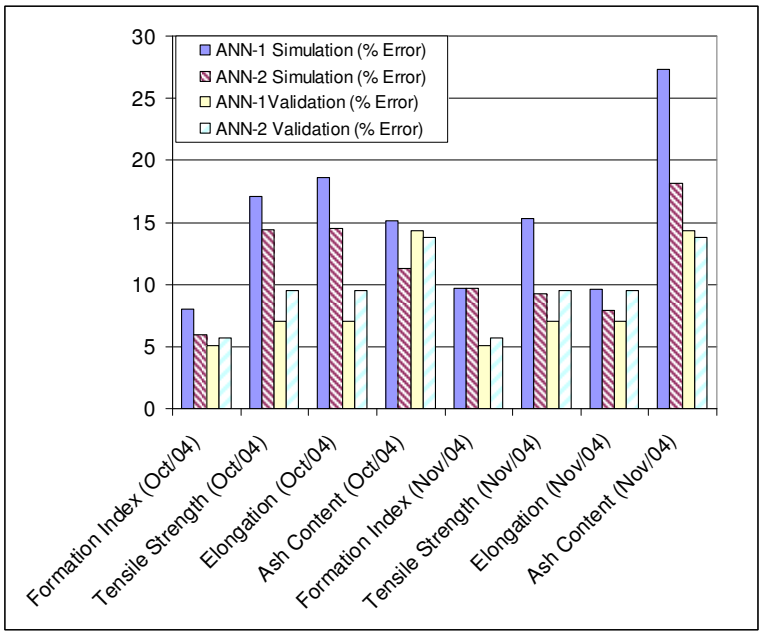


Figure 5. Average errors for the different product quality parameters with Neural Network modelling and random validation (ANN-1) or mixed validation (ANN-2). Errors are calculated with training data (validation) and with other data (simulation).

At validation, ANN-1 gives logically in general better results than ANN-2. That fact is due to the training process itself, with validation data from only the same period of time than training data (July-September 2004), which are quite similar. But when we simulate different periods of time, ANN-2 (built with validation data not only from the period July-September 2004, but also with October 2004) has better generalization capabilities.

Table 4 summarises the average value for all the errors considering all the ANNs' outputs. At validation, ANN-1 seems to be 13% better, but when we simulate, errors are 25% smaller if working with validation data with broader periods of time, even with those out of training periods (ANN-2). So, working in that way gives more robustness to the model.

This fact can be also denoted if we compare errors at validation with errors at simulation in each ANN. In ANN-2, values are much closer and thus, applying the model at different periods gives similar and robust results.

Table 4. Average errors of developed ANNs in each working step.

	Validation	**Simulation**	
		October 2004	**November 2004**
ANN-1	8.2	14.7	15.5
ANN-2	9.3	11.5	11.3

3. CONCLUSIONS

Artificial Neural Networks are flexible tools that give accurate and robust results when the behaviour of the process is difficult to understand and there are not equations modelling that process.

Best predictions are achieved for the formation index.

A very accurate model at a specific period of time can be very useful to diagnose troubles or optimize a very specific variable. However, when medium or long term predictions are of interest for the mill, models have to be robust, even when accuracy becomes smaller.

Validation becomes a crucial step when modelling. If used as stopping criterion when training, the quality of further predictions strongly depend on the methodology used.

Mixed validation achieves more robust results than random validation. Thus, trying with different possibilities of mixed validation would improve the robustness of the results.

Using ANN models with mixed validation for longer periods of time would require a more exhaustive pre-processing step, but would improve the robustness of such models. In further studies, an optimum will be established for the amount of time used to build the ANNs and then, future on-line models will include self-adaptive rules with time as 'learning new values' or 'forgetting the oldest ones'.

REFERENCES

[1]. A. Blanco, A. Alonso, C. Negro and I. San Pío, Advanced data treatment to improve quality in a newsprint machine. Proceedings of the "Simposium Internacional: Nuevos Desarrollos en Tecnologías del Reciclado del Papel", Instituto Papelero Español, Barcelona, Spain. 19-20 May 2005.

[2]. Blanco A., Fuente E., Negro C. and Tijero J., Flocculation monitoring: focused beam reflectance measurement as a measurement tool, Can. J. Chem. Eng., Vol. 80, Nº 4, pp 734-740. 2002.

[3]. Box G., Hunter W.G. and Hunter J.S., Statistics for experimenters, 1ª ed., John Wiley & sons, Inc., New York. 1978.

[4]. Himmelblau, D., Process Analysis by Statistical Methods, 1ª ed., John Wiley & sons, Inc., New York. 1969.

[5]. Wang H., Wang A.P. and Duncan S.R., Advanced process control in paper and board making, Pira Technology Series, Pira International, Surrey, UK. 1997.

[6]. Tsoukalas, L.H. and Uhrig, R.E., Fuzzy and Neural Approaches in Engineering, Ed. Wiley Interscience, . 1997.

A Simulation Study of the Validity of Multivariate Autoregressive Modeling

Olli Saarela
KCL
P.O. Box 70, Fin-02151 Espoo, Finland

Abstract

1. INTRODUCTION

A multivariate autoregressive (MAR) model is a time series model, where each new reading of a measurement is explained by a linear combination of the past values of the same measurement and of the other measurements in the model. MAR models are used (among other things) in the analysis of the dynamics of industrial processes, as this model structure provides a reasonable compromise between the ability of the model to describe the dynamics of the process and the ease of model identification and use.

MAR modeling has proved to be a valuable tool in solving process fluctuation problems. Solved cases have been published from paper mills [11], conventional power plants [3], nuclear reactors [10], a cement kiln [2], an aluminum roll mill [15], and a food processing plant [13]. However, the use of the MAR modeling approach has been impeded by insufficient understanding of the reliability of its results.

This paper describes how factors affecting the reliability of MAR modeling have been investigated in a simulation study, and how the results from the study have been used to automate process analysis, including model validation.

2. PROCESS FLUCTUATIONS

Temperatures, flows and other values in industrial processes rarely hold exactly at their target values. Fluctuation is caused by, e.g., wear in actuators and variations in the quality of raw materials. If the control equipment cannot attenuate the fluctuation sufficiently, the efficiency of the process is decreased, the quality of the end product is deteriorated and the wear of process equipment is accelerated.

Finding the origin of process fluctuations (*fault isolation*) is often complicated, process plants typically containing several tightly interconnected control loops. Fluctuations in one loop are reflected to other loops, and often the whole process or a significant part of it displays the same fluctuation symptoms. A large number of controllers can be involved in the fluctuation mechanism. On the other hand, all practical actuators have some degree of imperfection, contributing in some degree to the overall process fluctuation. Finding the origins of process fluctuations involves determining the mutual significance of the various fluctuation sources in the process.

Analysis of process fluctuations cannot be feasibly based on first principles modeling, as the building a sufficiently detailed model of the dynamics of an industrial process would be an overwhelming task. Hence, data-based black box modeling is an important tool in the analysis of process fluctuations.

3. THE MAR MODEL

3.1 The model structure

The canonical form of the MAR model is expressed as

$$\mathbf{x}(k) = \sum_{l=1}^{M} \mathbf{A}(l)\mathbf{x}(k-l) + \mathbf{e}(k) \qquad (1)$$

where the vector x(k) represents the measured time series at sample index k, $\mathbf{A}(l)$ is a square matrix of constant model coefficients for each value of delay l and the term $\mathbf{e}(k)$ represents the prediction error. M is the model order, which represents the maximum process delay that the model can describe.

The coefficient matrix **A** represents the transfer functions between the variables in the model. As transfer functions from each variable to every other variable are included, the model is capable of representing feedback controls and material recirculations. Model identification requires no a priori knowledge of which transfer functions actually exist.

The selection of the model order is most often based on a criterion such as the Akaike Information Criterion (AIC) [1], whereas the model coefficients are usually identified by the so-called LWR algorithm [8]. The model orders selected for the purpose of process analysis are typically significantly higher than the model orders used in control system design and controller tuning. Lower model order allows for more robust transfer function estimates, but less detail in the frequency responses.

3.2 The prediction errors as fluctuation sources

The transfer functions in an MAR model describe the propagation of process fluctuations from one variable to another. The prediction errors describe the fluctuation components that cannot be predicted from previous measurement values using the linear model structure. Process phenomena modeled in the prediction errors include, e.g., quality variations in the raw materials, non-linearity caused by malfunctions such as hysteresis in valve actuators, and sensor noise. As a practical simplification, the malfunctions causing process fluctuations are modeled in the prediction errors of the MAR model.

Consequently, the MAR model in Equation 1 can also be thought of as a digital filter, where the prediction errors $\mathbf{e}(k)$ are considered as inputs called the noise sources. Their effects propagate through the transfer functions **A** and results in process fluctuations observed in variables **x**.

A source of process noise is modeled in the prediction error of the variable in whose measured time series its effect is first observed. For example, the hysteresis of a valve actuator is most often represented in the prediction error of a flow measurement close to the valve. Naturally, the sample interval of the measurement data has to be sufficiently short compared to process delays, so that the propagation of fluctuation from one measurement to another can be observed.

In order to locate the origin of fluctuation, the mutual significance of the noise sources in the model has to be evaluated. If some noise sources can be determined to be insignificant, they can be excluded from further analysis. The actual malfunction can then be searched for in the neighborhoods of the remaining variables.

The mutual significance of the noise sources is most often evaluated in the frequency domain, in order to separate the effects of control loops resonating at different characteristic frequencies. A frequency transform of Equation 1 yields the estimated spectrum matrix $\mathbf{S_x}(f)$ of the variables

$$\mathbf{S_x}(f) = \mathbf{H}(f)\mathbf{\Sigma_e}\mathbf{H}^*(f) \tag{2}$$

Where $\mathbf{\Sigma_e}$ represents the covariance matrix of the noise sources, the asterisk denotes the complex conjugate transpose, and

$$\mathbf{H}(f) = \left(\mathbf{I} - \mathbf{A}(f)\right)^{-1} \tag{3}$$

is the transfer function matrix from each noise source to each variable in the model [9]. If the noise sources can be perfectly separated from each other, the covariance matrix $\mathbf{\Sigma_e}$ is diagonal and the power spectral density of each variable x_i can be expressed as the sum of effects from each of the noise sources

$$S_{\mathbf{x}_{ii}}(f) = \sum_{j=1}^{m} \left|H_{ij}(f)\right|^2 \Sigma_{\mathbf{e}_{jj}} \tag{4}$$

The relative effect of the noise source e_j to the fluctuation of x_i is given by

$$\Gamma_{ij}^{(\mathrm{diag})}(f) = \frac{\left|H_{ij}(f)\right|^2 \Sigma_{\mathbf{e}_{jj}}}{S_{\mathbf{x}_{ii}}^{(\mathrm{diag})}(f)} \cdot 100\% \tag{5}$$

where Γ_{ij} is called the noise source contribution and the superscript $^{(\mathrm{diag})}$ indicates the assumed perfect separation of the noise sources.

3.3 Inseparable noise sources

In practical process analysis, the noise sources can never be perfectly separated from each other. The most common approach to handling of the off-diagonal elements of $\mathbf{\Sigma_e}$ is to check whether they are smaller than some heuristic threshold value, and if so, to simply ignore them.

There have been various approaches to handle significant off-diagonal elements of $\mathbf{\Sigma_e}$ [14]. Methods requiring a priori process knowledge are not very practical for problems involving a large number of process measurements. Methods based on grouping mutually correlated noise sources together assume that clear group structures can be found and introduce new heuristic parameters for determining the groups. The approach adopted in this work is based on the Cholesky decomposition of the noise covariance matrix

$$\mathbf{\Sigma_e} = \mathbf{L}\,\mathbf{I}\,\mathbf{L}^* \tag{6}$$

Using this decomposition, a maximal noise source contribution

$$\Gamma_{i1}^{(\max)}(f) = \frac{\left[\mathbf{H}(f)\,\mathbf{L}\,\mathbf{I}^{(1)}\,\mathbf{L}^*\,\mathbf{H}^*(f)\right]_{ii}}{S_{x_{ii}}(f)} \cdot 100\% \tag{7}$$

can be defined [6]. This approach assigns to the noise source e_1 as much of the correlated part of the noise source power as possible, while being consistent with the identified $\mathbf{\Sigma_e}$. Equation 7 defines the maximal noise source contribution from the noise source e_1 to each variable x_i. To calculate the corresponding noise source contributions from other noise sources, the variables in the model have to be rearranged.

A fault isolation algorithm based on computing the maximal noise source contributions has been developed in [12]. This algorithm has been used to find the origins of fluctuations in several industrial processes.

4. INVESTIGATION OF FACTORS AFFECTING THE RELIABILITY OF MAR MODELLING

Despite the success in solving practical cases, the reliability of MAR modeling is inadequately understood. The theory supporting this technique makes several assumptions that are not valid when actual industrial processes are analyzed. These assumptions include the linearity and the stationarity of the process, the normal distribution and the whiteness of the noise sources and the infinite length of the data available for analysis. Theoretical confidence limits for MAR analysis have been derived [7], but they are based on assumptions not met in actual processes, such as process linearity or the Gaussian distribution of the fluctuation.

In most of the previous work to validate MAR modeling, data has been generated using a linear simulation model and the identified MAR model has been compared with the original simulation model. [4,5]

To achieve more realistic a setup, a relatively detailed simulation model was used. This model was built for operator training at unit No. 2 of the Ormond Beach power plant. The model is based on first-principles modeling and includes fault modes related to process dynamics, such as stiction in valve actuators. The simulation model has over three thousand variables.

Due to the non-linearity of the simulation model, there is no "correct" linear model with which the identified MAR model could be compared. Instead, a more application-oriented approach was developed. The validity of the MAR modeling was evaluated by considering the ability of the MAR analysis procedure to locate the inflicted “process” fault and to estimate its significance in each simulation set-up.

The text below describes a simulation experiment for determining the level of process noise required for reliable MAR modeling.

4.1 Level of process noise

MAR modeling is based on observing the propagation of process noise from one measurement to another (i.e., no step experiments are executed). Reliable model identification requires that some unique information is propagated through every transfer path in the model. Without sufficient noise sources in the process, the fluctuation propagates too deterministically in a feedback loop, which decreases the reliability of MAR model identification. However, the actual level of noise required for modeling isn’t known.

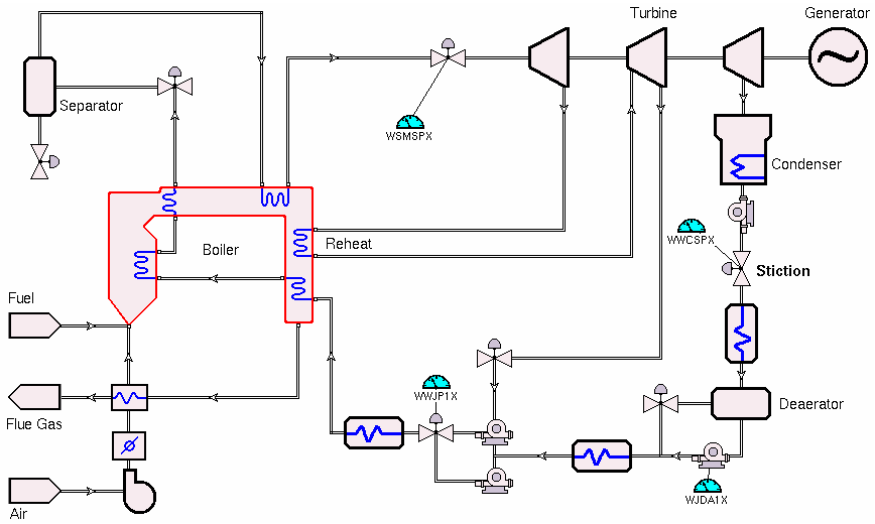


Figure 1. A simplified process diagram of the simulator. The measurements used in the experiment are shown in the diagram. The stiction was introduced to the inlet valve of the deaerator.

A considerable stiction (5%) was introduced to the valve controlling the deaerator input flow, and four closely related mass flow rate measurements were selected as variables for the MAR modeling:

- WJDA1X: Mass flow rate at the outlet of the deaerator (x_1),
- WWCSPX: Mass flow rate at the inlet of the deaerator (x_2),
- WWJP1X: Mass flow rate in feedwater pumps (x_3) and
- WSMSPX: Mass flow rate in the main steam pipe (x_4).

These measurements were selected, because they form a system that is closely connected, yet controlled by several control circuits. The measurements are shown in Figure 1.

In the simulation model, the level of noise was controlled by adjusting the level of sensor noise. Some of the sensors are used in feedback control, which connects the sensor noise back in the process. Figure 2 shows the trend plot of the x_4. The relatively sharp changes in the trend are caused by corrective actions of several controllers. As can be seen from the plot, the process is not very stationary. In addition, process non-linearity makes the fluctuation complicated: as the noise level is increased, the need for control actions becomes more frequent and the frequency of the fluctuation changes.

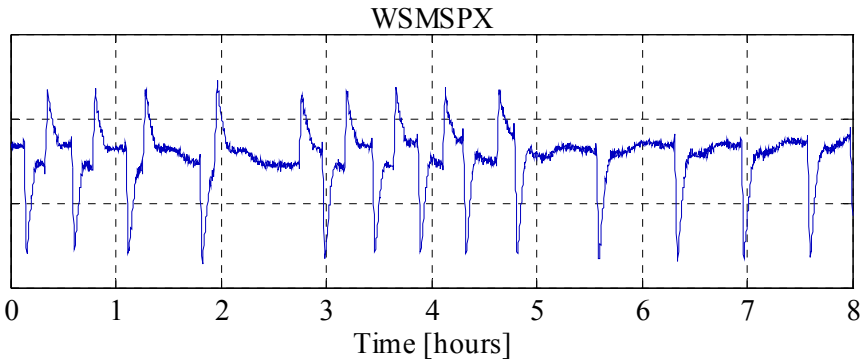


Figure 2. Fluctuation of the mass flow rate in the main steam pipe.

The simulation was executed using noise levels from (a somewhat realistic) 2% down to (unrealistically low) 0.001%, where the percentage figure indicates the standard deviation of the sensor noise in relation to the sensor reading. At each noise level, the simulation was executed both with and without the inflicted malfunction.

When the level of sensor noise was low, the fluctuation was dominated by the stiction in the deaerator input valve. The fluctuation of x_4 was almost completely eliminated when the stiction was removed. When the level of sensor noise was gradually increased, the plant-wide effect of the valve stiction was decreased, and the system became less deterministic. The process fluctuation became dominated by the sensor noise, and removing the stiction had a smaller effect on the fluctuation.

Figure 3 shows, at each noise level, the values of three characteristic figures:

- $\delta\sigma^2_{x4}$ indicating the change in the variance of x_4 when the stiction was removed. This change is plotted as a percentage to make it better comparable to the noise source contributions.

- $\Gamma_{42}^{(diag)}$ indicating the noise source contribution from e_2 (the noise source nearest to the actuator with the stiction), computed by ignoring the off-diagonal elements of the noise covariance matrix.
- $\Gamma_{42}^{(max)}$ indicating the maximal noise source contribution from e_2.

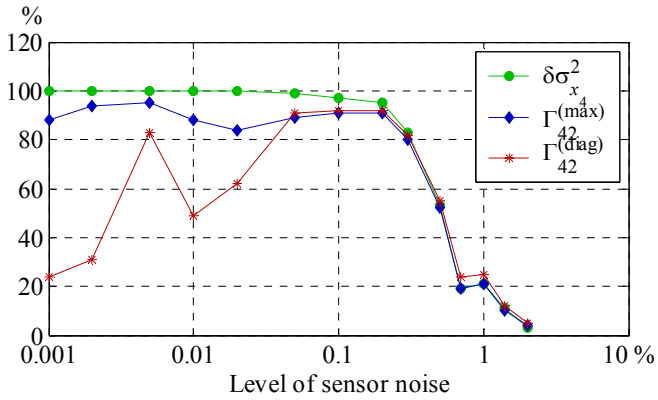


Figure 3. The maximal noise source contribution to variable x4 and the reduction in the power of the fluctuation of x4.

At reasonably realistic noise levels the three characteristic figures are in good agreement with each other, and the noise source contribution values could be used to predict the decrease in fluctuation level when the malfunction was fixed.

When using the maximal noise source contribution, adequate MAR model performance was achieved also with measurement signal-to-noise ratios much better than what is possible with contemporary instrumentation. This suggests that the level of process noise is not a practical concern when the maximal noise source contribution computed from a MAR model is used to locate the origin of a process fluctuation.

4.2 Various types of fluctuation sources

The main purpose of the simulation experiments was to gain further understanding about the applicability of the MAR modeling in solving fluctuation problems in continuous industrial processes. In the simulation model, process fluctuations were generated by simulating the following malfunctions:

- Mistuning a controller.
- Increased measurement noise.
- Valve stiction.
- Valve hysteresis.

These malfunctions were introduced at various control loops at several degrees of severity and the increased fluctuation was observed at various measurements across the process. The ability of the MAR modeling to locate the origin of fluctuation was evaluated by comparing the noise source contributions to the relative reduction of variance when the malfunction was removed, as discussed above.

Whenever a process malfunction caused a detectable change in process spectra, MAR analysis was able to locate the origin of that fluctuation. For valve hysteresis and stiction, the degree of simulated malfunction severity at which the spectral change became detectable was found to be comparable to those encountered in actual processes. The most important assumptions (process linearity, noise source properties) of the theory behind MAR analysis, which are not valid in actual processes, were found not to prevent successful application of the technique.

5. CONCLUSIONS

Validation of an identified process model can be quite a complicated task, and analyzing a process fluctuation can require identification of dozens of time series models. Furthermore, personnel analyzing process fluctuations don't necessarily have system identification or time series analysis as their primary interest or education. Hence, automatic model validation is highly desirable.

The formal theory supporting MAR modeling doesn't provide adequate tools model validation, as its assumptions (such as linearity) are not valid in practical processes. Better understanding of model validity can be gained with simulation experiments, where sufficiently realistic process model is used.

References

[1] Akaike, H., "A new look at the statistical model identification", IEEE transactions on automatic control, 19 (6), 1974.
[2] Akaike, H. & Nakagawa, T., "Statistical analysis and control of dynamic systems", Dordrecht: Kluwer Academic Publishers, 1988.
[3] Eklund, J., et.al., "Multivariate signal analysis – an efficient tool for power plant process analysis", in: Signal processing in measurement – 6th International symposium on measurement theory and its application to practice, Budapest, 10-12 June 1987.
[4] Heinola, K., "Testing multivariate autoregressive modeling", Thesis (MSc), Tampere University of Technology, 1988. In Finnish.
[5] Hoogenboom, J.E., et.al., "Summary of the benchmark test on artificial noise data", in: SMORN V, symposium on reactor noise, Munich 12-16 October 1987. Garching: Gesellschaft für Reaktorsicherheit.
[6] Ihalainen, H., "Multivariate signal analysis in Matlab", Tampere: Tampere University of Technology, 1988.
[7] Ljung, L., "System identification: theory for the user", 2nd ed., Upper Saddle River: Prentice-Hall, 1999.
[8] Kailath, T., "Linear estimation for stationary and near-stationary processes", in: T. Kailath, ed. Modern signal processing, Washington DC: Hemisphere Publishing, 1985.
[9] Marple, S.L., "Digital spectral analysis with applications", Englewood Cliffs; Prentice-Hall, 1987.
[10] Oguma, R., "Investigation of resonant power oscillation in Halden boiling water reactor based on noise analysis", Journal of nuclear science and technology, 17 (9), 1980.
[11] Rantala, S., et.al., "Process dynamics study based on multivariate AR modeling", in: IECON'88 14th annual conference of IEEE Industrial Society, Singapore 1988.
[12] Saarela, O., "Multivariate autoregressive analysis in locating the origin of fluctuation in continuous industrial processes", Tampere University of Technology, Publications 366, Tampere 2002.
[13] Saarela, O. & Ihalainen, H., "MAR Navigator – process diagnostics with simple user interface encapsulating advanced mathematics", Control systems 2000, Victoria, 1-4 May 2000.
[14] Suoranta, R., "Analysis of linear systems applying multivariate autoregressive model", Thesis (Licentiate), Tampere University of Technology, 1990.
[15] Upadhyaya, B.R., Glockler, O. & Eklund, J., "Multivariate statistical signal processing technique for fault detection and diagnostics", ISA transactions, 29 (4), 1990.

Dynamic validation of multivariate linear soft sensors with reference laboratory measurements

Kimmo Latva-Käyrä and Risto Ritala
Tampere University of Technology
Institute of Measurement and Information Technology
P. O. Box 692, FIN-33101 Tampere, Finland

Abstract

This paper presents a method for optimal detection of need for re-estimation of the parameters of multivariate soft sensors based on infrequent and irregular reference measurements. We assume that the validity of parameters deteriorates through a stochastic process. In particular, we give analytic distributions of parameters when deterioration is through random-walk diffusion. Method is demonstrated and its detection capabilities analyzed in cases where the actual deterioration is either a step change, a linear drift or a random walk.

1. INTRODUCTION

Soft sensors can be understood as steady-state simulators with process measurements as inputs. The soft sensor models may be physico-chemical, black-box or combinations of these two, often referred to as gray box models.

In black-box and gray box models the model parameters are identified from process and laboratory data with input measurements from the process and the output to be soft-sensed measured in laboratory. However, as the direct physico-chemical meaning of the model parameters is rather unclear, the parameters may depend on the process structures and conditions at the instant of model identification. Thus the parameters often change over time and the predictive power of the soft sensor diminishes.

In order to keep the soft sensor reliable, occasional laboratory measurements are made about the property soft-sensed and compared with the soft sensor predictions. Such occasional data sets are typically too small to serve as a basis for re-identification of the parameters. However two alternative questions may be answered on the basis of such data:

- is there a need for proper recalibration, or
- how much the model parameters should be adapted on the basis of such data?

In this paper we shall shortly describe a general Bayesian framework for answering these questions, assuming that the degradation of model parameters can be described as a stochastic differential equation. We shall give explicit results for the case of multivariate linear soft sensor and integrated white noise degradation of parameters.

2. BAYESIAN ESTIMATION OF PARAMETER DISTRIBUTION

2.1 General case

We shall denote the input signals to the soft sensor by vector s, the output by a scalar x, and model parameters by a vector β. The stochastic dependence describing the soft sensor model and the uncertainty of the soft sensor is the conditional probability density function (pdf)

$$f_{X|s,\beta}(x \mid s, \beta) \tag{1}$$

where the capital letters refer to stochastic variables and small letters to their values. The soft sensor output is the maximum likelihood value:

$$\hat{x} = \arg\max_x f_{X|s,\beta}(x \mid s, \beta) \tag{2}.$$

We shall consider the model parameters β also as random variables. After model identification the uncertainty in parameters is described by the joint probability distribution

$$f_B^{(0,+)}(\beta) \tag{3}.$$

Now let us assume that we know the pdf of parameters after *(n-1)th* laboratory reference measurement x_{n-1} to be $f_B^{(n-1,+)}(\beta)$. Furthermore, let us assume that after this reference measurement the (unknown) changes in model parameters are described with a stochastic differential equation

$$\frac{d\beta}{dt} = F(\beta, \xi) \tag{4},$$

, where ξ is some stochastic process.

We can use (4) to solve for the pdf of β at later time instants, in particular at time when we get the *nth* laboratory reference measurement, $f_B^{(n,-)}(\beta)$ (the minus sign in superscript denotes the distribution *before* reference measurement is taken into account).

By using Bayes formula, by assuming that – in case of several simultaneous reference measurements – the measurements are independent, and by using $f_B^{(n,-)}(\beta)$ as a priori information, we get

$$f_B^{(n,+)}(\beta) = N_1 * f_B^{(n,-)}(\beta) * \prod_{i=1}^{N} f_{X,S|\beta}(x_i, s_i \mid \beta) \tag{5}.$$

Here N_1 is an uninteresting normalization factor. Furthermore as s and b are statistically independent when marginalized over x [i.e. $f_{S,\mathrm{B}}(s,\beta) = f_S(s) * f_{\mathrm{B}}(\beta)$], we can write (5) as

$$f_B^{(n,+)}(\beta) = N_1 * f_B^{(n,-)}(\beta) * \prod_{i=1}^{N} f_{X|s,\beta}(x_i \mid s_i, \beta) \qquad (6).$$

Now using (1) and (4) defines a recursion between $f_B^{(n-1,+)}(\beta)$ and $f_B^{(n,+)}(\beta)$.

The need for recalibration can be detected through that the actual coefficients β_{actual} are exceptional according to $f_B^{(n,+)}(\beta)$ at the probability p_0 of false alarms, i.e. $\beta_{actual} \in A$, with

$$\begin{aligned} A &= \left\{\beta \mid f_B^{(n,+)}(\beta) < f_0(p_0)\right\} \\ p_0 &= \int\limits_{f_B^{(n,+)}(\beta) < f_0(p_0)} f_B^{(n,+)}(\beta) d\beta \end{aligned} \qquad (7)$$

The new best estimates of parameters – to be implemented at each step or when need for recalibration is detected – are given as the maximum likelihood values:

$$\hat{\beta}_n = \arg\max_{\beta} f_B^{(n,+)}(\beta) \qquad (8).$$

2.2 Linear – Gaussian – random walk case

Let us consider a linear model $x = \beta_0 + \beta_c^T s$. In what follows we denote $\beta \equiv \left[\beta_0 \quad \beta_c^T\right]^T$. Assuming model errors Gaussian we have

$$f_{X|s,\beta}(x \mid s, \beta) = (2\pi\sigma_e^2)^{-1/2} \exp\left[-\frac{1}{2\sigma_e^2}(x - \beta_0 - \beta_c s)^2\right] \qquad (9).$$

Hence

$$\begin{aligned} &\prod_{i=1}^{N} f_{X|s,\beta}(x_i \mid s_i, \beta) = \\ &N_2 * \exp\left[-\frac{N}{2\sigma_e^2}(\beta - B^{-1}\hat{\beta})^T B (\beta - B^{-1}\hat{\beta})\right] \end{aligned} \qquad (10)$$

with

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & B_{0.} \\ B_{.0} & B_{..} \end{bmatrix}$$

$$\begin{aligned} B_{.0} &= B_{0.}^T = \frac{1}{N}\sum_i s_i \\ B_{..} &= \frac{1}{N}\sum_i s_i s_i^T \\ \hat{\beta}_0 &= \frac{1}{N}\sum_i x_i \\ \hat{\beta}_c &= \frac{1}{N}\sum_i x_i s_i^T \end{aligned} \tag{11}$$

Let us assume that after the nth update with N_n simultaneous observations the distribution is normal and thus characterized by the mean $\mu_n^{(+)}$ and covariance matrix $\Delta_n^{(+)}$:

$$f_{\mathrm{B}}^{(n,+)}(\beta) = N_3 \exp\left[-\frac{1}{2}(\beta - \mu_n^{(+)})^T (\Delta_n^{(+)})^{-1}(\beta - \mu_n^{(+)})\right] \tag{12}$$

We assume that the uncertainty in *b* degrades through a random walk process:

$$\begin{aligned} &\frac{d\beta}{dt} = \Gamma(t) \\ &<\Gamma(t)>=0; <\Gamma(t)\Gamma(t')^T> = D*\delta(t-t') \end{aligned} \tag{13}$$

Then

$$f_{\mathrm{B}}^{(n+1,-)}(\beta) = N_3 \exp\left[-\frac{1}{2}(\beta - \mu_n^{(+)})^T (\Delta_n^{(+)} + D*(t_{n+1} - t_n))^{-1}(\beta - \mu_n^{(+)})\right] \tag{14}$$

Inserting (10) and (14) into (6), we see that $f_B^{(n-1,+)}(\beta)$ is Gaussian:

$$f_{\mathrm{B}}^{(n+1,+)}(\beta) = N_3 \exp\left[-\frac{1}{2}(\beta - \mu_{n+1}^{(+)})^T (\Delta_{n+1}^{(+)})^{-1}(\beta - \mu_{n+1}^{(+)})\right] \tag{15}$$

with

$$\begin{aligned} \mu_{n+1}^{(+)} &= \left(N_{n+1}B/\sigma_e^2 + (\Delta_n^{(+)} + D*(t_{n+1} - t_n))^{-1}\right)^{-1}(N_{n+1}\hat{\beta}_n/\sigma_e^2 + (\Delta_n^{(+)} + D*(t_{n+1} - t_n))^{-1}\mu_n^{(-)}) \\ \Delta_{n+1}^{(+)} &= (N_{n+1}B/\sigma_e^2 + (\Delta_n^{(+)} + D*(t_{n+1} - t_n))^{-1})^{-1} \end{aligned} \tag{16}$$

Therefore our assumption of Gaussian in (12) is self-consistent and an explicit recursion between the distribution parameters has been established. The recursion is initialized with the known mean and covariance estimates for parameters of linear multivariate models.

The test for exceptional values of parameters is that for Mahalanobis distance [3]

$$A_n^{(+)} = \left\{ \beta_{actual} \mid (\beta_{actual} - \mu_{n)}^{(+)})^T \left(\Delta_n^{(+)} \right)^{-1} (\beta_{actual} - \mu_{n)}^{(+)}) > F_d^{-1}(p_0) \right\} \tag{17}.$$

with k as the dimensionality of β.

Here we used the inverse function of

$$p_0 = F_d(C) = \frac{\Gamma(d/2, c/2)}{\Gamma(d/2)} \tag{18}.$$

The best estimate for the parameters during $t \in [t_n, t_{n+1})$ is $\mu_n^{(+)}$.

3. EXAMPLE

In this chapter, we shall analyze the performance of parameter updating of our algorithm with simulated cases. Also the detection capabilities of this algorithm are analyzed. These simulations include the response of the parameter updating algorithm to fault situations, demonstrating potential and actual situations happening in real factory.

Our signal is the form of $x = a_0 + a_1 s_1 + a_2 s_2 + a_3 s_3$, where s_1, s_2 and s_3 are random variables and a_0 =1, a_1 = 3, a_2 = 3, a_3 = 3. Figure 1 shows a simulated example of updating parameters when the true parameters undergo a diffusion. Original estimation of parameters at $t = 0$ is based on 100 signal-reference observation pairs. The reference values become available individually at irregular intervals evenly distributed between 2 and 4 time units. The diffusion coefficients for parameter estimation are selected as 0.001or 0.0001 for diagonal elements and 0 for off-diagonal.

The simulation results show that the estimated parameters follow the true parameters rather closely. Smaller variances would be obtained with smaller diffusion coefficients in the updating.

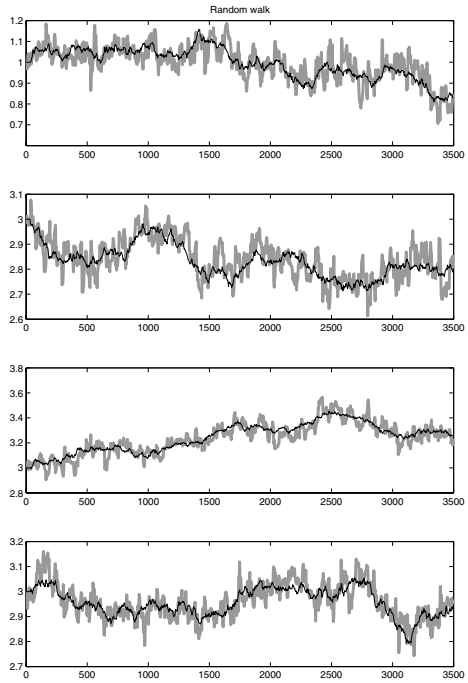


Figure 1. Updating of distribution parameters, assuming diffusion parameter D_{00}= D_{11}= D_{22}= D_{33}=0.001, D_{12}=0. Original parameters a_0 =1, a_1 = 3, a_2 = 3, a_3 = 3. μ (grey) and true (simulated) a (black).

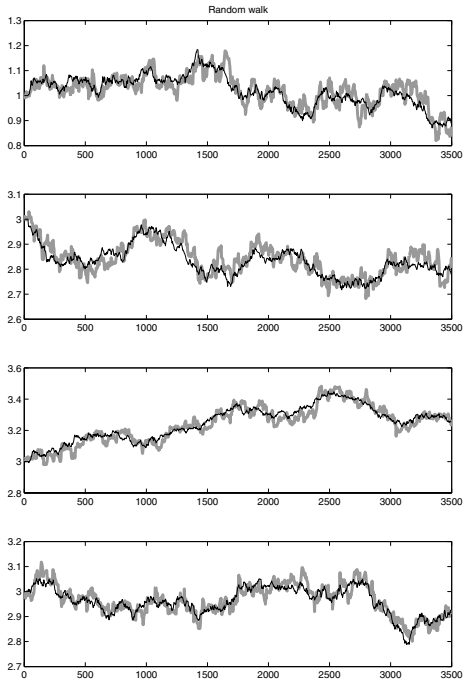


Figure 2. Updating of distribution parameters, assuming diffusion parameter D_{00}= D_{11}= D_{22}= D_{33}=0.0001, D_{12}=0. Original parameters a_0 =1, a_1 = 3, a_2 = 3, a_3 = 3. μ (grey) and true (simulated) a (black).

We simulated the detection capability of our algorithm in four cases. The time series of the test variable of the cases are shown in figures 1 and 2. The cases were:

- In the first case the true coefficients (a_0 =1, a_1 = 3, a_2 = 3, a_3 = 3) were constant throughout the analysis period.
- At the second case all true coefficients undergo a random walk, such that the standard deviation of per unit time of the random step size is 0.002.
- At the third case the true coefficient are constant till time t=1050 and then start to drift at the rates of + 0.0002/time unit in a_0 and -0.0002/time unit in a_1, a_2 and a_3
- And the fourth case, where step change occurs at time t=1748. True coefficients change (a_0) from 1 to 1.5, (a_1) from 3 to 2.5, (a_1) from 3 to 3.5 and (a_3) is constant.

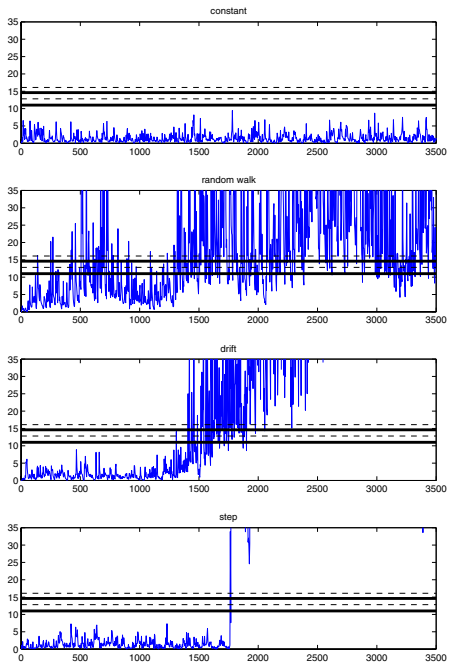


Figure 3. Detection of abnormalities at characteristic curve parameter estimation, assuming diffusion parameter D_{00}= D_{11}= D_{22}= D_{33}=0.001, D_{12}=0. From top to bottom: no change in calibration parameters, random walk, linear trend and step change.

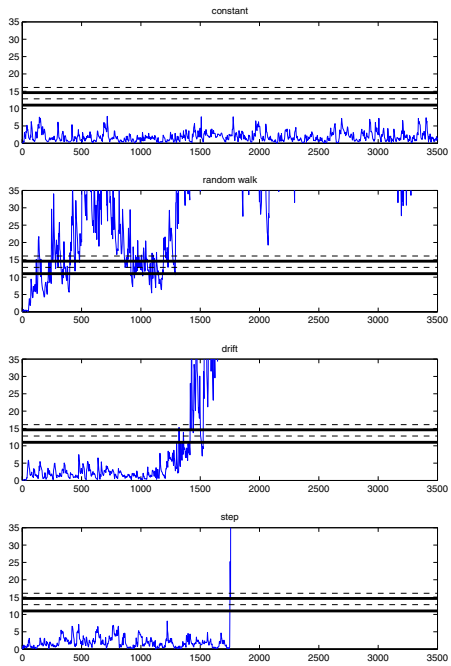


Figure 4. Detection of abnormalities at characteristic curve parameter estimation, assuming diffusion parameter D_{00}= D_{11}= D_{22}= D_{33}=0.0001, D_{12}=0. From top to bottom: no change in calibration parameters, random walk, linear trend and step change.

Abnormality set was calculated with Mahalanobis distance in Eq. (17) and detection limits given with expression Eq. (18).

Detection algorithm finds step change fast and reliable. Abnormality test signal is increasing about 10 time units later than step change in true coefficients occurred and all limits are crossed about 15 time units later compared to change. With smaller diffusion coefficients the detection is couple of time units slower.

At the case of linear drift, abnormality test signal begins increasing about 50 time units later than drift starts. All limits are crossed about 350 time units later compared to change in true coefficient. With smaller diffusion coefficients the detection is slightly slower but there is less noise in the test signal

The difference in diffusion coefficients seen in figures 1 and 2 can easily be seen at abnormality test signal in random walk case. In the case of larger diffusion coefficient the estimated parameters response is much faster compared to the small diffusion coefficients. Case where true coefficients were constant, if the limits are too tight the rate

of false alarms increases depending on the random part and measurement uncertainty. With smaller diffusion coefficients the abnormality test signal is more tranquil.

4. CONCLUSIONS

In this paper we have shown method of optimal detection of need for re-estimation of the parameters of multivariate linear soft sensor with integrated white noise degradation of parameters. The weakest link in this method is finding the diffusion parameters and they are to be understood as tuning parameters of this method.

This method can easily be applied to any automation system and used as

- early warnings system for operators or engineers,
- detection system for the need of re-estimation of the parameters of the soft sensor,
- to continuously update the parameters of the soft sensor,

and thus making the whole process more efficient. Also on-line sensor faults can be detected and isolated.

References

[1.] Ihalainen H., Latva-Käyrä K. & Ritala R.," Dynamic validation of on-line measurements: a probabilistic analysis", submitted to Measurement.
[2.] Latva-Käyrä K. & Ritala R.," Sensor Diagnostics based on Dynamic Characteristic Curve Estimation", 10th International Conference on Technical Diagnostics, Budapest, Hungary, June 9-10th, 2005.
[3.] Dillon W. R.& Goldstein M., "Multivariate Analysis – Methods and Applications", Wiley, USA 1984, 587 p.

Fibre Classification – Model Development and Validation

Klaus Villforth and Samuel Schabel
Darmstadt University of Technology
Chair of Paper Technology and Mechanical Process Engineering
Alexanderstraße 8, 64283 Darmstadt, Germany

Abstract

This paper presents the classification of fibres in technological motivated classes for quality control and process optimization. Fuzzy membership functions weigh the fibre morphology and associate every single fibre with memberships to fibre classes. Particular models for every class estimate the fibre mass, which are summed up to fibre fractions.

The fuzzy fibre classifier has been validated by visual classified fibres. The results of the fibre classification are correlating with the Bauer-McNett classifier and process trials.

1. INTRODUCTION

The properties of fibres are important for quality control and process optimization. Innumerable scientific papers and numerous research projects are involved in the measurement of fibre properties. The Bauer-McNett classifier (BMC) is a common source of information on fibre length distribution [1]. The other sources are fibre analyzers mainly based on laser optics and image analysis. They can provide information on fibre length, coarseness, curl and kinks, width, cell wall thickness and fibrillation.

The diversity of fibres and the various process steps lead to distribution functions which are difficult to interpret. Very often the effects of beating and sorting can be seen in the BMC fractionations while the fibre length distribution shows no significant differences. To overcome this phenomenon, the analyzers should classify the fibres according to technological aspects. Based on the fibre classes in table 1 an even more detailed classifier has been developed.

Table 1: Fibre classes based on technological aspects [2]

Type of fibre	**Dimension**	**Notes**
Fibril	length: 50 to 200 µm width: 1 µm	part of fines
Flour	length: 20 to 30 µm width: 1 to 30 µm	cubic structure, small specific surface, part of fines
Short fibre	length: 25 to 800 µm width: 2 to 3 µm	mostly latewood fibres and fibre fragments
Long fibre	length: 0.8 to 4 mm width: 25 to 80 µm	mostly earlywood
Shives	length: < 5 mm width: 1 to 3 mm	fibre bundles

The developed classifier distinguishes between fibrils, flour, plain short fibre, fibrillated short fibre, short fibre bundle, plain long fibre, fibrillated long fibre, long fibre of extreme width and long fibre bundle. Since the fibre analyzer delivers not enough information for a definite grouping, a model based on fuzzy membership functions classifies the fibre.

The study is based on FiberLab™ which also stores fibre images. Metso Automation provided a preliminary software version with supplementary measurements for the FiberLab. UPM-Kymmene contributed a fibre analysing software which has been developed by Hirvonen, VTT [3,4]. Every measurement consists of thousands of stored fibre images with FiberLab and VTT data sets for the single fibres.

2. CLASSIFYING FIBRE MORPHOLOGY

A membership represents the probability of being a member of a fibre class. These memberships functions overlap for the most part. A fuzzy interference system filters and combines the memberships to a single fuzzy set, which represents the membership to the distinct fibre classes. Stored fibre images have been visual classified by experts to create a data base. The fuzzy membership functions and the interference system are derived from this data base of visual classified fibres, their images and the measured properties.

2.1 Membership functions

The data base of visual classified fibres has been processed with MATLAB. The Distribution Fitting Tool helps to characterise the fibre classes. Almost all fibre properties, i.e. length, width, wall thickness, fit a lognormal distribution as shown in Fig. 1 for the fibre length of fines, short and long fibres.

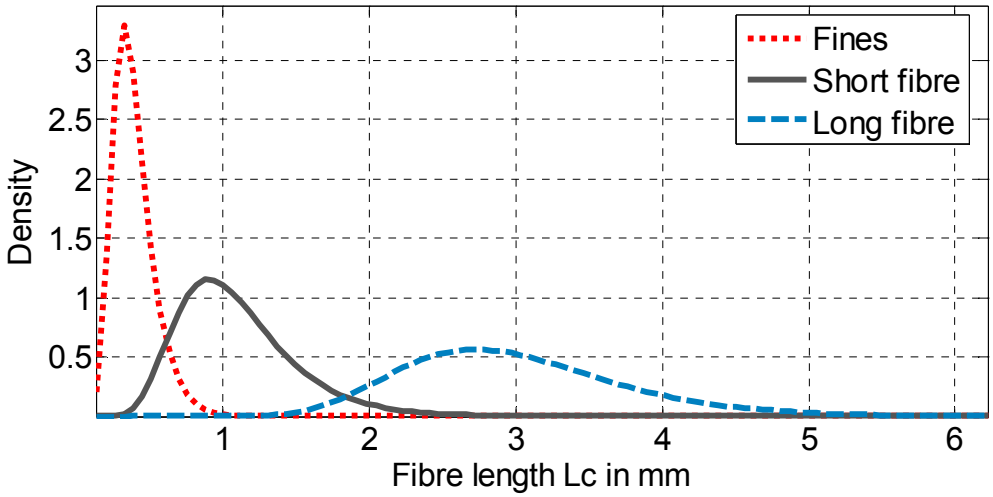


Figure 1. Lognormal fibre length distribution

The developed fuzzy membership function encloses the confidence bounds of the log-normal distribution. The interval $\mu_1 <= \mu <= \mu_2$ defines the region of membership = 1.

$$f(x,\mu_1,\sigma) = \frac{1}{x} e^{\frac{-(\ln x-\mu_1)^2}{2\sigma^2}+(\mu_1-\frac{1}{2}\sigma^2)} \quad \text{for} \quad x \le e^{\mu_1-\sigma^2} \tag{1}$$

$$f(x) = 1 \quad \text{for} \quad e^{\mu_1-\sigma^2} < x \le e^{\mu_2-\sigma^2} \tag{2}$$

$$f(x,\mu_2,\sigma) = \frac{1}{x} e^{\frac{-(\ln x-\mu_2)^2}{2\sigma^2}+(\mu_2-\frac{1}{2}\sigma^2)} \quad \text{for} \quad x > e^{\mu_1-\sigma^2} \tag{3}$$

From these equations the fuzzy memberships for L_c are calculated as shown in Fig. 2.

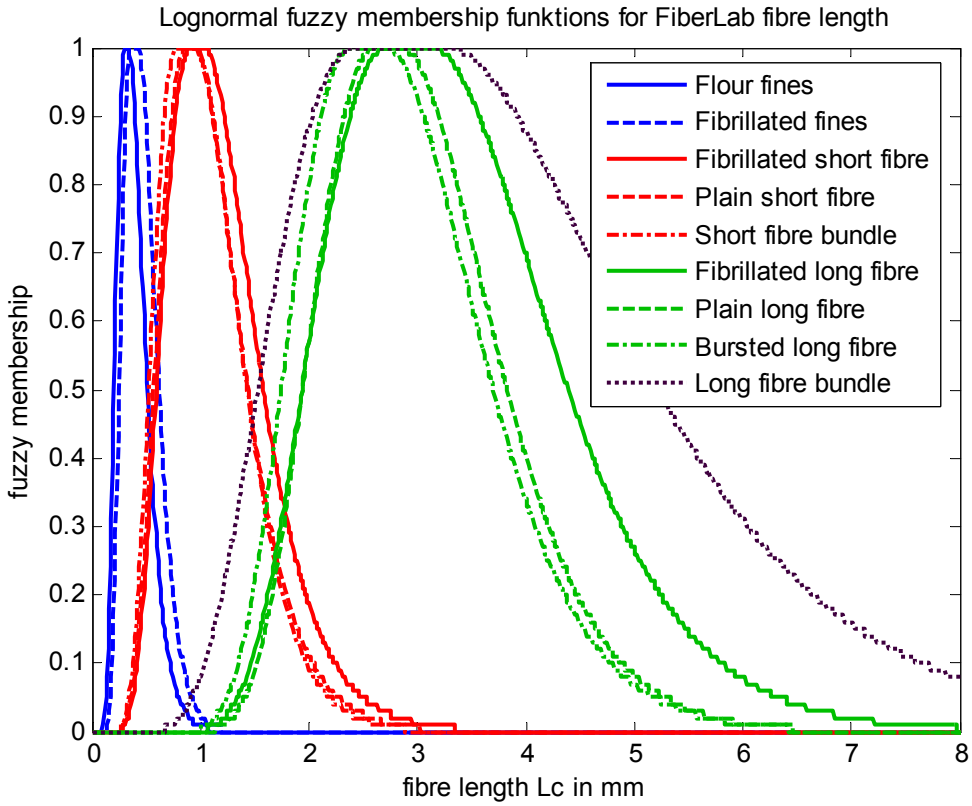


Figure 2. Fuzzy memberships for distinct fibre classes

Sigmoid membership functions are used to convert indistinct criteria, i.e. fibre width and wall thickness. Since FiberLab and VTT have different methods to measure fibre properties, the fuzzy interference system has to assess all available information to calculate the fibre class memberships.

2.2 Modelling the fibre mass

Fibre class-specific models evaluate the mass portion for the different fibre fractions. The mass models take the particular coarseness of the fibre classes into consideration. Mathematical models can be used, since the fibre properties are well known within the fibre classes. The mass of a fibre is estimated by multiplying the length by the cross-sectional area and the density of the cell wall.

The mass of a normal fibre is estimated as a pipe

$$m_{Fiber} = \pi \cdot s \cdot (D - s) \cdot l \cdot \rho \quad (4)$$

with wall thickness s, fiber width D, fiber length l and specific density ρ.

The mass model of fibre bundles is based on the assumption of two associated normal fibres.

$$m_{Bundle} = 2\pi \cdot s \cdot (D - s) \cdot l \cdot \rho \quad (5)$$

The FiberLab measures the width of entire object. The wall thickness reaches halve of the width. Both values have to be divided by two to get the appropriate figures. The VTT width is approx. 1.5-times higher than a normal fibre but the wall thickness is reliable.

The characteristics of the diverse measurement methods have to be compensated in the fibre mass models.

3. VALIDATION OF THE CLASSIFIER

Different fibre types and various process samples from TMP and recovered paper processing plants have been analyzed with the Bauer-McNett classifier. The BMC fractions again have been analyzed with the FiberLab for model validation.

Table 2 shows the correspondence between both methods for a TMP sample. R14 consists mainly of inflexible long fibres. R30 has a higher share of flexible long fibres. Flexible long fibres and inflexible short fibre dominate in the R50 fraction. The R100 fraction consists of flexible short fibres and fines.

Table 2. Mass distribution of Bauer-McNett fibre fractions in %

Class	Fines		Short fibre			Long fibre			
	Flour	Fibrils	Fibrilated	Plain	Bundle	Fibrilated	Plain	Extreme width	Bundle
R14	1	0.9	0.7	1.6	0	8.6	**57**	**22**	**7.8**
R30	0.1	0.7	2.8	8.4	0.7	**11.6**	53.5	14.8	7.1
R50	0.5	3.1	14.1	**33**	**1.5**	9	27.1	10	1.2
R100	**16.1**	**24.5**	**21.9**	29.8	0.8	1.7	3.2	1.5	0.1

The shortening and the fibrillation of fibres are the basic mechanism of beating. The fibre classifier allows a closer look into the refining process. Tab. 3 illustrates a trial in TMP plant with a steady rise of flour, fibrils and fibrillated short fibres fractions when the gap width drops.

Table 3. Mass distribution of refined TMP in %

Refiner	**Fines**		**Short fibre**			**Long fibre**			
Gap width	Flour	Fibrils	Fibrilated	Plain	Bundle	Fibrilated	Plain	Extreme width	Bundle
~ 0.4 mm	4.6	9.3	9.8	16	0.3	11.9	35.6	8.7	3.2
~ 0.2 mm	4.3	9.6	10	15.2	0.6	12.6	33.8	12	1.6
~ 0.1 mm	5	10	10.1	14.9	0.5	13.8	35.2	8.4	1.6
~ 0.0 mm	**5.1**	**11.1**	**11.1**	16	0.5	12.2	31.3	10.1	2.3

Because of the varying input quality, more trials are needed to study the process.

4. CONCLUSIONS

The fibre classifier has been developed for the FiberLab™ of Metso Automation. The optical system with two cameras distinguishes the FiberLab from other fibre analyzers. Its fibre width images can be stored and externally processed for visualisation and classification.

The fibre morphology can be assessed by lognormal and sigmoidal membership functions by a fuzzy interference system. Class-specific mass models estimate the fibre mass which is summed up to fibre fractions.

5. Acknowledgement

We would like to thank UPM-Kymmene, Schongau mill, particular Mr. Frank Zimmermann for classifying fibres visual and providing the measurements for validation. Special thanks to Mr. Ulrich Vitzthum, Metso Automation, for making a preliminary software version for the FiberLab™ available.

References

[1] Gooding, R., W.; Olson, J., A.: *Fractionation in a Bauer-McNett Classifier.* Journal of Pulp and Paper Science, Vol. 27, No. 12, 2001

[2] Gruber, E.: *Chemische Technologie des Zellstoffs und Papiers - Mechanische Faserstoff-Gewinnung*. Vorlesungsmanuskript, TU Darmstadt; 2003

[3] Hirvonen, J.: *Fiber wall thickness measurement by machine vision*. Vision Club of Finland, Helsinki, Machine Vision News, Vol. 4; 1999

[4] Hirvonen, J.: *Measurement of the quality of mechanical pulp by machine vision*. Vision Club of Finland, Helsinki, Machine Vision News, Vol. 5; 2000

3-D Simulation of handsheets made of different pulps

Rémi VINCENT[1, 2], Martine RUEFF[2] and Christian VOILLOT[2]

[1] TEMBEC (R&D) - [2] EFPG, LGP2

Ecole Française de Papeterie et des Industries Graphiques
461 rue de la Papeterie, BP 65, 38402 Saint Martin d'Hères CEDEX, France

Abstract

A new computational model of paper three-dimensional structure was developed in order to simulate the effect of fibre morphology upon paper properties. Over the past ten years, several automatic devices for determining geometrical parameters of the fibres in a pulp appeared on the market. The results, being displayed as discrete distributions, give the users a good insight on the heterogeneity of the pulp. Our model creates virtual pieces of handsheets using these distributions, in order to examine their effect on the 3-D structure of paper. The data produced by the simulation are used to determine the sample grammage and thickness and, thereafter, its apparent density and porosity. Other properties related to the geometry, such as the relative bonded area and the apparent specific surface area, can also be assessed. With a relatively short computing time, i.e. a few hours, it is possible to estimate the evolution of texture parameters in function of variations of one morphological parameter, e.g. the mean fibre length.

1. INTRODUCTION

Investigations about fibrous network structure aim at getting a better understanding of the process mechanisms or the effect of the raw material upon this structure. Various experimental methods have been used in order to characterise the structure. At the beginning of the eighties, the scanning electron microscopy became a conventional tool for examining the surface of samples or their cross sections. The images could be analysed, using concepts of stereometry, to get quantitative information. More recently, the X-ray microtomography was successfully applied for characterising 3-D structures [1-2]. However this method has some limits. Although paper is made of individual fibres bonded together, it is pictured as a compact structure. Thus, some information is missing. Another way to study the structure of a fibrous network is to use the computational modelling. Many models have already been developed; nevertheless, most of them are limited to two dimensions [3-6]. Nowadays, the statistical geometry of two-dimensional random fibre networks is well understood. It gives great information on paper formation and local grammage. Nevertheless, as the 2-D models do not consider the thickness of the network and the fibre conformation, they cannot give information about the bulk structure or the roughness of the material. Indeed, it is the three-dimensional pore structure which directly controls the porosity, density, permeability and some optical properties. It also plays an indirect role upon the mechanical properties through the relative bonded area. Thus, only a 3-D model can predict the properties of a paper sheet. Several models taking into account the three dimensions [7-9] already exist. These models provide qualitative or relative results, such as the effect of fibre flexibility on porosity. Unfortunately they are not precise enough. They either accept too important fibre interpenetration or do not consider the heterogeneity of the fibre morphology. Moreover, no comparison with real samples was done. A better quantitative modelling requires more elaborate microscopic models [9].

The aim of this article is to present a new 3-D model of paper structure applied to handsheets and the validation methods used. The model limits fibre interpenetration by a precise discretization of the elements. It enables us to introduce a large heterogeneity in the fibre morphology and gives quantitative results that can be compared to experimental values.

2. 3-D HANDSHEET GENERATION AND DATA TREATMENT

2. 1 3-D handsheet generation

Handsheet making is a succession of complex physical operations whereby a web is formed and then consolidated by means of pressing and drying. In order to simplify the problem, we only consider the final product. Thus, assuming that the fibres are deposed one by one, as if they settle on the underlying mat by filtration of a dilute suspension, the computational simulation consists of a loop with three steps:

- generation of a fibre,
- fibre deposition,
- and, finally, fibre conformation.

Each fibre is characterised by three parameters: its width, length and wall thickness. The fibre length, fibre width and fibre thickness distributions are modelled by normal distributions. For each parameter, the probability density is given by:

$$f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\exp\left[-\frac{1}{2}\left(\frac{x-\sigma}{\mu}\right)^2\right] \qquad (1)$$

where μ and σ are respectively the mean and the standard deviation of the parameter. The Box Muller method [10] was chosen to generate the random series. Several techniques were used to check the validity of the normal law: measurement of fibre length and fibre width distributions using a Morfi automatic fibre analyser [11, 12] and light microscopy, measurement of fibre width and wall thickness by means of light and SEM microscopy. The three distributions are considered to be independent of each other. This means for example that the larger fibres are not necessarily the thicker.

In this paper only Kraft pulps were modelled. Thus, we could assume that all the fibres are collapsed. The fibre cross-section is meshed with 6 points, repeated every 3 µm along the length. Therefore, the fibres are described by two thousand points per millimetre, as shown in Fig.1. Such a precision is necessary to avoid fibre interpenetration and simulate the actual conformation during the numerical generation of the network. The fibre curl is also considered: the fibre shape is modelled by a polynomial or a sine function.

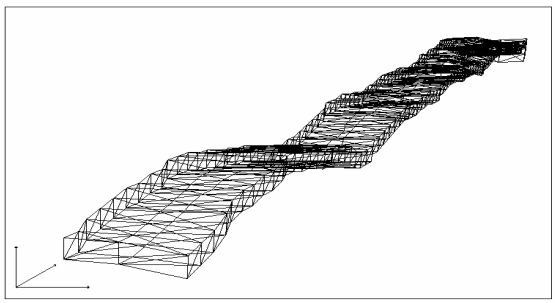


Figure 1: Meshing of a fibre.

After being generated, a fibre is placed horizontally on the network surface. According to the work by Kerekes and Schell [13], as the fibre suspension is sufficiently diluted for giving a Crowding factor around unity (case of a handsheet), there is no flocculation. Thus, the position of the centre of gravity of the fibres is calculated following a random uniform process. In order to include the influence of the 3-D structure on the fluid flow and fibre

arrangement during the forming process, a micrometric displacement of the fibre on the horizontal plane is created. The algorithm takes into account the small elements, i.e. shorter than 200 µm, namely the fines which fill the holes of the structure and give gradients in the z-direction [14, 15]. The last step is the conformation of the fibre on the network surface depending on its moment of inertia. As shown in Figs 2 and 3, the fibre interpenetration is negligible and the fibre meshing is sufficient for getting a good precision. The simulated samples can be virtually manipulated in real time, which enables us to make qualitative comparisons with SEM images.

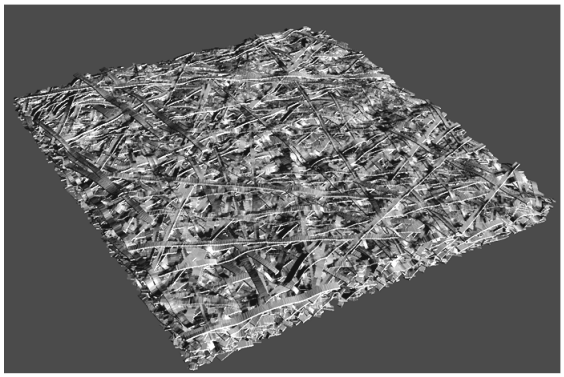


Figure 2: View of a simulated network with a surface area of 2.5 mm x 2.5 mm.

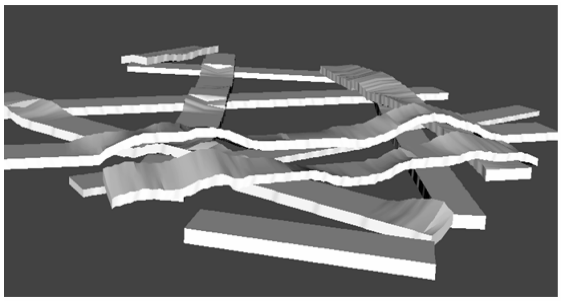


Figure 3: View of fibre conformation on a plane surface.

2.2 Data Treatment

All the model results are numerically calculated. The 3-D visualisation of the virtual sample is used only for qualitative validation. When the fibres are generated, their volumes are known. All the fibres being collapsed, their density is considered to be close to that of cellulose, i.e. 1.54 kg/L. Thus, it is easy to determine the value of the network grammage and, then, the apparent density, which is the ratio of the grammage by the thickness and the porosity. The relative bonded area (RBA) of the network is defined as the expected fraction of the surface of fibres that is bonded to other fibres [12]. The value of the contact area is assessed by means of an interpolation of the calculated points.

3. MODEL VALIDATION

3. 1 Repeatability

Many input parameters are partly governed by stochastic factors. Therefore, it is important to check the model repeatability. The model was tested with four different pulps corresponding to different wood species (eucalypt, pine, silver birch, northern softwood). Thirty networks were generated for each type of pulp and the worst error obtained was three percent.

3. 2 Sample Size

For reasons of computing time and memory capacity, the simulated sample size should be as small as possible. But a minimum size is necessary for neglecting the inevitable edge effects and obtaining stable values of the texture parameters. Indeed, these values may vary owing to

the heterogeneity of the structure. In order to examine the effect of this parameter, several networks with various sizes were generated for two pulps. As shown in Fig. 4, when the sample length is smaller than the average fibre length, the values of thickness are affected. For this reason, all the results presented hereafter concern square samples with a 6 mm^2 area.

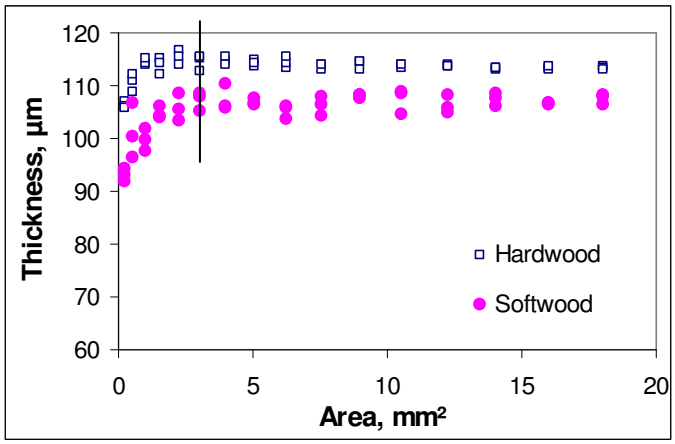


Figure 4: Variation of the thickness of the network as a function of the surface area.

3. 3 Comparison of the Simulated Handsheets with Real Handsheets

The final validation was made by comparing the properties of simulated samples to that of real handsheets. Four different pulps, representing a large panel of morphologies, were chosen. Some results are depicted in Fig. 5, where the very good agreement between the simulated and experimental values can be noticed. Such a good agreement, i.e. less than 5 % difference, was obtained for all pulps and texture properties.

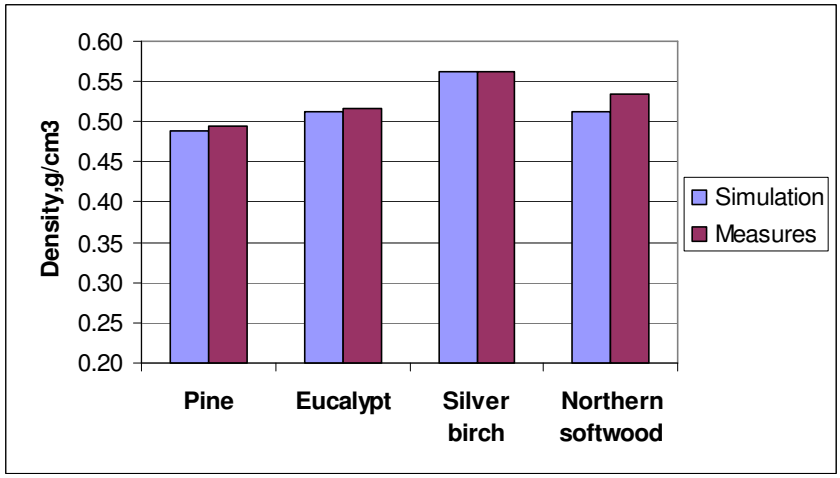


Figure 5: Comparison of the density obtained by simulation with handsheet values.

4. PREDICTION OF THE INFLUENCE OF FIBRE MORPHOLOGY UPON HANDSHEET PROPERTIES

In order to analyse the effect of fibre morphology on density and RBA, several monodisperse networks were generated. To this effect, two typical fibres were defined: a hardwood fibre (800 µm in length, 19 µm in width and 4.5 µm in thickness) and a softwood fibre (2500 µm in

length, 30 µm in width, 6.5 µm in thickness). For each type of simulation, only one dimension was varied. Then, in order to examine the influence of the fibre size distribution, polydisperse networks were generated, using "typical fibres" with one dimension varying according to a normal distribution. In this case, the mean value was kept constant and the standard deviation was varied.

5. 1 Influence of Fibre Length

Fig. 6 shows the variations of the density as a function of fibre length. A linear trend can be noticed for both types of fibres. Increasing the length slightly decreases the density, but for a large range of variation, i.e. 500 to 3000 µm, the density variation is only 0.05g/cm^3. This is in agreement with Lu and Carlsson [5] who stated that the fibre length has very little influence on the sheet apparent density. The variations of the RBA with fibre length are presented in Fig. 7. Conversely to Wang and Shaler [7], who found that the density and the contact area increase with the mean fibre length, we observe that the RBA also decreases when the fibre length increases. This tendency is due to the decrease of the density. In the range explored, the variations of the RBA with fibre length may be fitted by a power law function:

$$RBA = a + \frac{c}{L^b} \qquad (2)$$

in which a, b and c are three positive constants.

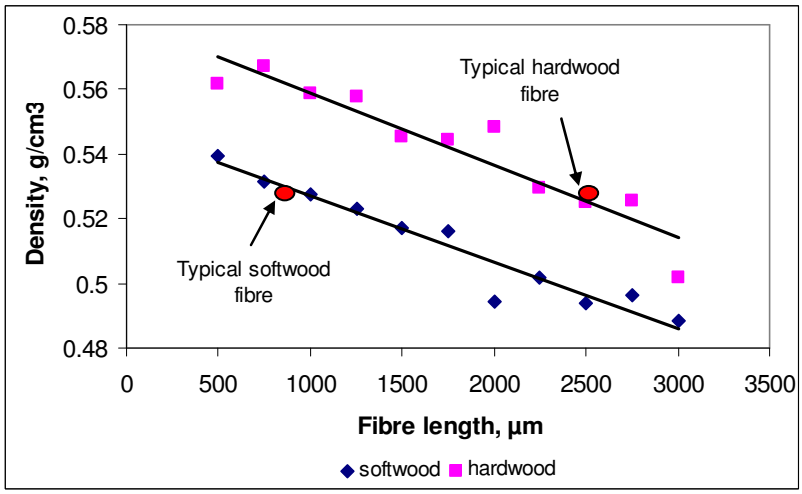


Figure 6: Variations of the network density with the fibre length.

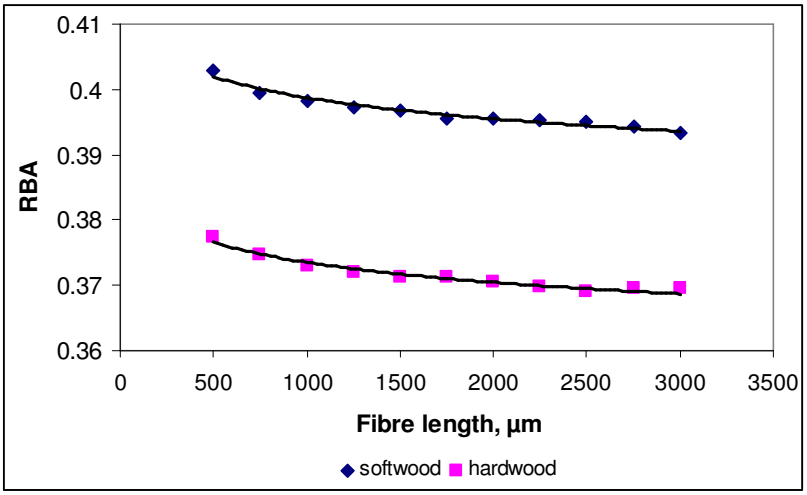


Figure 7: Variations of the relative bonded area with the fibre length.

5. 2 Influence of Fibre Width and Thickness

It was found that the density of the networks increases when the fibre width is increased or when the fibre thickness is decreased, but the variations are not linear. However, the most interesting parameter seems to be the ratio width by thickness, cf. Fig. 8. Its influence on the density of the network may be expressed as follows:

$$\rho = a + b \cdot \sqrt{w/t} \qquad (3)$$

where a and b are two positive constants, w the fibre width and t the fibre thickness.

It was also found that the fibre width only has a very small influence on the RBA value. This is in agreement with Lu and Carlsson's results [5]. The variations of the RBA as a function of the thickness of the fibres are depicted in Fig. 9. This figure shows that the fibre thickness has a strong influence on the RBA values, which vary linearly with the reciprocal of fibre thickness:

$$RBA = a + b/t \qquad (4)$$

where a and b are two positive constants.

We must notice that this trend differs with the Wang's results [7]. This author obtained the same type of function for his results but with a negative slope b, which means that the RBA increases with fibre thickness. We explain our results by the decrease of the fibre flexibility when fibre thickness increases, which induces a decrease of the contact area of the fibres.

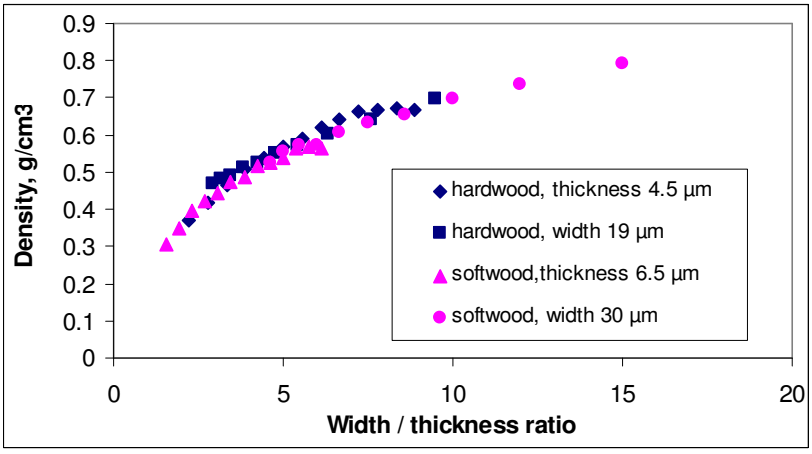


Figure 8: Variations of the network density as a function of the ratio fibre width / fibre thickness.

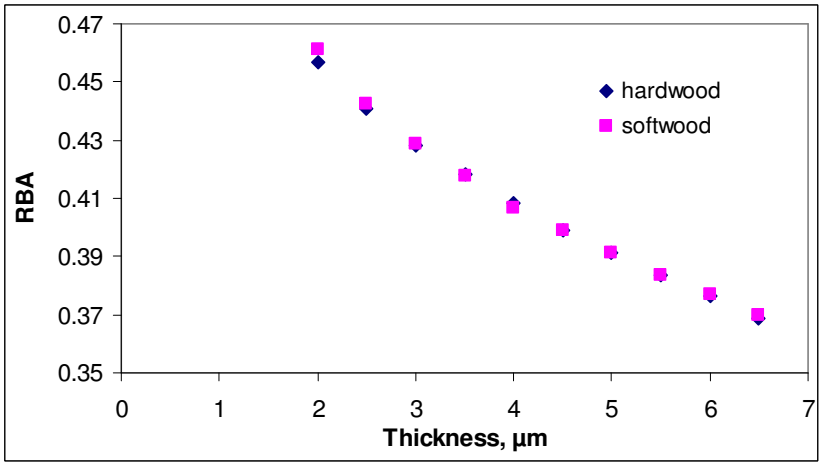


Figure 9: Influence of the fibre thickness upon the relative bonded area.

5. 3 Influence of the coefficient of variation of a distribution

Regarding the polydisperse networks, it was found that increasing the coefficient of variation of the length slightly decreases the density, similarly to the effect of the length for the monodisperse networks. Moreover, the relation is almost linear, with a negative slope. As shown in Fig. 10, the density also decreases when the coefficient of variation of the thickness is increased, and the trend is linear.

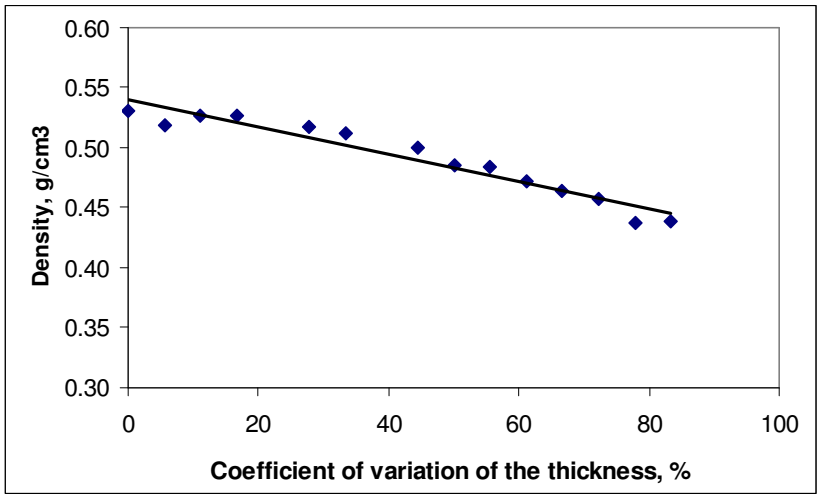


Figure 10: Influence of the coefficient of variation of fibre thickness upon network density.

Finally, it was found that the dispersion of the size distributions (length, width and thickness) does not affect the relative bonded area of the network.

6. CONCLUSION

Our model enables us to simulate 3-D samples of handsheets, with a much better accuracy than the models published in the literature. The simulation results do not only consist of numerical values or statistical data. We are able to visualise and manipulate the generated samples with our computer and compare them with SEM images.

This simulation has shown its capacity to predict the effect of geometrical characteristics of the fibres upon the structure of the network and some texture properties. It was validated by comparison with experimental values. In addition, simulation results were compared to the trends of existing models.

The very large number of discretization points on the fibres enables us to simulate particles as small as 1 µm. A future development will be to include the influence of the fines on the density and relative bonded area of paper. Other extensions will be made in order to analyse the effect of fibre morphology upon mechanical properties, e.g. the tensile strength. Moreover, an analysis of the pores in the structure could be used to determine the permeability of the simulated network.

ACKNOWLEDGEMENTS

This research was supported by the EFPG and the TEMBEC R&D Company.

REFERENCES

1. E. J. Samuelsen, P-J. Houen, O. W. Gregersen, T. Helle and C. Raven. Three-dimensional imaging of paper by use of synchrotron X-ray microtomography. 1999 TAPPI International Paper Physics Conference Proceedings, Kailua Kona, HI, pp. 307-312. 1999.

2. C. Antoine, P. Nygard, O. W. Gregersen, R. Holmstad, T. Weitkamp, C. Rau. 3D Images of Paper obtained by Phase-Contrast X-ray Microtomography: Image Quality and Binarisation. 2002 Nuclear Instrument and methods in Physics Research A 490, pp. 392-402.

3. O. Kallmes, H. Corte. The Statistical Geometry of an Ideal Two Dimensional Fiber Network. Tappi Journal vol. 43 (9), pp. 737-752. 1960.

4. O. Kallmes, H. Corte, G. Bernier. The Structure of Paper, Part II. The Statistical Geometry of an Multiplanar Fiber Network. Tappi Journal vol. 44 (7), pp. 519-528. 1961.

5. W. Lu and L. A. Carlsson. Micro Model of Paper. Part 2. Statistical Analysis of Paper Structure. Tappi Journal vol. 79 (1), pp. 203-210. 1996.

6. M. Deng, C. T. J. Dodson. Paper an Engineered Stochastic Structure. Tappi Press. 1994.

7. H. Wang, S. M. Shaler. Computer-Simulated Three-Dimensional Microstructure of Wood Fibre Composite Materials. Journal of Pulp and Paper Science vol. 24 (10), pp. 314-319. 1998.

8. K. J. Niskanen, M. J. Alava. Planar Random Networks with Flexible Fibers. Physical Review Letters vol. 73 (25), pp. 3475-3478. 1994.

9. N. Porvatas, M. Haataja, J. Asikainen, S. Majanien, M. Alava, T. Ala-Nissila. Fiber Deposition Models in two and three Spatial Dimension. Colloids and Surface A, 165 pp. 209-229. 2000.

10. Numerical Recipes in C++: the Art of Scientific Computing. Chapter 7. pp. 293-294. Cambridge University Press 2nd edition. 2002.

11. Instruction manual of the MorFi apparatus. Techpap – Saint-Martin d'Hères – France.

12. G. Eymin Petot Tourtollet, G. Gillet, P. Lafaverges, J. Kerneis, J. Sabatier, D. Moineau, R. Pascal, C. Voillot, R Passas. MorFi : une avancée décisive dans l'analyse des pâtes. Revue ATIP vol. 54 n° 3-4, pp. 106-114. 2000.

13. R. J. Kerekes, C. J. Schell. Characterization of Fibre Flocculation Regimes by a Crowding Factor. Journal of Pulp and Paper Science vol. 18 (1) J33-38. 1992.

14. Z. J. Majewski. Effect of forming processes on sheet structure. In Formation and Structure of Paper. Trans. 2nd Fundamental Research Symposium. Oxford, pp. 749-766. 1961.

15. M. Hasuike, T. Kawasaki, K. Murakami. Evaluation Method of 3-D Geometric Structure of Paper Sheet. Journal of Pulp and Paper Science vol. 18 (3) J114-120. 1992.

16. J. He, W. J. Batchelor, R. E. Johnston. An Analytical Model for Fibre-fibre Contacts in Paper and Expressions for Relative Bonded Area (RBA). 2003 International Paper Physics Conference, Paptac, Montreal, Canada, pp. 77-83. 2003.

An Approach for Modeling and Optimizing of Industrial Fine Screening Processes

Jukka Valkama, Klaus Villforth and Samuel Schabel
Darmstadt University of Technology
Chair of Paper Technology and Mechanical Process Engineering
Alexanderstraße 8, 64283 Darmstadt, Germany

Abstract

This paper presents an approach for modelling industrial screening systems by using a dynamic single matrix model. The aim was to compare different screening systems and find out the right strategies for single screening systems.

The validation of the model was done by using 12 quality parameters and the results were compared with the measured process data. The simulation with the model gives sufficient results when using mass balances and can be used for optimization of fine screening processes.

1. INTRODUCTION

Several studies have been made to model the behavior of suspension in pressure screening process /1,2,3/. The plug flow and mixed flow models are well predicting the correlation between the single components in relation to the mass reject rate /4,5/. Advanced models have also been introduced for fibre passage. Most of the studies have been made for modeling of fractionation of fiber material in pressure screens, but also are suitable for general modeling in multi-component recovered paper process.

The approach for this study was to generate a model for optimization of sticky separation in fine screening in the recovered paper field. The target was the validation of a quality propagation model. The model was tested with research data from several packaging paper mills by using a plug flow model to describe the behavior of fine screens.

The fine screening result is strongly dependent on the process parameters like the total opening area of the slots/holes, rotor type or rotor velocity. Feed quality parameter also affect on the result. As a control parameter the controllers can change the volumetric overflow ratio and feed amount but also affect the feed quality i.e. through dilution. A screening system usually has three to five different screening units and several installation possibilities. The input quality to the system is a changing multi-compontent sum of different type of particles.

Changes in machine parameters are normally made through the investment in a new screening basket or rotor. Feed quality is controlled through defined consistency. The control parameters are set in the mill according to fine screening strategy and not continuously controlled. Therefore the free capacity of the screening units cannot be

used. No dynamic control exists and the existing screening strategies are not optimal. Every mill has it own strategy on optimal quality for the paper machine. Every optimization criteria is needed to configure for each mill separately.

Within this study compared different screening units and defined the relevant constants according to the theoretical model on screening behavior.

2. MATHEMATICAL EQUATIONS FOR BALANCING A FINE SCREENING PROCESS

The mass balance in a screening unit for a component ***i*** can be defined as

$$\dot{m}_f c_{i,f} = \dot{m}_o c_{i,o} + \dot{m}_t c_{i,t} \qquad (1)$$

where flow is defined with a mass flow rate, $\dot{m}$ and the concentrations *c*. The subscripts ***f***, ***o*** and ***t*** indicate feed flow, overflow and throughflow. Some authors are using inlet, reject and accept, respectively. As the overflow is often still valuable raw material the term reject is not suitable for a single screening unit. In practice the processes are always controlled by volumetric flow rates $\dot{V}$ and consistencies ***C***.

The thickening factor $T = \frac{C_o}{C_f}$ can be calculated for every sceening unit as a quotient of the consistency of the overflow and feed. The passage ratio $P = \frac{C_t}{C_o}$ is defining the consistency ratio between throughflow and overflow, when the pulp is assumed to be perfectly mixed inside the screen and the upstream consistency equals the overflow consistency i.e. the mixed flow conditions.

Because the volumetric flow ratio is defined as,

$$R_V = \frac{\dot{V}_o}{\dot{V}_f} = 1 - \frac{\dot{V}_t}{\dot{V}_f} \qquad (2)$$

and mass overflow ratio or mass reject ratio as,

$$R_m = \frac{\dot{m}_o}{\dot{m}_f} = \frac{\dot{V}_o C_o}{\dot{V}_f C_f} = 1 - \frac{\dot{m}_t}{\dot{m}_f} \qquad (3)$$

the equation for thickening factor may be rewritten as $T = \frac{R_m}{R_V}$.

Two main removal equations can be defined for the singular mass components like macro stickies. The removal efficiency $\boldsymbol{E_R}$ for the component *i* is defining as

$$E_R = \frac{\dot{m}_o c_{i,o}}{\dot{m}_f c_{i,f}} = R_m \frac{c_{i,o}}{c_{i,f}} \tag{4}$$

Cleanliness efficiency $\mathbf{E_C}$ is describing the difference between the concentration of the component *i* in outgoing streams in correlation to feed concentration.

$$E_C = \frac{c_{i,o}}{c_{i,f}} = 1 - \frac{c_{i,t}}{c_{i,f}} \tag{5}$$

2.1 Plug Flow Model

The plug flow model was introduced year 1955 /4/. In the plug flow model is assumed that the no axial mixing exists in a screening drum but a perfect radial mixing. The capability of a pressure screen according to plug flow model can be decribed through reject thickening factor $\boldsymbol{T}$

$$T = \frac{c_{i,o}}{c_{i,f}} = R_V^{(P-1)} \tag{6}$$

and component removal efficiency $\boldsymbol{E_R}$

$$E_R = R_m^{\beta} \tag{7}$$

where $\boldsymbol{\beta}$ is a unit specific constant for every component defined as a passage ratio difference between the component *i* and the homogenous fibres.

$$\beta = \frac{P_i}{P_{fibre}} \tag{8}$$

In reality it is hard to define homogenous fibres and the $\boldsymbol{\beta}$ can be calculated for example for every fibre fraction separately.

2.2 Mixed Flow Model

The mixed flow assumes a perfect radial and axial mixing inside of the screening unit. Mixed flow model was introduced by Nelson 1981 /5/. For the mixed flow model the component (reject) thickening factor $\boldsymbol{T}$ is given as

$$T = \frac{1}{P - R_V P + R_V} \tag{9}$$

and component removal efficiency $\boldsymbol{E_R}$

$$E_R = \frac{R_m}{1 - Q + QR_m} \tag{10}$$

Q is a device specific screening quotient and equals Q=1-β.

3. AN APPROACH TO MODEL A SCREENING PROCESS

The process parameters are defined according to the theoretical models of the screening process. We shall present the unit specified process parameters by a vector *p*. For the optimizisation we let the control parameter i.e. the mass reject rate to be a variable. When optimizating an existing process with constant input quality we shall denote the input quality to be presented as a vector *q*. The quality matrix in a system is then presented in the time k and the location l.

A three dimensional matrix (*fig.1*) is used providing the quality data to the process model.

- 1st dimension represents the time (k,k-1,...,k-n)
- 2nd dimension represents the test point/place (l_1,…l_r)
- 3rd dimension distinguish the quality parameters (q_1,…q_s)

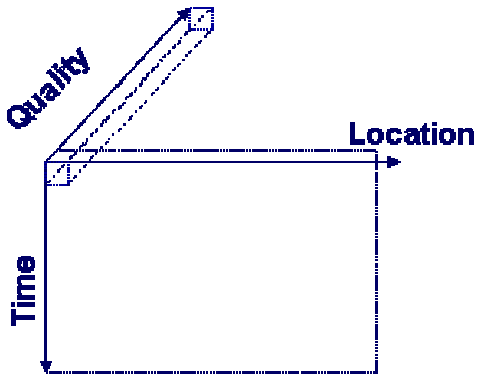


Figure 1. Schematic figure from quality matrix.

The present quality is shown with the time *k*. The quality data matrix presents the actual state of the component in time and place and a four dimensional array (*fig.2*) models the process.

- 1st dimension represents the time (k,k-1,…,k-n)
- 2nd dimension represents the process steps, source (l_{s1},…l_{sr})

- 3rd dimension represents the process steps, sink ($l_{d1},\ldots l_{dr}$)
- 4th dimension distinguish the quality index ($q_1,\ldots q_s$)
- 5th dimension contains the optional parameter set of a transfer function ($t_1,\ldots,t_t$)

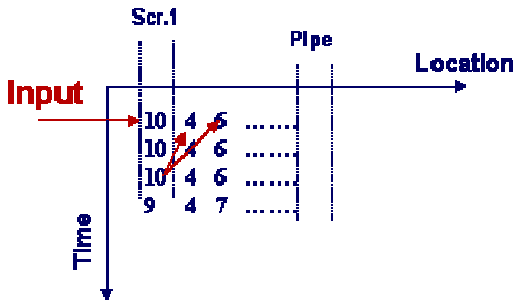


Figure 2. Schematic figure from quality data transfer in two-dimensional plane matrix.

Local transfer function-vectors *t* are supporting the quality propagation.

The location *l* shall now be presented as a source denoted with a letter *n* and the sink with the letter *m*. Then for every quality component *q* in time axis *k*=0 (sink) is the value a sum of all tranfered values:

$$Y_{(m,q)} = X_{(k=0,m,q)} = \sum_{k}^{K} \sum_{n}^{N} p_{(k,m,n,q)} Q_{(k,n,q)} \qquad (11)$$

Y represents the actual quality at the distinguished locations. The transferred quality will be defined as a function of the source quality and process parameter. The multi-dimensional array propagates quality in time and place.

5. VALIDATION OF THE MODEL

The processes parameters (β-values) were defined for 12 quality components. The used components were volumetric flow, mass flow, specific total macro sticky area, number of macro stickies, total fibre mass, fibre fines, filler content and 5 fractions from the Bauer-McNett classifier (R14, R30, R50, R100 and 150µm // screen).

The model was tested by constant known input quality with R_m-values varying between 0.1-0.9. Quality output plane-matrix (with real R_m values) was compared with laboratory measurement data from the mills. The model delivered equal results when using mass flows. The correlation existed for all the quality components, when the measurement error was included.

6. OPTIMIZATION OF A STICKY REMOVAL IN AN INDUSTRIAL SEPARATION PROCESS

Let us set a constant input (measured actual input) to the model and the model gives unit parameters, in plug flow case the β-values for every component considered. For every control parameter or model variable combination $R_{m,i}$ there is a two-dimensional quality matrix-plane as a result.

The optimization of separation processes, especially fine screening, is a question of optimization criteria. The fine screening system has always a total accept, which can be considered as a raw material for the paper machine, but also a total reject which needs to be treated.

The accept flow consists mainly of five components which need to be considered.

- Fibres as a function of quality (length, fibrillation degree, flexibility),
- Fibre fines,
- Fillers,
- Stickies (colloidal, micro, macro), and
- Trash.

The total reject can be evaluated by two components: fibre material and total reject.

A two-dimensional value matrix can be created to evaluate the maximal profitability. Let us use as a simple example following parameters:

Fibres and fines	Accept:	350€/t (as an average selling price for test liner)
	Reject:	-86€/t (as a lost of raw material)
Fillers	Accept:	300€/t (poorer raw material)
Trash	Accept:	-300€/t (for quality losses)

The proposal for macro stickies is a cost function with concentration and mass flows.

$$f(c_{SS,A}, \dot{m}_{SS,A}, \dot{m}_{tot,A}) = \frac{\dot{m}_{SS,A}^2}{\dot{m}_{tot,A}^2} \cdot G_{Cons} + \dot{m}_{SS,A} \cdot G_{Mass} = c_{SS,A}^2 \cdot G_{Cons} + \dot{m}_{SS,A} \cdot G_{Mass} \qquad (12)$$

where G_{Cons} and G_{Mass} are the cost constants for concentration and amount of stickies. The total profit is taken to be the optimization criteria. Figure 3 presents the results from an industrial system as an example.

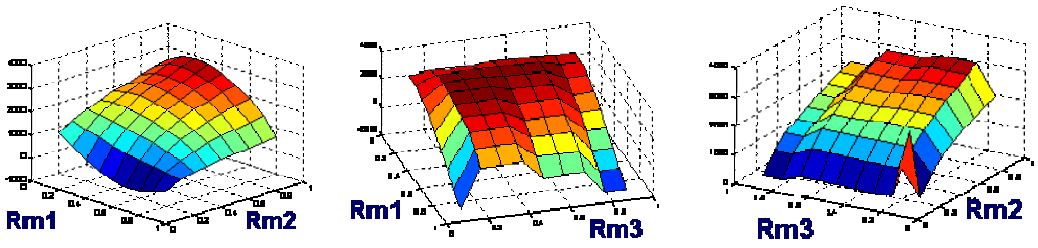


Figure 3. Profit as a function of mass reject rate in 3 stage fine screening system.

The optimum of these results is to set the R_{m1} between 0.3-0.4 and the R_{m2} as high as possible whereas the R_{m3} should remove the final reject in the level of 0.3-0.7. The most critical stage as expected is the first stage. No other limits were set for the optimization in this example and no thickening was considered.

7. CONCLUSIONS

The model introduced is suitable for optimization of industrial separation processes. The model is offering a powerful tool for dynamic simulation, when there are dynamic input data available. The results of the optimization strongly depend on the criteria given for the optimization tool. There are yet no published parameters available for more accurate optimization and therefore one must pay attention on the results given by the simulation. Especially the value function of stickies need to be tuned.

8. Acknowledgement

The research project (AiF 13990) is founded by the Arbeitsgemeinschaft industrieller Forschungsvereinigungen "Otto von Guericke" e.V. (AiF) with budgets of the German Ministry of Economics and Technology (BMWi). The authors would like to thank AiF for financing. Also special appreciation to the Finnish Paper Engineering Association for the support of my Ph.D. work.

References

[1] Niinimäki, J., “On the fundamentals of pressure screening”, Ph.D. thesis, University of Oulu, Department of Process Engineering, Oulu, Finland, 1998.

[2] Ämmälä, A., “Fractionation of thermomechanical pulp in pressure screening”, Ph.D. thesis, University of Oulu, Department of Process and Environmental Engineering, Oulu, Finland, 2001.

[3] Schabel, S., “Improvements in Screening System Configuration Using Simulation and Mill Verification”, Res. Forum Recycling, (2001), pp.79-82.

[4] Kubat, J. & Steenberg, B., “Screening at low particle concentrations”, Svensk Papperstidning, 58(9): pp.319-324.

[5] Nelson, G.L., “The screening quotient: a better index for screening performance”, TAPPI Journal, 64(5), p.133.

Validation in mill practice: Accuracy of water/COD, heat and stickies simulations in P&P applications

Dr. Johannes Kappen, Dr. Wolfram Dietz
Papiertechnische Stiftung PTS
Hess-Strasse 134, 80797 Munich, Germany

Abstract

Simulation models have been built for over two decades within the pulp and paper industry. Still today most of these models are being built within the research environment of universities and research centres. In many cases their results are not being used in mill practice. The industry use of computational simulation methods is lacking far behind. The question is why? One could put the argument aside stating that the industry is a conservative one and that it simply takes time or that the profit margin is too low to allow innovative techniques to be applied. But on the other hand mills are highly sophisticated in technical terms and many technologies promising a higher margin have been quickly adopted.

As it seems, industry has not yet found models valid enough to be applied onto their processes and procedures. Valid in this context is meant to be taken in the broadest sense of an apparent value.

In the first part of this presentation simple examples are given of projects were mill models have actually been used in the design phase. The examples describe projects with the following objectives: Water balance optimisation, COD prognosis, heat optimisation and stickies reduction. These examples represent typical cases and concern some basic necessities paper mills have. The approaches on modelling, calibration, validation and on the degree of accuracy reached are described. Special attention is given the question of finding the appropriate degree of complexity in setting up the model.

In a second part of the presentation an attempt is made to generalise the knowledge reached within the projects. This is done by questioning the success of a mill related simulation project: What is important to achieve a successfully validated model? What makes a model a valid tool to the industry? Is it only the accuracy?

Published by

Series title, number and report code of publication

VTT Symposium 238
VTT–SYMP–238

Author(s)
Kappen, Johannes, Manninen, Jussi & Ritala, Risto (eds.)

Title

COST ACTION E36

Modelling and simulation in pulp and paper industry

Proceedings of Model Validation Workshop

Abstract

This workshop by COST Action E36 focuses on model validation which is considered as one of the current central topics in practical modeling and simulation of industrial processes, such as those in pulp and paper industry. We have invited contributions from COST E36 members and all other researchers interested in model validation. The topics of the workshop cover applications, practical aspects and theoretical considerations of

- model validation as a part of model development and parameter identification

- maintenance of models

- dynamic model validation

- sensitivity analysis

- uncertainty analysis.

Keywords
COST Action E36, parameter identification, maintenance of models, dynamic model validation, sensitivity analysis, uncertainty analysis, pulp and paper industry, wet-end chemistry, runnability, emissions reduction

Activity unit
VTT Processes, Lämpömiehenkuja 3 A, P.O.Box 1604, FI–02044 VTT, Finland

ISBN			Project number
951–38–6300–X (URL: http://www.vtt.fi/inf/pdf/)			
Date	Language	Pages	Price
October 2005	English		-
Name of project		Commissioned by	
Series title and ISSN VTT Symposium 1455–0873 (URL: http://www.vtt.fi/inf/pdf/)		Published by VTT Information Service P.O.Box 2000, FI–02044 VTT, Finland Phone internat. +358 20 722 4404 Fax +358 20 722 4374	

This workshop by COST Action E36 focuses on model validation which is considered as one of the current central topics in practical modeling and simulation of industrial processes, such as those in pulp and paper industry. We have invited contributions from COST E36 members and all other researchers interested in model validation. The topics of the workshop cover applications, practical aspects and theoretical considerations of –
model validation as a part of model development and parameter identification
- maintenance of models
- dynamic model validation
- sensitivity analysis
- uncertainty analysis.

VTT SYMPOSIUM 238 COST ACTION E36. Modelling and simulation in pulp and paper industry

VTT TIETOPALVELU
PL 2000
02044 VTT
Puh. 020 722 4404
Faksi 020 722 4374

VTT INFORMATIONSTJÄNST
PB 2000
02044 VTT
Tel. 020 722 4404
Fax 020 722 4374

VTT INFORMATION SERVICE
P.O.Box 2000
FI–02044 VTT, Finland
Phone internat. + 358 20 722 4404
Fax + 358 20 722 4374

ISBN 951–38–6300–X (URL: http://www.vtt.fi/inf/pdf/)
ISSN 1455–0873 (URL: http://www.vtt.fi/inf/pdf/)