(article starts on next page)

## Mechanical pulping

Fredrik Bengtsson*, Anders Karlström and Torsten Wik

# On the modeling of pulp properties in CTMP processes

**Abstract:** The goal of this paper is to model the pulp properties fiber length, shives width and freeness. This will be done utilizing specific energy, flat zone inlet consistency and the internal variables, consistencies and fiber residence times estimated from refining zone soft sensors. The models are designed using more than 3600 hours of data from a RGP82CD refiner. The pulp properties are sampled using a measurement device positioned after the latency chest. Such measurements are noisy and irregularly sampled which opens for a number of challenges to overcome in modeling procedures. In this paper it is shown that the models for shives width and fiber length are capable of predicting most of the major dynamics. However, for freeness no reliable linear models can be derived. When estimating fiber length, the specific energy together with flat zone inlet consistency, fiber residence times and the consistency in the conical zone were the dominant inputs. For shives width it was found that a similar set of inputs resulted in the best models, except that the consistencies during normal process conditions did not significantly influence shives width. Furthermore, fiber residence times were shown to have considerably more pronounced impact on fiber length compared with shives width estimates.

**Keywords:** CTMP; fiber length; linear modelling; pulp properties; shives.

## Introduction

It is well-known that pulp properties are often difficult to measure reliably in TMP- and CTMP-processes. This, to-gether with the fact that most commercial automatic measurement devices have several minutes long measurement times (Hirn and Bauer 2006) and irregular sampling rate has made implementation of automatic control of pulp properties has been a challenge for many years.

As the measurement quality of these devices is questioned from a control engineering perspective, considerable benefits could be achieved if models are implemented to predict pulp properties based on process conditions sampled before the latency chest. Such models are useful for a number of reasons as they allow predictions which guide operators to find proper operating conditions. For controller design on pulp and paper refiners an understanding of the process is necessary (Blanco et al. 2009). The approach in mechanical pulping processes, has been that external variables, such as specific energy (i. e. the ratio between motor load and production), dilution water added to the refiners, plate gaps (disc clearance) etc., should be used for control and modelling of pulp and handsheet properties (Strand 1996, Härkönen et al. 2000, Sabourin et al. 2001, Härkönen et al. 2003, Strand and Grace 2014, Nelsson 2016). However, when using external variables as predictors, process non-linearities tend to negatively affect the result. To cope with that, soft sensors describing physical phenomena in the refining zone have been developed during the last decade (Karlström and Eriksson 2014a, 2014b, 2014c, 2014d). The soft sensor's outputs can be seen as estimates of internal variables (such as fiber residence time, consistency profile, forces on bars, distributed defibration, thermodynamic work etc.) which are difficult to measure directly in the process. Typically, such soft sensors are non-linear but still have become important for advanced process optimization (Karlström and Hill 2017a, 2017b, 2017c, Bengtsson et al. 2020).

In this paper, a mix of external variables (such as specific energy and flat zone (FZ) inlet consistency) together with internal variables (consistencies and fiber residence times) will be used to design models for future use in control applications. While previous research has used specific energy to model pulp properties (Schwartz et al. 1996, Luukkonen et al. 2012), research on modeling the influence of internal variables is limited.

*Corresponding author: Fredrik Bengtsson, Department of Electrical Engineering, Chalmers University of Technology, SE 412 96 Göteborg, Sweden, e-mail: fredben@chalmers.se, ORCID: https://orcid.org/0000-0002-6670-7493
Anders Karlström, Torsten Wik, Department of Electrical Engineering, Chalmers University of Technology, SE 412 96 Göteborg, Sweden, e-mails: anderska@chalmers.se, torsten.wik@chalmers.se
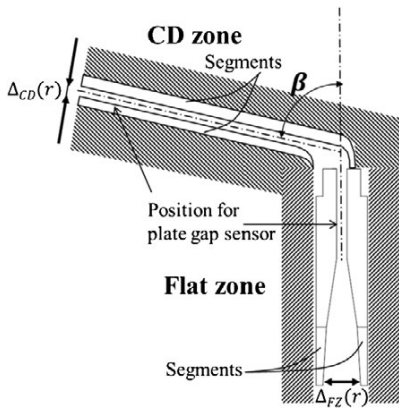
**Figure 1:** A schematic drawing of a CD refiner. The vertical flat zone (FZ) is directly linked to the conical zone (CD) via an expanding point. Plate gaps are represented by $\Delta_{CD}$ and $\Delta_{FZ}$.



**Figure 2:** A figure showing the output (shives width) demonstrating the slow and time varying sampling rates.

Among the numerous different pulp properties available, this paper will focus on predicting freeness, fiber length and shives width.

The main focus will be to determine which inputs are best suited for modeling the chosen pulp properties and how extensive the times series should be to derive reliable models of the refining process.

Two strategies for linear modelling are explored, one is to use the entire data set to derive a static linear model and the other is to use a recursive algorithm to derive dynamic model which is continuously updated over time (Ljung 1999).

## Materials and methods

The process data available is obtained from a full-scale production line with a yearly production capacity of about 250 000 tons CTMP-pulp. The main line refiner is a RGP82CD refiner with three small 6 MW refiners (RLP54) in parallel. The CD82-refiner consists of two serially linked refining zones called the flat zone (FZ) and the conical zone (CD) as shown in Figure 1. In each zone, sensor arrays with eight temperature sensors have been mounted to measure the entire temperature profiles. The temperature measurements can be seen as internal variables and from these, variables such as the consistencies and fiber residence times of the different zones can be estimated (Karlström and Eriksson 2014a, 2014b, 2014c, 2014d).

In this paper six model inputs are considered:
1. The specific energy.
2. The inlet consistency of the flat zone, which mainly depends on the percentage of sawmill chips fed to the refiner (high inlet consistency implies low amount of sawmill chips).
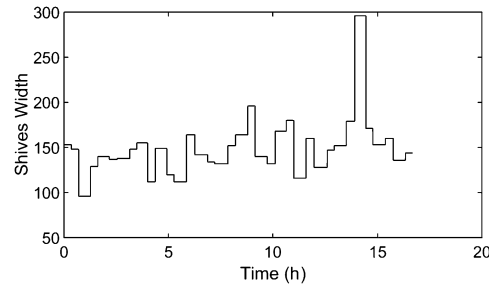3. The fiber residence times of each zone.
4. The consistency, when leaving each refining zone.

The pulp produced in the refiner is analyzed in a measurement device positioned after the latency chests. The device periodically samples and records the values of the pulp properties (freeness, fiber length and shives width). Shives width is measured as number of shives in a sample with a width greater than 150 µm. The sampling rate is non-equidistant and considerably slower than the dynamics in the refiners. Therefore, it is normally not used for control purposes.

The process variables (inputs) were taken during two different periods, each around 1800 hours long (sampling rate: 6 seconds) to get enough data for the training and validation procedures. The process was running in normal conditions during the time the measurements were taken, i. e. no experiments were performed to deliberately excite the inputs during the period.

Two controllers were used during the complete test period. Firstly, the consistencies were controlled by manipulation of the dilution water feed rates in the flat zone and the conical zone, according to specifications set by the operators. Secondly the maximum temperature in the FZ was controlled by manipulating the production. These parameters were under automatic control for around 80 % of the test period, while manually adjusted for the remaining 20 % of the time.

## Data preprocessing

The outputs are measured using a measurement device (PulpEye), with a sampling rate that is time varying and considerably lower than the sampling rate of the inputs (see Figure 2). To resolve this we take the mean value of the inputs during the period of each output sample, which gives one set of input values for each output value. This seems a reasonable step to take as there is a latency chest between the process and the output measurements, and

the mixing in the latency chest can be seen to have an averaging effect. Furthermore the latency chest will impose an unknown and random delay (due to variations in the volume stored in the latency chest). Further preprocessing includes de-trending the data by removing the means as well as removing measurements 3 standard deviations from the mean as outliers (this removed around 3 % of the data, mainly extreme outliers believed to be due to measurement equipment errors).

We explored the possibility of adding delays to the input measurements to compensate for the delay from the latency chest. However this was found to yield no improvements, probably due to the fact that the slow sampling rate of the outputs meant that the delay of the latency chest could easily be less then the output sampling time.

## Methods for pulp property modeling and evaluation

### Static linear regression model

The static linear regression model in this article is

$$\hat{y} = \theta x,$$

where $\hat{y} \in R^{1 \times 1}$ is the prediction of the output, $x \in R^{p \times 1}$ is a vector of the predictors (inputs), and $\theta \in R^{1 \times p}$ is a vector of what is called the parameters. $p$ is the number of predictors. The goal is to derive values of $\theta$, such that from measurements of $x$, one gets as accurate estimate of $y$ as possible.

The output and the predictors are grouped in matrices such that a data set of size $N$ can be expressed as

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \qquad X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{bmatrix}$$

where $y_1$ corresponds to the first measurement of the output, and $x_1$ are the measurements of the predictors at this time. Thereafter the parameters $\theta$ which best describe the relationship between $Y$ and $X$ are derived. This relationship can be assessed by examining the mean square error of the difference between predicted and measured outputs, i.e

$$\sum_{i=1}^{N} (y_i - \theta x_i)^2 = (Y^T - \theta X^T)(Y - X\theta^T),$$

which is minimized by

$$\theta = (X^T X)^{-1} X^T Y.$$

Static linear models have the advantage that they are simple to derive and interpret. However, it is important to note that they are not able to model the system dynamics. Moreover the models will be of poor quality if the effect of the predictors on the outputs is highly non-linear.

### Dynamic models

An alternative to the static models are dynamic or adaptive models which change over time. These models have the advantage that they can change to adjust to changing process conditions. There are a few different variants of adaptive models and here we will examine a sliding window adaptive model (Young 2011) in which data is discarded after a certain time, such that the models are based only on the most recent data.

When implementing sliding window recursive models there are two key properties. Firstly, the window length which is a measure of the amount of data used to determine the models (so a window length of 1000 means that the 1000 most recent data points are used to derive models). Secondly, the prediction horizon, which is a measure of how long forward in time predictions are made. So a prediction horizon of 50 would mean that we are making predictions 50 samples (300 seconds) ahead.

A dynamic alternative to the static linear models is the ARX model in which the output is predicted from a combination of past inputs and outputs, i. e.,

$$A(q)y(k) = B_1(q)u_1(k) + B_2(q)u_2(k)$$
$$+ ... + B_p(q)u_p(k) + e(k) \qquad (1)$$

where $y(k)$ is the output at time $k$, $e(k)$ describes the noise, $A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + ... + a_n q^{-n}$, $B_x = b_1 + b_2 q^{-1} + b_{n+1} q^{-n}$ and $q$ is the shift operator (i. e. $q^{-1}y(k) = y(k-1)$). $n$ is a tuning parameter used to determine the size of the model. By setting $q = 1$ in the above equation the DC-gain can be calculated, which is the ARX models gain at low frequencies, and is comparable with the parameter vector from the static linear model.

The ARX model can, unlike the static models, also reflect the dynamic behavior of the system. The ARX models can be derived in the same way as static linear regression models, by simply expanding the predictor matrix $X$ to also contain past input and output measurements.

Conventional implementation of ARX models in our case is problematic as the output sampling rate is irregular and slow. We cannot resolve this by simply averaging the inputs over the period of each output sample, as was done when implementing linear regression, as this would not only mean that the sampling time would be time-varying

but also result in a sampling rate too low to be able to encompass the dynamics of the refiner. Instead to solve this we use a method described in Sanchis and Albertos (2002), where the model is designed recursively and used to estimate the output at each input sampling interval.

There are other more complex dynamic models such as the ARMAX and Box-Jenkins models (Ljung 1999), which include noise dynamics. However these utilize output measurements to estimate the noise terms and thus are not appropriate in our case with highly sparse output measurements.

### Evaluation of models

When evaluating individual models the coefficient of determination,

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2}, \quad (2)$$

will be used. Here $\sum(y_i - \hat{y})^2$, represents the sum of the squared residuals from the regression and $\sum(y_i - \bar{y})^2$ the sum of the squared differences from the mean of the dependent variable while $n$ is the number of observations (Draper and Smith 1998). If the coefficient of determination is 1 a perfect fit is obtained, while an $R^2$ below zero indicates that the model is worse than a horizontal line (i. e. the models predictions are very poor).

When designing models with multiple predictors there is always a risk of multicollinearities, i. e. a linear interdependency between the different predictors. This can be a consequence of both poorly designed experiments, where various predictors are not exited separately, and a true physical link between the predictors, which is often found in serially linked processes. To analyze and detect multicollinearities the Variance Inflation Factors (VIF) will be used in this paper (Belsley et al. 1980). VIF quantifies how much the variance is inflated, that is

$$VIF_k = \frac{1}{1 - R_k^2}, \quad (3)$$

where $R_k^2$ is the $R^2$ value obtained by regressing the $k$th predictor on the remaining predictors. A $VIF_k = 1$ means that there is no linear correlation between the $k$th predictor and the other remaining predictor variables, i. e. the variances in the estimated coefficients are not inflated. If $VIF > 4$, a general rule is that further analysis should be performed, while $VIF > 10$ indicates serious multicollinearities and a need to find a modified set of predictors. However, in our serially linked processes in Figure 1, some collinearities occur naturally and further analysis should be considered from a more holistic perspective.

**Table 1:** Variance inflated factors for the different predictors. C = Consistency, R = fiber residence time, SPE = Specific energy, FZ = flat zone, CD = Conical zone, In.CFZ = Inlet consistency of the flat zone.

| Variance inflation factors (VIF) | | | | | |
|---|---|---|---|---|---|
| SPE | C.FZ | C.CD | R.FZ | R.CD | In.CFZ |
| 3.6 | 14.7 | 16.5 | 75.5 | 76.9 | 1.1 |

**Table 2:** Correlation coefficients for pairs of potential predictors.

| Correlation Coefficient | | | | | | |
|---|---|---|---|---|---|---|
| | SPE | C.FZ | C.CD | R.FZ | R.CD | In.CFZ |
| SPE | 1 | 0.35 | 0.43 | 0.67 | 0.67 | −0.13 |
| C.FZ | | 1 | 0.96 | −0.18 | −0.19 | −0.24 |
| C.CD | | | 1 | −0.12 | −0.13 | −0.25 |
| R.FZ | | | | 1 | 0.99 | 0.09 |
| R.CD | | | | | 1 | 0.07 |
| In.CFZ | | | | | | 1 |

## Results and discussions

Using the VIF as described above, it is obvious that there is considerable interdependency for the fiber residence times and consistencies in our data set, see Table 1. Specific energy and FZ inlet consistency both have a VIF below 4 which indicates that the interdependencies are most likely not problematic.

If we further evaluate the results in Table 1 by looking at the Pearson's correlation coefficient, which evaluates linear interdependency between pairs of parameters, a broader picture is recieved, see Table 2. As seen in Table 2 the main interdependencies are between the measurements in the flat and conical zones which is expected as the process is serially linked. There is also some interdependency between specific energy and the fiber residence times, due to the fact that fiber residence times are linked to the production and the plate gaps, which both affect the specific energy.

The very high linear correlation between the different fiber residence times indicate that one of them should be discarded. However when doing that, the resulting models become considerably worse. The fiber residence times are partly influenced by the plate gaps of the individual zones. As these can be set individually, changes here will differ between the zones and thereby influence the pulp properties. To take this into account both fiber residence times may need to be included.

As a large amount of data is available, there is some flexibility when including several predictors as the risk for overfitting is relatively low. It should, however, be noted
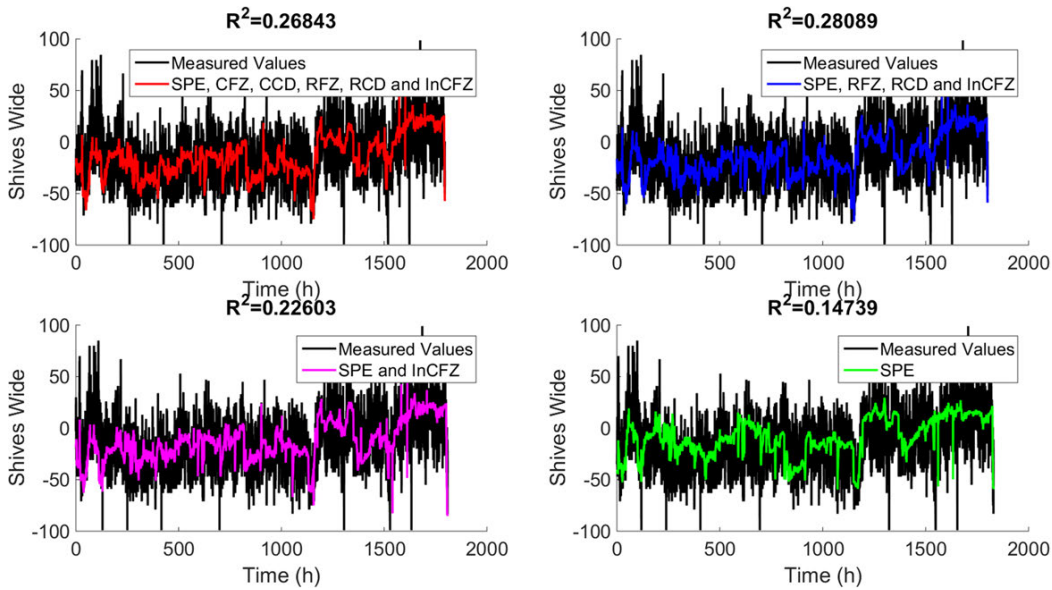
**Figure 3:** The predictions of shives width from models based on 4 different sets of predictors compared to the measured values on validation data.

that the high cross correlation of the fiber residence times and also the consistencies will make inference of the impact of the individual fiber residence times and consistencies difficult. If inference is an important aspect of the model, then the plate gap measurements can be used directly.

## Models for shives width

### Static models

To evaluate the models the data is split into two equally large sets: a training set and a validation set. The training set is used to design the model, while the validation set is used to evaluate the model. The evaluation is performed by examining the $R^2$ of the models, calculated using the validation set. To analyze the linear fit, all possible combinations of the inputs were considered when the models were designed, see Table A1 in Appendix. In Figure 3, the $R^2$ as well as the predictions are shown for some of the best and most interesting combination of inputs.

There are a few conclusions that can be drawn from Figure 3, firstly that the contribution of the consistencies to model quality seem to be marginal compared to a model with only specific energy, fiber residence times and FZ inlet consistency. The $R^2$ of the different models are all quite poor, which can be due to the considerable amount of measurement noise. This noise also makes visualisation of the models difficult and makes it hard to visually judge how well they predict shives width. To alleviate this, impact of

high frequency noise is reduced by smoothing the inputs and outputs with a size 50 moving average filter.

From Figure 4, which shows our models implemented on filtered data, it can be seen that with the models no longer trying to follow high frequency noise, the $R^2$ is considerably higher for all the models. We also see that our models can generally follow the major changes in shives width.

It is noteworthy that it seems to be sufficient to use specific energy and FZ inlet consistency as predictors to explain most of the variations. This, as the general shape of our model based on these two parameters follow the measured values quite well. One possible reason for this is that the residence times and consistencies influence specific energy to a certain degree, thus including them may not yield significant new information. However, it still seems to be beneficial to include residence time in our models, as this improves the model quality considerably.

Regarding the consistencies of each zone, it is worth mentioning that their impact on the models seems to be detrimental, which would imply that their influence on shives width seems to be marginal. This could be a consequence of the fact that the consistencies were controlled during the time the measurements were taken which means that the consistencies did not change very much during the test period. As the measurements were taken during normal operating conditions, this is a quite important conclusion as it implies that if consistency control is utilized then the shifts in consistency are so small that they will not significantly influence shives width.
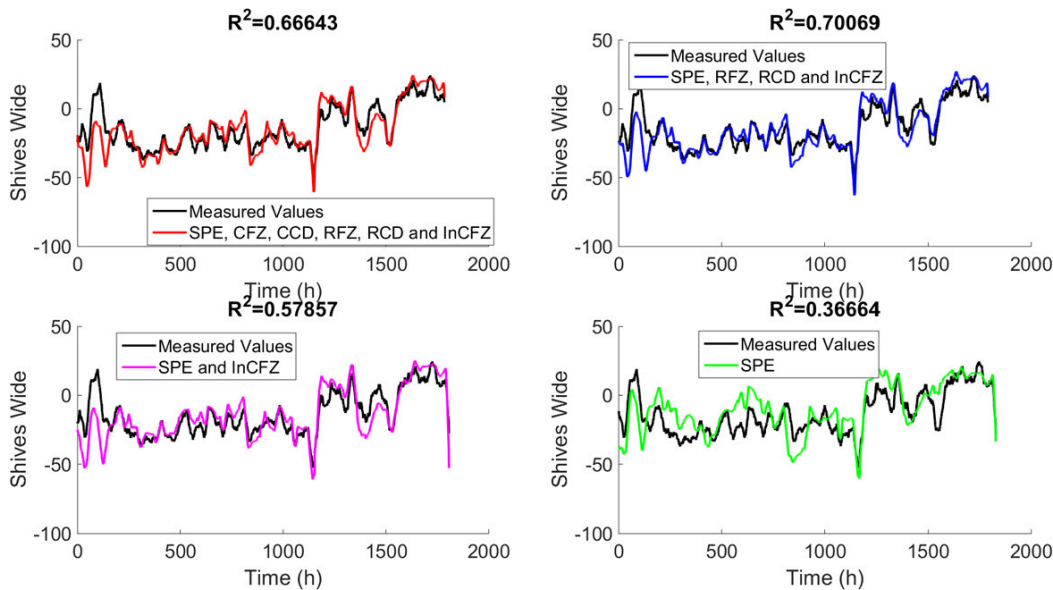
**Figure 4:** The predictions from models based on 4 different sets of predictors compared to the measured values on validation data with filtered inputs and outputs.

## Dynamic models

An alternative to the static models are dynamic or recursive models which change over time, as described previously. When designing our recursive models it was decided to use the set of parameters that our best static models used, that is, specific energy, fiber residence times and FZ inlet consistency. Thereafter, models with different window lengths and prediction horizons were implemented (note that the maximum window length is 7225 so a window length of that length would mean using all available data). In Table 3, the $R^2$s for the recursive models are presented with different window sizes and prediction horizons.

From Table 3 we can see that (with the exception of when the prediction horizon is very low), the largest window always yields the best result. In general this implies that there is little benefit to be extracted from discarding old data, which in turn can be seen as a indicator that the process did not change significantly during the measurement period.

## ARX models

As discussed previously an alternative are ARX models which can also model the system dynamics. If we implement the ARX models using different values of $n$ and compare it to static linear regression models we get the result shown in Table 4.

**Table 3:** $R^2$ for different window sizes and prediction horizons for recursive models.

| | | Size of window | | | | |
| | | 10 | 100 | 500 | 4000 | 7225 |
|---|---|---|---|---|---|---|
| Prediction | **0.5** | 0.96 | 0.93 | 0.84 | 0.85 | 0.87 |
| Horizon (min) | **5** | 0.25 | 0.61 | 0.72 | 0.84 | 0.87 |
| | **10** | −0.19 | 0.36 | 0.62 | 0.84 | 0.86 |
| | **50** | −2.20 | −0.29 | 0.64 | 0.85 | 0.87 |
| | **200** | −0.86 | −1.22 | 0.58 | 0.83 | 0.85 |

**Table 4:** A comparison between the $R^2$ for ARX and static linear regression models for shives width.

| Model | SPE, R.FZ, R.CD, In.CFZ | SPE, In.CFZ |
|---|---|---|
| Linear | 0.74 | 0.63 |
| ARX $n = 1$ | 0.66 | 0.67 |
| ARX $n = 3$ | 0.69 | 0.64 |
| ARX $n = 5$ | 0.63 | 0.64 |

As can be seen from Table 4 the ARX models do not yield a substantial improvement compared to static linear models, with considerably worse models when utilizing specific energy, FZ inlet consistency and fiber residence times, while being marginally better when only using specific energy and FZ inlet consistency. If we graphically compare the best models in Figure 5 we see that the resulting models are quite similar with only minor differences. Note that the $R^2$ for static linear models in Table 4 do not perfectly match those in Figure 4. As with ARX mod-
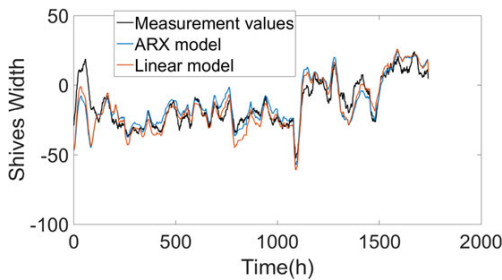
**Figure 5:** A comparison between the best ARX and static models for shives width.

**Table 5:** Parameter values for different predictors for shives width.

| Model | Unfiltered Data | Filtered data | ARX static gain |
|---|---|---|---|
| SPE (kWh/T) | −0.53 | −0.56 | −0.48 |
| R.FZ (s) | −336.55 | −326.99 | −87 |
| R.CD (s) | 2193.82 | 2169.07 | 526 |
| In.CFZ (%) | 49.75 | 48.94 | 55 |

els we are interested in the dynamics we chose to split the training and validation data sets so that the data covers an equal amount of time of data in the training and validation sets, while previously we split it so there was an equal amount of output measurements in the training and validation data sets.

As our ARX models do not yield an improved result compared to static linear models it seems to indicate that it is difficult to model the system dynamics from this data. This is, however not, very surprising as output data is only available at intervals of 10–30 minutes, while the expected dynamics of the system are considerably faster then this.

### Parameter values

The parameters values for the predictor for our different models are presented in Table 5. The more complex nature of the ARX models means that their parameters do not lend themselves to an easy interpretation. However the static gain for each predictor can be examined. This is similar to that of the parameters in the linear models in that it shows the amplification of low frequencies.

As can be seen from Table 5, the parameter values for filtered and unfiltered data are quite similar, which indicates that not much information was lost when applying the moving average filter. This is probably because that even with unfiltered data we are unable to model the high frequency dynamics.

Furthermore, we can see that an increase in FZ inlet consistency corresponds to an increase in shives width, while an increase in specific energy naturally causes shives width to decrease. From the model it is seen that the residence time in the flat zone adversely affects the shives width, however, as previously mentioned there is a very high linear dependency between the fiber residence times in the different zones. This means that it is unlikely that there is sufficient excitation of the different fiber residence times to accurately gauge their separate impact.

The static gain of the ARX model is also quite similar to the parameters of the static linear model except is for the residence times. However, as these residence times are highly correlated the individual parameter values are difficult to interpret. The similarity in parameter values between the models is expected as they are both linear models based on the same data set.

### Models for fiber length

If we repeat the modeling procedure described previously but for fiber length, a corresponding analysis can be made to yield a set of viable models. Table A1 in Appendix has the result for all the different combinations, and Figure 6 shows some of the more interesting results from the different combinations. From Figure 6 it is seen that the best combination seems to be when using specific energy and the fiber residence times in both refining zones. However this is somewhat misleading as visual inspection of the graphs shows that using more predictors will yield considerably improved results over nearly the entire data set, with the exception of the first 200 hours of the data set, where more predictors led to a very large overshoot. Further investigations indicate that this overshoot seems to be caused by high residence times. It is unclear whether this is due to measurement errors, non-linear dynamics or anomalous process conditions. However a single overshoot is not sufficient to disregard a model and thus we remove the first 200 hours of the validation data and investigate the remaining data. The results for this are shown in Figure 7 and in Table A1 in Appendix.

From Figure 7 we can see that without the first 200 hours considerable benefits can be obtained from utilizing more predictors. The best model utilizes specific energy, fiber residence times, FZ inlet consistency and the consistency in the conical zone. While this may seem similar to the results for shives width it differs in a few key aspects. Firstly the consistency of the conical zone seems to have a noteworthy effect on fiber length, considerably improving the model. While it is hard to gauge the exact impact (especially as we noted previously the consistencies were not individually exited to a particular degree) it seems to
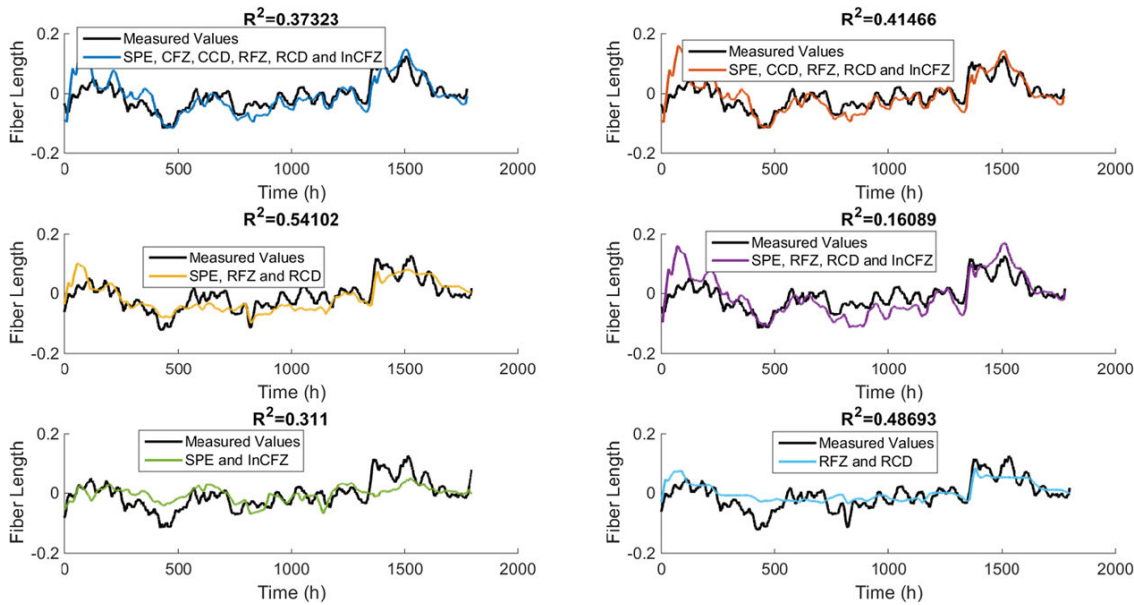
**Figure 6:** The predictions of fiber length from models based on 6 different sets of predictors compared to the measured values on validation data with filtered inputs and outputs.



**Figure 7:** The predictions of fiber length from models based on 6 different sets of predictors compared to the measured values on validation data with filtered inputs and outputs without the first 200 hours of validation data.
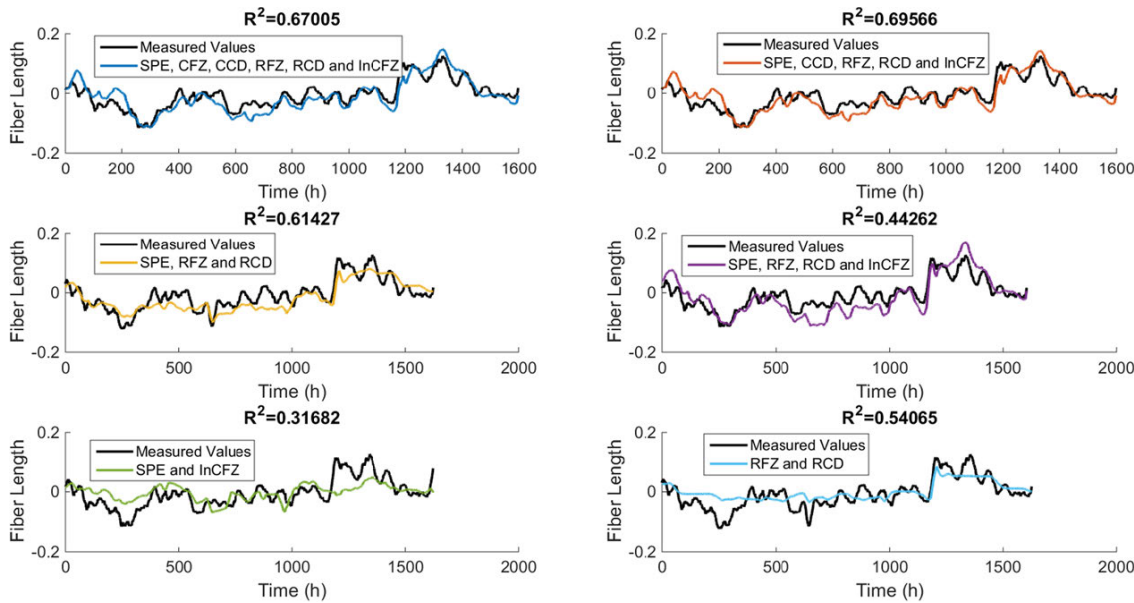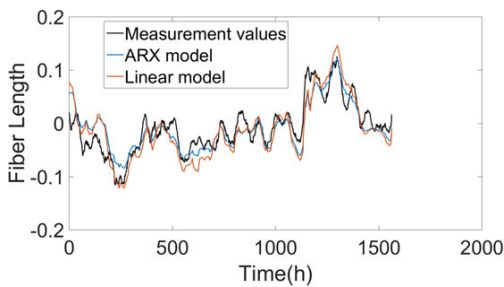
contribute to a notable improvement of model quality. It is also noteworthy that the effect of residence time on fiber length seems to be considerably more pronounced than on shives width. For shives width specific energy and FZ inlet consistency were sufficient to accurately predict when the major shifts would occur, while for fiber length it is apparent from Figure 7 that residence times are also needed in order to do the same.

If we also implement ARX models for fiber length and compare them to the static linear model we get the results portrayed in Table 6. From here we can see that the ARX model outperforms the static linear model somewhat when it comes to $R^2$. This may be seen to imply that there may be slow dynamics affecting the fiber length that our ARX models can detect. However if we compare our best ARX and linear models graphically in Figure 8 we see that

**Table 6:** A comparison between the $R^2$ for ARX and static linear regression models for fiber length.

| Model | SPE, R.FZ, R.CD, In.CFZ | SPE, C.CD, R.FZ, R.CD, In.CFZ |
|---|---|---|
| Linear | 0.41 | 0.67 |
| ARX $n = 1$ | 0.69 | 0.75 |
| ARX $n = 3$ | 0.67 | 0.74 |
| ARX $n = 5$ | 0.66 | 0.75 |

**Table 7:** Parameter values for different predictors for fiber length.

| Model | Unfiltered Data | Filtered data | ARX Static gain |
|---|---|---|---|
| SPE (kWh/T) | −0.0011 | −0.0013 | −0.0012 |
| Cons. CD (%) | 0.0031 | 0.0052 | 0.0044 |
| Res.FZ (s) | −2.16 | −2.22 | −2.45 |
| Res.CD (s) | 13.88 | 14.49 | 16.2 |
| In.CFZ (%) | −0.085 | −0.088 | −0.11 |



**Figure 8:** A comparison between the best ARX and linear models for fiber length.

they are quite similar with only minor differences in behaviour.

From these results we propose a model of fiber length, utilizing specific energy, fiber residence times, consistency in the conical zone and FZ inlet consistency. The parameters of this model are presented in Table 7. We also present the static gain for the best ARX models (utilizing specific energy, fiber residence times and FZ inlet consistency).

As can be seen from the parameters in Table 7, they differ somewhat compared with those for shives width in Table 5. Inlet consistency seems to have a detrimental impact on fiber length, while it increased shives width. This is not unexpected as high inlet consistencies are due to low amount of sawmill chips, which tend to have longer fibers. Moreover the regression analysis indicates that increasing specific energy decreases fiber length, which matches well with previous results (Lecourt et al. 2006).

## Models for freeness

Repeating the same procedure for freeness gives the results shown in Table A1. As seen in Table A1 no combinations of inputs has a positive $R^2$. This means that no models yield a better result than a horizontal line, which in turn implies that there is no clear linear relationship between freeness and the tested inputs.

If we examine some of the models more closely in Figure 9, it is seen that the major changes in freeness is caused by something other than the predictors used. Even disre-

garding this, it is apparent that the models predict freeness very poorly in most cases. There are some indications that our models shape follow the freeness somewhat at certain places in the validation data. For instance some models predict the general shape to some extent for the last 200 hours of the validation data, which might indicate some correlation between the inputs and outputs. This indicates that to properly model freeness other predictors are required, or possibly more complex non-linear models should be used to handle freeness estimations. These results correspond well to those presented by Luukkonen et al. (2012), who found the relationship between freeness and refiner properties to be highly non-linear.

## Parameter convergence

Another element that can be investigated is how fast the parameters converge to their final values, as this can give a indication of how much data is required to derive models, as well as an idea of how much the model is likely to change with the addition of new data. Therefore models are derived using different amounts of data. The parameters are analyzed and visually it is seen how they converge to the final values for shives width in Figure 10.

As can be seen from Figure 8 the parameters for residence time seem to converge after approximately 1800 hours, with specific energy converging after approximately 1400 hours of data. The FZ inlet consistency's parameter converges already after 500 hours of data. It is interesting that the fiber residence times converge around 1800 hours as this is after approximately half the data. As two approximately equally sized data sets were used this means that the fiber residence times converge when data from the other data set started to be used. To further explore this we plot the parameter convergence when only using the second data set in Figure 11.

As can be seen from Figure 11 when using only the second data set the parameters converge around 700 hours of data. So considerably less data is required here. This may be indicative that the data of the second set may be more useful from a modelling perspective, which in turn may be
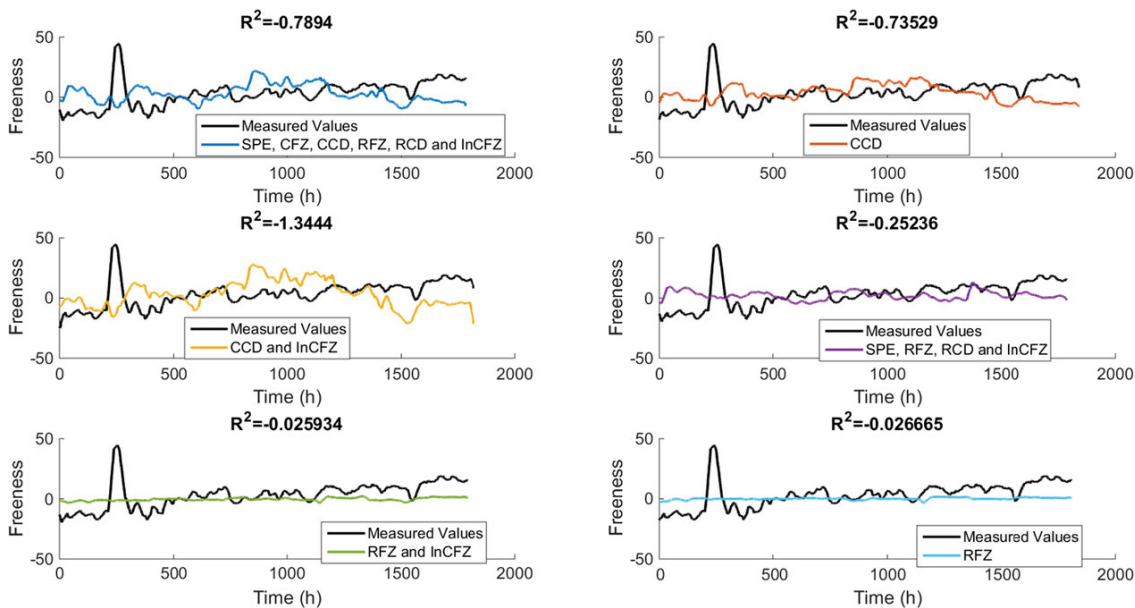
**Figure 9:** The predictions of freeness from models based on 6 different sets of predictors compared to the measured values on validation data with filtered inputs and outputs.
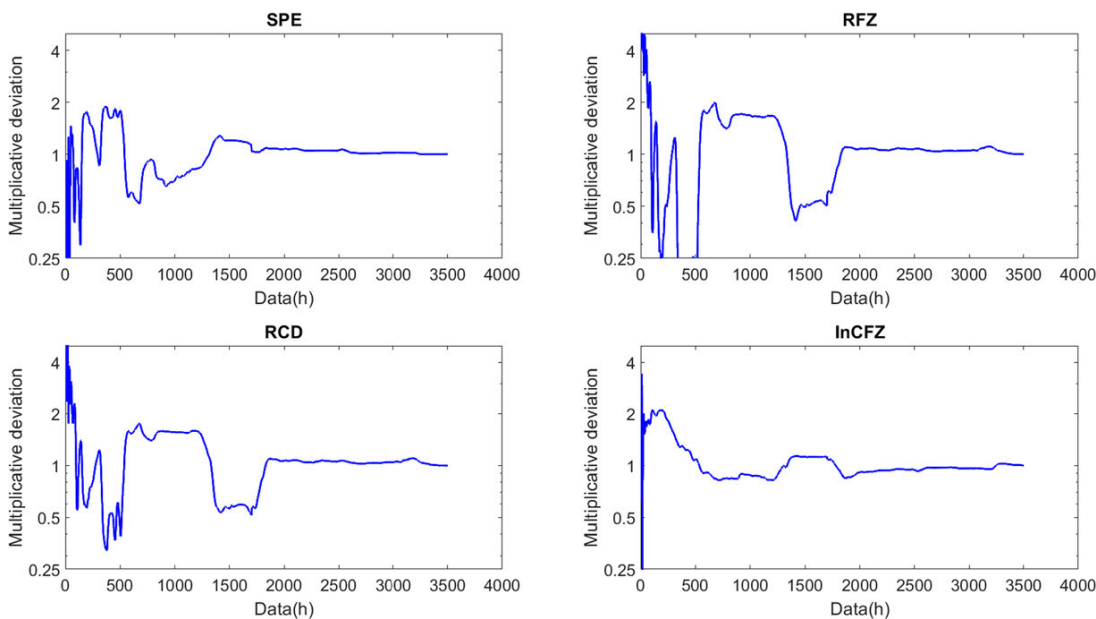


**Figure 10:** Parameter convergence as percentage of final value for shives width, using both data sets.

construed to indicate that fewer disturbances were present during the time this data set. Alternatively this can be seen to indicate that the excitation of the predictors was more pronounced for the second data set than for the first.

Ultimately this highlights the problem with determining the amount of data required to derive reliable models, as it is highly dependent on the quality of the data available. In our case as much as 1800 hours or as little as 700 hours of data could be enough for the parameters to con-

verge. However, even with the data from the first set good quality models could be acquired, as demonstrated when we designed models using training and validation data (as the data sets were appended to each other chronologically, the training data consisted nearly exclusively of the data of the first data set).

If we repeat the same investigations for fiber length (figures shown in Appendix) we see that the results differ somewhat from shives width. Though we still need around
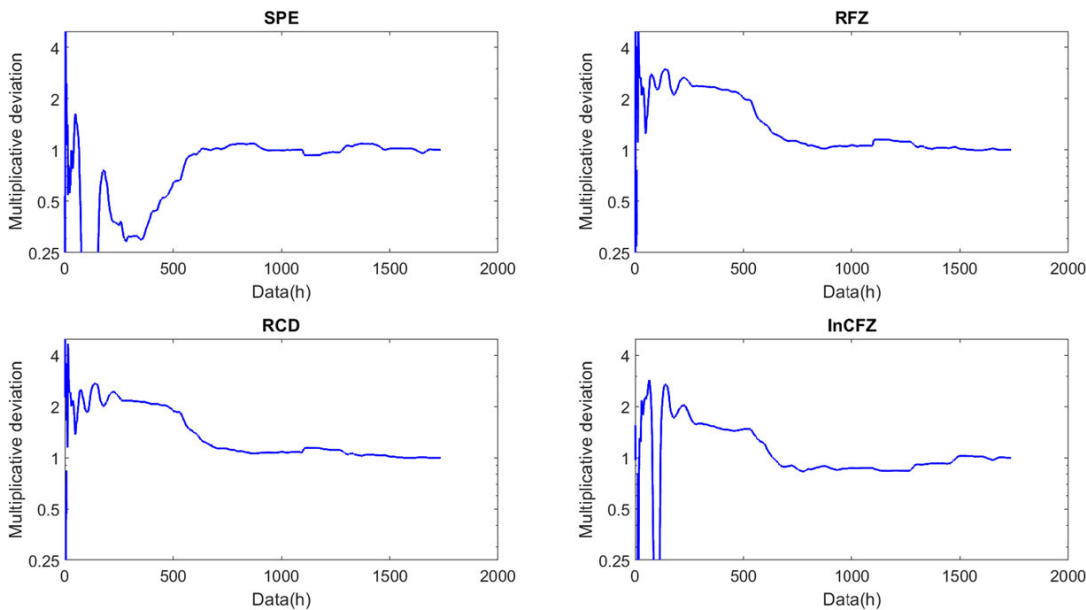
**Figure 11:** Parameter convergence as percentage of final value for shives width using only the second data set.

1500 hours of data for the parameters to converge when we look at the entire data set. For the case when using only the second data set the parameter for specific energy converges quite quickly, but for FZ inlet consistency and fiber residence times nearly the entire data set is required for convergence.

## Conclusion

In this paper, models for shives width, fiber length and freeness (outputs) have been designed using a set of external and internal variables (inputs). For shives width and fiber length models capable of predicting most of the major variations can be derived. For freeness, no good linear models could be derived, implying either highly nonlinear relationships between the inputs and outputs or, alternatively, that the main factors affecting freeness could not be derived from the inputs we used.

When estimating shives width the inputs specific energy, FZ inlet consistency and fiber residence times were the inputs with the most pronounced effect. Generally, FZ inlet consistency and especially specific energy seem to be the most influential inputs. For fiber length we found that a similar set of inputs resulted in the best models, with the addition of the consistency in the conical zone. Fiber residence times however had a considerably more pronounced impact on fiber length then they had on shives width.

A limitation regarding the measurements taken from normal process conditions, was that the inputs were not exited to the extent that would be ideal from a modelling perspective. Due to colinearities in the inputs this made it somewhat difficult to distinguish the impact of individual inputs. This was particularly problematic for the fiber residence times, as they were barely exited separately at all. Further experiments with deliberately changing the residence time in only one zone may allow one to better assess the impact of the residence time. It is important to note that many of the parameter used, such as the residence times and consistencies, were derived from internal and external measurements of the refining process. Thus, by examining these non-linear models, it may be possible to anticipate the colinearities, and find new, less co-linear parameters.

The high noise level in the outputs meant that only low frequency dynamics were examined. Better and more frequent output measurements would have allowed us to further explore the high frequency dynamics. More frequent output measurements would also allow for model structures capable of capturing the noise dynamics, such as AR-MAX and Box-Jenkins models.

Finally, the poor estimate of freeness can also be interpreted in other ways as it indicates that freeness should be avoided when it comes to pulp property control, even though it is still a measure commonly used in the pulp and paper industry.

**Conflict of interest:** The authors declare no conflicts of interest.

# Appendix

**Table A1:** $R^2$ for different inputs and outputs on filtered and unfiltered data. A one denotes that this output was used and a zero denotes it was not utilized.

| Predictors | | | | | | Shives width | | Fiber length | | | Freeness | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SPE | C.FZ | C.CD | R.FZ | R.CD | In.CFZ | Unfiltered | Filtered | Unfiltered | Filtered | Filtered, reduced data set | Unfiltered | Filtered |
| 1 | 0 | 0 | 0 | 0 | 0 | 0.15 | 0.37 | 0.03 | 0.14 | 0.18 | −0.11 | −0.22 |
| 0 | 1 | 0 | 0 | 0 | 0 | −0.12 | −0.09 | −0.17 | −0.40 | −0.40 | −0.42 | −0.59 |
| 0 | 0 | 1 | 0 | 0 | 0 | −0.08 | −0.01 | −0.21 | −0.55 | −0.58 | −0.48 | −0.74 |
| 0 | 0 | 0 | 1 | 0 | 0 | −0.17 | −0.54 | −0.02 | −0.02 | −0.02 | −0.02 | −0.03 |
| 0 | 0 | 0 | 0 | 1 | 0 | −0.18 | −0.58 | −0.03 | −0.05 | −0.05 | −0.04 | −0.06 |
| 0 | 0 | 0 | 0 | 0 | 1 | −0.10 | −0.31 | 0.07 | 0.12 | 0.11 | −0.07 | −0.09 |
| 1 | 1 | 0 | 0 | 0 | 0 | 0.13 | 0.35 | −0.13 | −0.21 | −0.12 | −0.42 | −0.66 |
| 1 | 0 | 1 | 0 | 0 | 0 | 0.13 | 0.32 | −0.19 | −0.39 | −0.33 | −0.47 | −0.77 |
| 1 | 0 | 0 | 1 | 0 | 0 | 0.15 | 0.41 | 0.04 | 0.12 | 0.15 | −0.14 | −0.25 |
| 1 | 0 | 0 | 0 | 1 | 0 | 0.16 | 0.44 | 0.06 | 0.17 | 0.20 | −0.14 | −0.26 |
| 1 | 0 | 0 | 0 | 0 | 1 | 0.23 | 0.58 | 0.11 | 0.31 | 0.32 | −0.13 | −0.26 |
| 0 | 1 | 1 | 0 | 0 | 0 | −0.06 | 0.01 | −0.22 | −0.65 | −0.69 | −0.48 | −0.75 |
| 0 | 1 | 0 | 1 | 0 | 0 | −0.39 | −0.95 | −0.06 | −0.12 | −0.12 | −0.41 | −0.73 |
| 0 | 1 | 0 | 0 | 1 | 0 | −0.31 | −0.72 | −0.10 | −0.23 | −0.24 | −0.46 | −0.81 |
| 0 | 1 | 0 | 0 | 0 | 1 | 0.01 | 0.07 | −0.02 | −0.11 | −0.13 | −0.60 | −0.92 |
| 0 | 0 | 1 | 1 | 0 | 0 | −0.40 | −1.12 | −0.10 | −0.24 | −0.26 | −0.46 | −0.87 |
| 0 | 0 | 1 | 0 | 1 | 0 | −0.31 | −0.80 | −0.14 | −0.37 | −0.39 | −0.51 | −0.93 |
| 0 | 0 | 1 | 0 | 0 | 1 | 0.06 | 0.19 | −0.04 | −0.25 | −0.29 | −0.79 | −1.35 |
| 0 | 0 | 0 | 1 | 1 | 0 | −0.17 | −0.58 | 0.19 | 0.49 | 0.54 | −0.19 | −0.31 |
| 0 | 0 | 0 | 1 | 0 | 1 | −0.04 | −0.17 | 0.04 | 0.10 | 0.08 | −0.02 | −0.03 |
| 0 | 0 | 0 | 0 | 1 | 1 | −0.09 | −0.31 | 0.04 | 0.09 | 0.07 | −0.04 | −0.06 |
| 1 | 1 | 1 | 0 | 0 | 0 | 0.13 | 0.35 | −0.20 | −0.50 | −0.48 | −0.47 | −0.78 |
| 1 | 1 | 0 | 1 | 0 | 0 | 0.06 | 0.23 | −0.24 | −0.67 | −0.51 | −0.33 | −0.55 |
| 1 | 1 | 0 | 0 | 1 | 0 | 0.11 | 0.34 | −0.31 | −0.91 | −0.68 | −0.44 | −0.73 |
| 1 | 1 | 0 | 0 | 0 | 1 | 0.24 | 0.59 | 0.01 | 0.05 | 0.07 | −0.62 | −1.02 |
| 1 | 0 | 1 | 1 | 0 | 0 | 0.05 | 0.19 | −0.49 | −1.63 | −1.56 | −0.37 | −0.69 |
| 1 | 0 | 1 | 0 | 1 | 0 | 0.11 | 0.33 | −0.55 | −1.78 | −1.70 | −0.50 | −0.90 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0.23 | 0.54 | −0.04 | −0.18 | −0.16 | −0.79 | −1.37 |
| 1 | 0 | 0 | 1 | 1 | 0 | 0.16 | 0.40 | 0.23 | 0.54 | 0.61 | −0.19 | −0.33 |
| 1 | 0 | 0 | 1 | 0 | 1 | 0.22 | 0.54 | 0.14 | 0.34 | 0.34 | −0.17 | −0.29 |
| 1 | 0 | 0 | 0 | 1 | 1 | 0.23 | 0.55 | 0.17 | 0.43 | 0.43 | −0.18 | −0.31 |
| 0 | 1 | 1 | 1 | 0 | 0 | −0.42 | −1.13 | −0.11 | −0.33 | −0.35 | −0.46 | −0.87 |
| 0 | 1 | 1 | 0 | 1 | 0 | −0.33 | −0.81 | −0.15 | −0.45 | −0.49 | −0.51 | −0.93 |
| 0 | 1 | 1 | 0 | 0 | 1 | 0.07 | 0.23 | −0.05 | −0.34 | −0.38 | −0.79 | −1.44 |
| 0 | 1 | 0 | 1 | 1 | 0 | −0.53 | −1.36 | 0.19 | 0.46 | 0.51 | −0.29 | −0.53 |
| 0 | 1 | 0 | 1 | 0 | 1 | −0.06 | −0.14 | 0.05 | 0.11 | 0.08 | −0.67 | −1.22 |
| 0 | 1 | 0 | 0 | 1 | 1 | −0.05 | −0.10 | 0.02 | 0.00 | −0.02 | −0.71 | −1.25 |
| 0 | 0 | 1 | 1 | 1 | 0 | −0.70 | −2.05 | 0.19 | 0.46 | 0.51 | −0.28 | −0.54 |
| 0 | 0 | 1 | 1 | 0 | 1 | −0.03 | −0.14 | 0.03 | 0.00 | −0.03 | −0.84 | −1.65 |
| 0 | 0 | 1 | 0 | 1 | 1 | 0.00 | −0.03 | −0.01 | −0.12 | −0.15 | −0.86 | −1.64 |
| 0 | 0 | 0 | 1 | 1 | 1 | −0.12 | −0.44 | 0.18 | 0.43 | 0.59 | −0.14 | −0.27 |
| 1 | 1 | 1 | 1 | 0 | 0 | 0.05 | 0.21 | −0.48 | −1.71 | −1.73 | −0.37 | −0.69 |
| 1 | 1 | 1 | 0 | 1 | 0 | 0.11 | 0.35 | −0.54 | −1.83 | −1.83 | −0.50 | −0.90 |
| 1 | 1 | 1 | 0 | 0 | 1 | 0.22 | 0.49 | −0.05 | −0.31 | −0.32 | −0.78 | −1.44 |
| 1 | 1 | 0 | 1 | 1 | 0 | −0.03 | 0.01 | 0.17 | 0.41 | 0.62 | −0.26 | −0.46 |
| 1 | 1 | 0 | 1 | 0 | 1 | 0.22 | 0.55 | −0.09 | −0.37 | −0.29 | −0.56 | −0.98 |
| 1 | 1 | 0 | 0 | 1 | 1 | 0.23 | 0.57 | −0.16 | −0.58 | −0.44 | −0.66 | −1.14 |
| 1 | 0 | 1 | 1 | 1 | 0 | −0.11 | −0.27 | 0.13 | 0.23 | 0.37 | −0.24 | −0.45 |

**Table A1:** (continued)

| Predictors | | | | | | Shives width | | Fiber length | | | Freeness | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SPE | C.FZ | C.CD | R.FZ | R.CD | In.CFZ | Unfiltered | Filtered | Unfiltered | Filtered | Filtered, reduced data set | Unfiltered | Filtered |
| 1 | 0 | 1 | 1 | 0 | 1 | 0.23 | 0.53 | −0.32 | −1.47 | −1.45 | −0.76 | −1.52 |
| 1 | 0 | 1 | 0 | 1 | 1 | 0.23 | 0.54 | −0.38 | −1.57 | −1.53 | −0.87 | −1.65 |
| 1 | 0 | 0 | 1 | 1 | 1 | 0.28 | 0.70 | 0.13 | 0.16 | 0.44 | −0.14 | −0.25 |
| 0 | 1 | 1 | 1 | 1 | 0 | −0.70 | −2.07 | 0.19 | 0.46 | 0.51 | −0.29 | −0.54 |
| 0 | 1 | 1 | 1 | 0 | 1 | −0.06 | −0.17 | 0.02 | −0.08 | −0.11 | −0.84 | −1.67 |
| 0 | 1 | 1 | 0 | 1 | 1 | −0.04 | −0.09 | −0.01 | −0.20 | −0.24 | −0.86 | −1.66 |
| 0 | 1 | 0 | 1 | 1 | 1 | −0.07 | −0.20 | 0.06 | −0.12 | 0.12 | −0.36 | −0.70 |
| 0 | 0 | 1 | 1 | 1 | 1 | −0.10 | −0.46 | −0.06 | −0.65 | −0.39 | −0.42 | −0.94 |
| 1 | 1 | 1 | 1 | 1 | 0 | −0.11 | −0.25 | 0.13 | 0.22 | 0.34 | −0.25 | −0.44 |
| 1 | 1 | 1 | 1 | 0 | 1 | 0.21 | 0.46 | −0.32 | −1.64 | −1.69 | −0.76 | −1.55 |
| 1 | 1 | 1 | 0 | 1 | 1 | 0.22 | 0.47 | −0.38 | −1.67 | −1.70 | −0.87 | −1.67 |
| 1 | 1 | 0 | 1 | 1 | 1 | 0.26 | 0.66 | 0.15 | 0.30 | 0.61 | −0.32 | −0.60 |
| 1 | 0 | 1 | 1 | 1 | 1 | 0.28 | 0.71 | 0.18 | 0.41 | 0.70 | −0.37 | −0.83 |
| 0 | 1 | 1 | 1 | 1 | 1 | −0.11 | −0.41 | −0.07 | −0.80 | −0.53 | −0.41 | −0.91 |
| 1 | 1 | 1 | 1 | 1 | 1 | 0.27 | 0.67 | 0.17 | 0.37 | 0.67 | −0.36 | −0.79 |
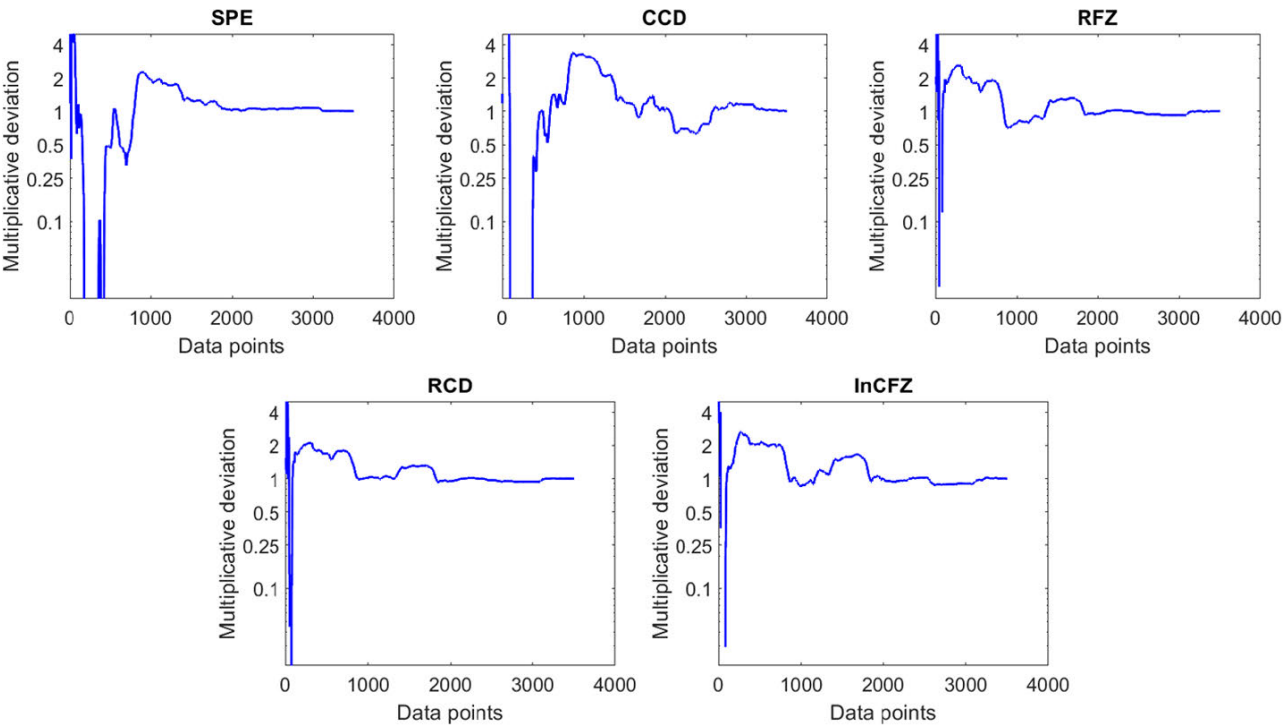


**Figure A1:** Parameter convergence as percentage of final value for fiberlength.
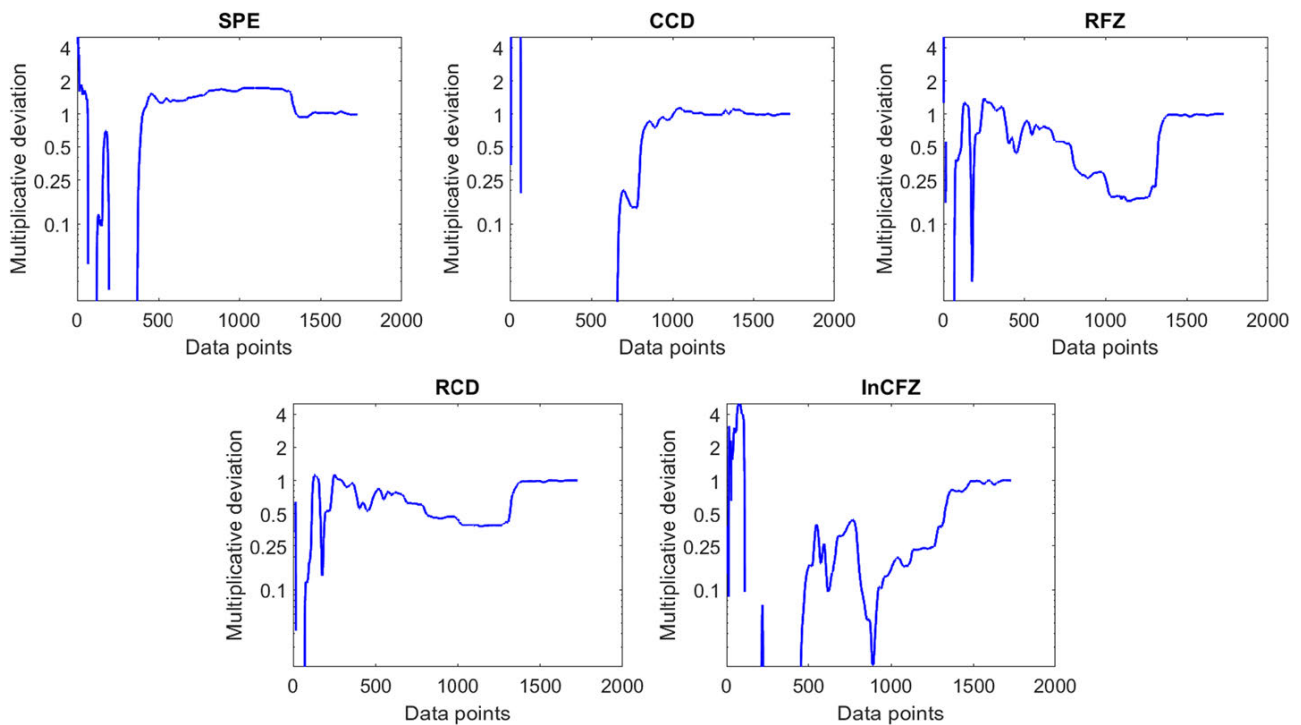
**Figure A2:** Parameter convergence as percentage of final value for fiber length using only the second data set. Note that missing data is due to it being of a negative sign which cannot be portrayed with a logarithmic axis.

# References

Belsley, D.A., Kuh, E., Welsch, R.E. Regression Diagnostics. Identifying Influential Data and Sources of Collinearity. John Wiley & Sons, Hoboken, USA, 1980.

Bengtsson, F., Karlström, A., Wik, T. (2020) Modeling of tensile index using uncertain data sets. Nord. Pulp Pap. Res. J. 35:231–242.

Blanco, A., Dahlquist, E., Kappen, J., Manninen, J., Negro, C., Ritala, R. (2009) Use of modelling and simulation in the pulp and paper industry. Math. Comput. Model. Dyn. Syst. 15:409–423.

Draper, N.R., Smith, H. Applied Regression Analysis. vol. 326, John Wiley & Sons, New York, USA, 1998.

Härkönen, E., Huusari, E., Ravila, P. (2000) Residence time of fibre in a single disc refiner. Pulp Pap. Can. 101:330–335.

Härkönen, E., Kortelainen, J., Virtanen, J., Vuorio, P. (2003) Fiber development in TMP main line. In: Int. Mech. Pulping Conf., Quebec. Canada. pp. 171–178.

Hirn, U., Bauer, W. (2006) A review of image analysis based methods to evaluate fiber properties. Lenzing. Ber. 86:96–105.

Karlström, A., Eriksson, K. (2014a) Fiber energy efficiency Part I: Extended entropy model. Nord. Pulp Pap. Res. J. 29:322–331.

Karlström, A., Eriksson, K. (2014b) Fiber energy efficiency Part II: Forces acting on the refiner bars. Nord. Pulp Pap. Res. J. 29:332–343.

Karlström, A., Eriksson, K. (2014c) Fiber energy efficiency Part III: Modeling of bar-to-fiber interaction. Nord. Pulp Pap. Res. J. 29:401–408.

Karlström, A., Eriksson, K. (2014d) Fiber energy efficiency Part IV: Multi-scale modeling of refining processes. Nord. Pulp Pap. Res. J. 29:409–417.

Karlström, A., Hill, J. (2017a) CTMP process optimization Part I: Internal and external variables impact on refiner conditions. Nord. Pulp Pap. Res. J. 32:35–44.

Karlström, A., Hill, J. (2017b) CTMP process optimization Part II: Reliability in pulp and handsheet Measurements. Nord. Pulp Pap. Res. J. 32:253–265.

Karlström, A., Hill, J. (2017c) CTMP process optimization Part III: On the predicition of Scott-bond, z-strength and tensile index. Nord. Pulp Pap. Res. J. 32:266–279.

Lecourt, M., Bouvet, A., Deleuze-Brezins, C. (2006) Modelling TMP fibre morphology and pulp properties from wood and forest data: The example of Norway spruce. N. Z. J. For. Sci. 36:435.

Ljung, L. System Identification: Theory for the User. Prentice Hall information and system sciences series. Prentice Hall PTR, 1999. ISBN 9780136566953.

Luukkonen, A., Olson, J., Martinez, D.M. (2012) Mechanical Pulping: Low consistency refining of mechanical pulp: Relationships between refiner operating conditions and pulp properties. Nord. Pulp Pap. Res. J. 27:882–885.

Nelsson, E. Improved energy efficiency in mill scale production of mechanical pulp by increased wood softening and refining intensity. Ph.D. thesis. Mid Sweden University, 2016.

Sabourin, M., Wiseman, N., Vaughn, J., et al. (2001) Refining theory considerations for assessing pulp properties in the commercial manufacture of TMP. In: 55th Appita Annual Conference. Appita Inc. pp. 195–204.

Sanchis, R., Albertos, P. (2002) Recursive identification under scarce measurements—convergence analysis. Automatica 38:535–544.

Schwartz, H.M., Chang, G., Liu, Y., Phung, T. (1996) A method of modeling, predicting and controlling TMP pulp properties. In: Proceeding of the 1996 IEEE International Conference on Control Applications. IEEE. pp. 846–851.

Strand, B. (1996) Model-based control of high-consistency refining. Tappi J. 10:140–146.

Strand, B., Grace, B. (2014) Implementation of advanced supervisory control within a TMP refiner quality control system. In: Int. Mech. Pulping Conf., Helsinki, Finland.

Young, P.C. Recursive estimation and time-series analysis: An introduction for the student and practitioner. Springer Science & Business Media, 2011.