

# Data-Driven Soft Sensors in Pulp Refining Processes Using Artificial Neural Networks

Anders Karlström,<sup>a</sup> Jan Hill,<sup>b</sup> and Lars Johansson<sup>c</sup>

Pulp refining processes are most often complicated to describe using linear methodologies, and sometimes an artificial neural network (ANN) is a preferable alternative when assimilating non-linear operating data. In this study, an ANN is used to predict pulp properties, such as shives (wide), fiber length, and freeness. Both traditional process variables (external variables) and refining zone variables (internal variables) are necessary to include as model inputs. The estimation of shives (wide) results achieved an  $R^2$  (coefficient of determination) of 0.9 (0.7) for the training and (validation) sets. Corresponding measures for fiber length and freeness can be questioned using this methodology. It is shown that the maximum temperature in the flat zone can be modeled using the external variables motor load and production instead of the specific energy. This resulted in an  $R^2$  of approximately 0.9 for the training sets, while the  $R^2$  for the validation set did not reach an acceptable level – most likely due to inherent non-linearities in the process. Additional results showed that the consistency profile is difficult to estimate properly using an ANN. Instead, a model-driven sensor is preferred to be used. The main results from this study indicate that shives (wide) should be the prime candidate when introducing advanced pulp property control concepts.

DOI: 10.15376/biores.19.1.1030-1057

Keywords: ANN models; Soft sensors; Consistency; Temperature; Pulp property estimation

Contact information: a: Chalmers University of Technology, Department of Electrical Engineering, SE-412 96 Gothenburg, Sweden; b: QualTech Jan Hill AB, SE-282 21 Tyringe, Sweden; c: Woodworks Cluster, Skognæringa Kyst, Skolegata 22, Landbruksklyngen bygg P2, NO-7713 Steinkjer, Norway;

\* Corresponding author: lars@woodworkscluster.no

## INTRODUCTION

Many systems in the pulp and paper industry can be classified as nonlinear. Some of them are possible to describe using first principle models (FPM). If the nonlinearities are weak, linear models based on gain scheduling can be derived as an alternative. In other situations, the complexity can be almost impossible to describe, which calls for the implementation of complementary modeling techniques. When producing chemi-thermo-mechanical pulp (CTMP), this is certainly true, as the needed input-output variables can be unknown as well as non-linear, noisy, and imprecise (Karlström and Hill 2017a,b,c; Karlström *et al.* 2018a; Karlström and Hill 2018b; Sund *et al.* 2021).

To approach the challenge, a hybrid concept of both model-driven and data-driven soft sensors is used in this paper. These two soft sensors are well described by Bohlin (2006) and Kadlec *et al.* (2008), and a short overview is given below of pulp refining processes. The traditional process variables are most often inputs to a model-driven sensor. Most often, the model-driven soft sensors are based on FPM, extended Kalman filter, and adaptive observers (Bastin and Dochain 1990; Chruy 1997; Jos de Assis and Filho 2000).

In refining processes, the physical models (FPM) are of special interest, as non-linear states can be derived (Miles and May 1990, 1991; Harkonen *et al.* 1999, 2003; Huhtanen 2004; Eriksson 2005; Karlström *et al.* 2008; Eriksson 2009). However, such models are often not suitable for online applications.

Through measuring the temperature profile inside the refining zone, Karlström and Eriksson (2014a) introduced a simplified entropy modeling concept to cope with defibration/ fibrillation in three-phase systems. This model allowed for online estimation of internal variables, such as consistency profile, backward and forward flowing steam (the direction of the steam is dependent on where to find the maximum temperature (Karlström and Eriksson 2014b, 2014c, 2014d)), shear forces between chips/pulp and refining segments, *etc.* Thereby, together with the traditional process variables (such as motor load, production (also called throughput), and dilution water feed rates (also called external variables)), the number of process variables expanded considerably, which also allowed for a broader use of data-driven soft sensors where for instance parametric black-box models have been used successfully in several processes (Ljung 1987; Principe *et al.* 2000; Wold *et al.* 2001; Jolliffe 2002). However, parametric models are often based on linear methodology, requiring updated routines of the model parameters to cover an entire process operation window that also changes over time (Berg and Karlström 2005; Eriksson 2005; Eriksson 2009; Karlström *et al.* 2015b, 2016a,b). This limits the use of such models (Karlström *et al.* 2022). Therefore, data-driven models based on artificial neural networks (ANN), with multiple layers of neurons, can be an alternative to handle inherent non-linearities (Demuth and Beale 2004; Rajesh and Ray 2006).

In this paper, the authors focused on ANN-models based on a gradient- or Jacobian-based Levenberg-Marquardt method (LM) (Levenberg 1944; Marquardt 1963). This algorithm was used to solve non-linear least square problems. Other backpropagation methods can be used as well (Hagan and Menhaj 1994; Björck 1996; Hagan *et al.* 1996; Kermani *et al.* 2005; Bonnans *et al.* 2006), but LM has proven to be both robust and relatively fast, which is favorable when handling large industrial data sets of information.

In this paper, the authors concentrated the analysis on the estimation of a) the maximum temperatures in the flat zone (FZ) and the conical zone (CD); b) consistencies out from FZ and CD; and c) the pulp properties shives (wide), fiber length, and freeness. It is worth mentioning that the measured shives can be wide or long, where the term “wide” refers to the thickness of the fibers. The term “long” refers to the length of the fibers on average. Here, the focus is on the measured shives (wide), as the fiber length is also being used as a measure.

## EXPERIMENTAL

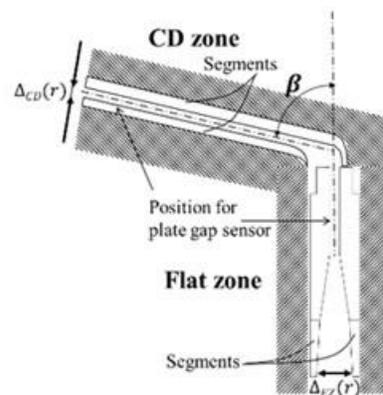
### Materials and Methods

This section covers an introduction to special process conditions in a full-scale CTMP production line followed by information about the ANN-structure used. Some of all possible model input candidates are thereafter penetrated in more detail. The output candidates will be temperatures and consistencies in the refining zones and pulp properties with specific criteria for the selected training and validation sets.

## Fundamentals

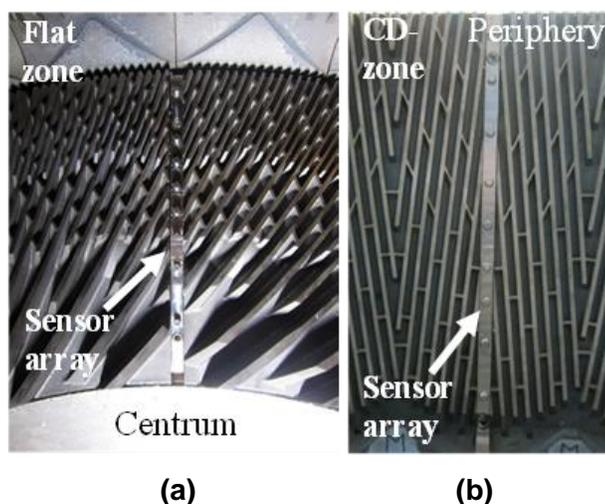
In this study, the production line consisted of an RGP82CD (with a capacity of about 36 T/h and a normal operation of 16 to 20 MW). Three small 6 MW refiners (RLP54) ran in parallel but the focus was the CD-refiner.

The CD-refiner, consists of two serially linked refining zones called the flat zone (FZ) and the conical zone (CD), see Fig. 1. In each refining zone, refining segments were mounted on a stationary and a rotating disc. The chips were fed into the FZ refining zone and moved towards the periphery of the CD by centrifugal forces.



**Fig. 1.** A schematic drawing of a CD refiner. The vertical FZ is directly linked to the CD *via* an expanding point. Distance between the stationary and rotating segments (plate gaps) is represented by  $\Delta_{CD}$  and  $\Delta_{FZ}$ . For clarification, the pulp is introduced from below in the figure, *i.e.* at the label “ $\Delta_{FZ}$ ”.

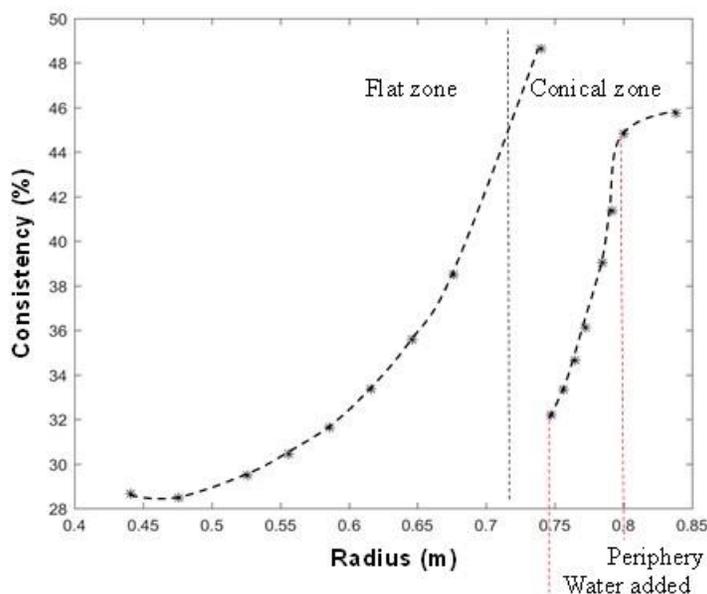
The segments in Fig. 1 have a specific pattern with several bars to defibrilate chips and to fibrillate the pulp along the way from the centrum (Fig. 2a) to the periphery of the refiner (Fig. 2b).



**Fig. 2.** Sensor arrays with eight temperature measurements each, mounted along the radius and between refining segments in a CD82-refiner, see Fig. 1.

It is a real challenge to describe the final pulp quality (Ferritsius *et al.* 2018; Ferritsius 2021). As Fig. 2 shows, the temperature profiles were measured using sensor arrays mounted between two segments in each refining zone. Eight temperature sensors were used in each zone. Thereby, the temperature profile can be seen as an internal variable vector that is measured together with traditional process variables (external variables) such as production rate, dilution water flows, motor load, and distance between the refining segments.

Through using the temperature profile, it is possible to get access to other important internal variables utilizing a non-linear physical model derived by Karlström and Eriksson (2014a,b,c,d). This model opens for fast dynamic follow-up (ranging from 0.5 to 1 s) of *e.g.*, consistency profiles (Karlström and Hill 2017a,b,c). In Fig. 3, a snapshot of the consistency profile is given. It is essential to control the shape of the profile by manipulating the added dilution water to the refining zones. Thereby, the process inside the refining zones becomes observable and controllable, which is essential when controlling both the flat zone and the conical zone simultaneously (Karlström *et al.* 2018b).



**Fig. 3.** Consistency profile for one sample. The outlet consistency in the blow-line (last sample) has been included. Note that the drastic change in the conical zone is a consequence of the position at which the dilution water is added.

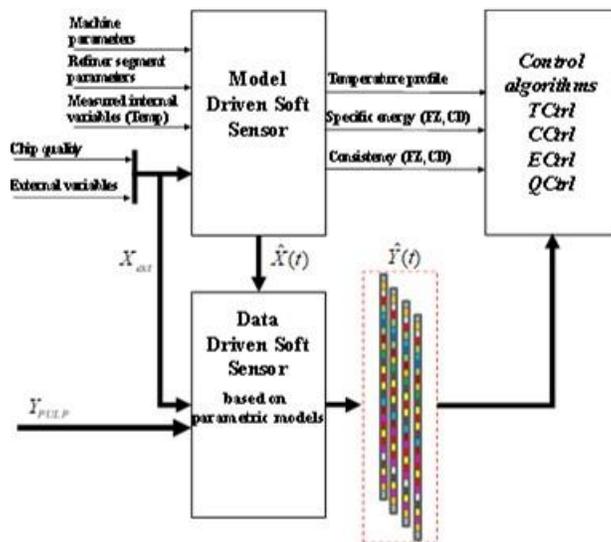
Other important profiles are related to the fiber shear force, steam balance, and the fiber residence time in FZ and CD, which affect the final pulp properties as well. However, the pulp properties are still a challenge to derive only by using physical models, which calls for additional efforts such as the use of empirical models (Karlström *et al.* 2022).

The non-linear physical model, which can be classified as a model-driven soft sensor, is preferably combined with empirical models. These models are often classified as data-driven soft sensors according to the Grey-box terminology presented by Bohlin (2006), as shown in Fig. 4.

The output data of interest in this article is,

$$\hat{Y} = [\hat{Y}_{\text{internal}}, \hat{Y}_{\text{PULP}}] \quad (1)$$

where the internal variables can be derived from the model-driven soft sensor,  $\hat{Y}_{\text{internal}} = \hat{X}$ , while  $\hat{Y}_{\text{PULP}}$  comes from the data-driven soft sensor (Karlström *et al.* 2022).



**Fig. 4.** The concept includes a non-linear model based on first principles (model-driven soft sensor) and the data-driven soft sensor, which can be e.g. pulp and handsheet models. The idea is to use the outputs for control purposes in *TCtrl*: Temperature control, *CCtrl*: Consistency control, *ECtrl*: Specific energy control and *QCtrl*: Quality control.

### ANN Models

In this paper, the MATLAB Neural Network Toolbox is referenced for basic information about the ANN concept. The overview is written by Demuth and Beale (2004) and describes how the ANN can be trained with a backpropagation algorithm based on the Levenberg-Marquardt algorithm (Levenberg 1944; Marquardt 1963). As such, the network has a good characteristic for solving non-linear multi-dimensional mapping problems, given consistent data and enough neurons in its hidden layer (Hagan and Menhaj 1994; Björck 1996; Kermani *et al.* 2005; Rajesh and Ray 2006).

The hidden layers in the authors' ANN were based on a hyperbolic tangent sigmoid transfer function. To prevent overfitting, five layers were used as an initial setting, although it is possible to expand the size dependent on the size of the input matrix. The hidden neurons for each layer was set to ten according to the MATLAB default setting. In this paper, the focus will not be on the ANN-methodology. Instead, the Levenberg-Marquardt algorithm is used for all cases as it is fast and robust when it comes to the non-linear least squares curve-fitting (Gavin 2020).

### Input Candidates to the Data-Driven Soft Sensor

As indicated in Fig 4, the model-driven soft sensor provides several candidates that can be used as additional inputs to a data-driven soft sensor.

Examples, such as temperature profiles (> 20 variables); consistency profiles (> 20 variables); backward and forward flowing steam (>20 variables); water content profiles (> 20 variables); forces on bars (> 20 variables); and defibration and thermodynamical work in each refining zone (> 4 variables), are just a few internal input candidates. In this paper, the internal variables were represented by the matrix  $\hat{X}$  in Fig. 4. For a more comprehensive overview of other possible candidates, see Karlström and Eriksson (2014a,b,c,d). Note that

the reason for using both consistency and steam as inputs is a consequence of the fact that there were both backward and forward flowing steam in the refining zone.

It is also natural to combine such inputs with traditional external variables like specific energy, dilution water added to FZ and CD, as well as the distance between the segments (also called plate gaps), as visualized in Fig. 1.

The external input candidates are quite reliable as measures and constitute the matrix  $X_{ext}$ , which means that the complete set of independent variables used in the data-driven soft sensor can be represented by Eq. 2:

$$X_{extended} = [X_{ext}, \hat{X}] \quad (2)$$

From a practical perspective, it is impossible to introduce all inputs from  $X_{extended}$  in the ANN modeling procedure. However, some candidates are more relevant than others (Karlström and Hill (2017a,b,c); Bengtsson *et al.* (2019, 2020, 2021)) see below:

External variables ( $X_{ext}$ )

- Specific energy (SpE) or the split into motor load (Mload) and production (Prod);
- Dilution water to FZ and CD (DilwFZ, DilwCD);
- Distance between segments in FZ and CD (GapFZ, GapCD);
- The amount of sawmill chips (Sawmill), which affect *e.g.*, the inlet consistency to FZ.

Internal variables ( $\hat{X}$ )

- Backward flowing steam in FZ (primarily two selected close to the maximum temperature; Steam5, Steam8. The figures correspond to the position of the temperature measurements in Fig. 2 (counted from the bottom of the pictures);
- Consistency out from FZ as well as the consistency near the periphery in CD, (CFZ, CCD);
- Fiber residence time in FZ and CD (RFZ, RCD);
- Temperature profile with a focus on maximum temperature in FZ and CD (TmaxFZ and TmaxCD) and periphery temperature in CD (TCDper).

To handle outliers in the input/output data sets, pre-processing was performed to reject data from process stops and traditional process failures. This is handled by removing all measurements larger than 3 standard deviations from the mean. Other failures, like malfunctions in measurement devices, will be rejected as well.

Another obstacle to handle is the fact that the internal and external variables as inputs can introduce (larger or smaller) process-related dependencies. To analyze such situations and detect multicollinearities the Variance Inflation Factors (VIF) are worth studying.

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

In Eq. 3,  $R_k^2$  represents the coefficient of determination obtained by regressing the  $k^{\text{th}}$  inputs on the remaining inputs, *i.e.*, VIF quantifies how much the variance is inflated, (Karlström *et al.* 2022). At this stage, it is not known whether the dependencies are linear or non-linear but if a standalone model is considered, handling the collinearities is a good initiative to analyze input data in the modeling procedure.

A  $VIF_k = 1$  means that there is no linear correlation between the  $k^{\text{th}}$  input and the other remaining inputs. If  $VIF_k > 4$ , a general rule is that further analysis should be performed, while  $VIF_k > 10$  indicates serious multicollinearities, which may call for further analysis and perhaps a modified set of inputs (Belsley *et al.* 1980). In general, if only external variables are used, then the collinearities are not severe, which will be shown

below. However, when using internal variables, the collinearities occur naturally, meaning that several aspects must be considered. For example, if the aim is to derive “standalone” estimators where the contribution from each independent variable is analyzed, the collinearities in  $X_{extended}$  become cumbersome for all data-driven soft sensors. In contrast, if the aim is to find a data-driven soft sensor that can use information from the model-driven soft sensor, we can accept that the *VIF* between the internal variables can exceed the rule of thumb if the data-driven soft sensor is not used as a standalone estimator.

This paper analyzes whether the external variables in  $X_{ext}$  will do fine as model inputs or if additional information from the internal variables is needed when estimating the outputs  $\hat{Y}$  in Eq. 1.

### Estimation of Internal Variables in Data-Driven Soft Sensors

As declared above, some of the internal variables in Eq. 1, can be worth studying as output candidates as well. This is especially true if the aim is to develop simulation tools for pulp refining processes.

A natural set of internal variables to study in more detail will include the maximum temperatures and the consistencies in FZ and CD. These variables are central when it comes to simulation tools as well as advanced control concepts (Karlström and Hill 2018b). In this paper, the following internal variables are of special interest:

$$\hat{Y}_{int} = [\hat{Y}_{TmaxFZ}, \hat{Y}_{CFZ}, \hat{Y}_{TmaxCD}, \hat{Y}_{CCD}] \quad (4)$$

To minimize the risk for input collinearities in the model, the focus here is placed on the external variables in  $X_{ext}$  as inputs when estimating  $[\hat{Y}_{TmaxFZ}, \hat{Y}_{CFZ}]$ , while  $X_{extended}$  in Eq. 2 can be needed when estimating  $[\hat{Y}_{TmaxCD}, \hat{Y}_{CCD}]$ . The latter statement will be discussed further below.

Normally, the operators use the specific energy as a measure of how the defibration/fibrillation of the chips/fibers is performed. This is, however, a rather rough approach, as the specific energy is dependent on both the production and the motor load *i.e.*, the force distribution related to the steam and fiber-to-bar interaction along the radius of the refining segments (Karlström and Eriksson 2014a,b,c,d). Therefore, there is motivation to analyze both specific energy as well as the split between motor load and production as model inputs when estimating internal variables in Eq. 4.

### Pulp Property Estimation in Data-Driven Soft Sensors

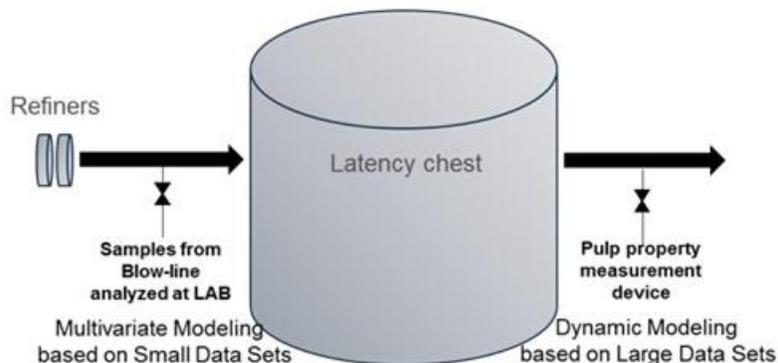
As shown in Fig. 5, small data sets can be sampled before the latency chest, but such sampling procedures are difficult to perform on a regular basis as they are based on tedious laboratory-based measurement procedures. The pulp samples become a set of under-sampled variables, which must be expanded into an oversampled data set of traditional process variables (normally sampled with a fast-sampling rate) to obtain a common time frame (Karlström *et al.* 2016a,b). This undermines the possibility of obtaining enough data to cover the entire operating window over the segment lifetime.

To obtain enough data for modeling purposes, large data sets spread over a long period must be available from measurement devices. For practical reasons it is common that such measurement devices are positioned after the latency chest, see Fig. 5.

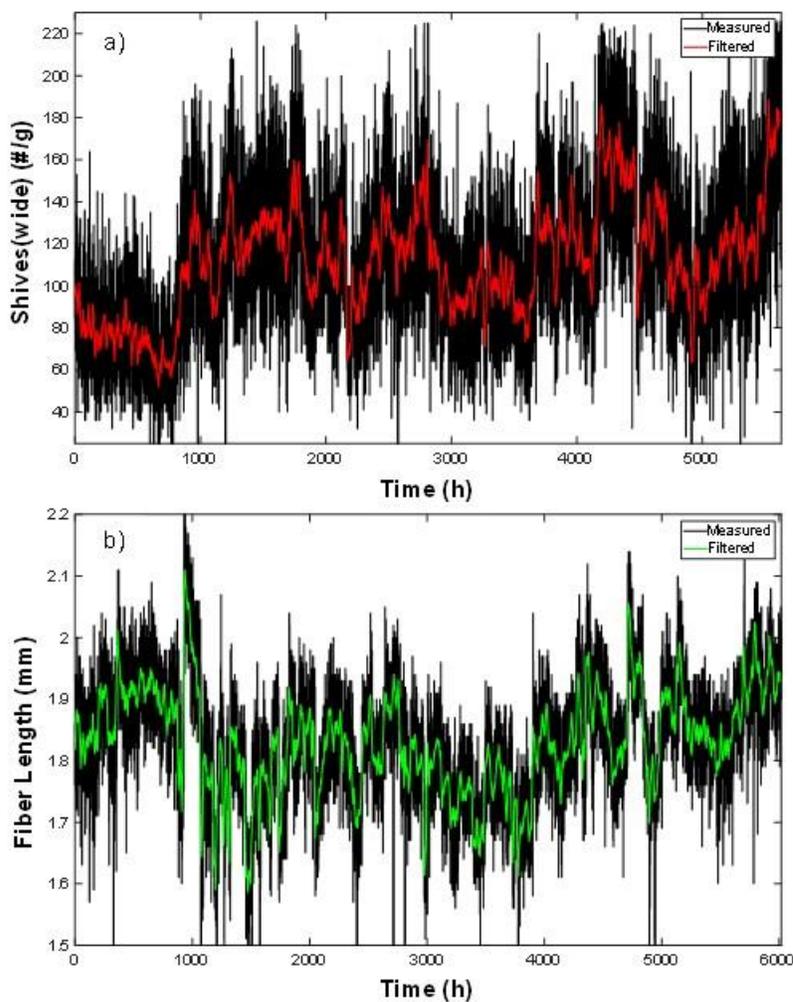
This paper focuses on the estimation of the commonly used pulp properties,

$$\hat{Y}_{PULP} = [\hat{Y}_{Shives}, \hat{Y}_{FL}, \hat{Y}_{Freeness}] \quad (5)$$

where the subscripts “*Shives*, *FL*, and *Freeness*” correspond to shives (wide), fiber length, and freeness. Here freeness is the total volume of water discharged from a side orifice of a specific configuration while the pulp suspension drains freely under gravity.



**Fig. 5.** Schematic drawing of the sampling points for blow-line samples before the latency chest and the pulp property measurements after the latency chest



**Fig. 6.** Measured and filtered pulp properties (a) shives (wide) and b) fiber length) according to a moving average filter based on a horizon of 4000 samples. The sampling rate was 6 s.

Karlström *et al.* (2015a,b, 2022) highlighted several obstacles to overcome when preparing data for modeling purposes. Some of them are related to the absence of relevant signals for long periods (sometimes longer than 3 h). Other complications are associated to the position of the device after the latency chest that introduces a long time constant (20 min). They also indicated that sometimes mixing problems in the chest can be a problem. This makes it hard to guarantee that pulp property variations, caused by changes in the refining conditions, are possible to derive properly.

Finally, it is worth mentioning that the pulp property measurement device often has a non-equidistant sampling rate with a varying interval of 15 to 20 min, which needs to be taken into consideration as well (Sund *et al.* 2021).

In Fig. 6, two of the pulp properties in  $Y_{PULP}$  (see Fig. 4) are included, and it is obvious that the noisy signals must be properly analyzed before using them as outputs. Here, filtering of the signals is most likely required, which can be performed in many ways. In this paper, a moving average filter was used (a low pass FIR filter as seen in Fig. 6), to handle the noisy data. A set of filters [2500, 4000, 8000, 10000] were analyzed. However, the main analysis was performed using 4000 samples as a reference, which corresponds to a time horizon of about 7 h.

### Criteria for the Modeling Procedure

As the Levenberg-Marquardt algorithm can end up finding a local minimum, which is not necessarily the global minimum, it was not expected to get one optimal model. Instead, 200 iterations were run and the best models fulfilling specific criteria were selected for training (*train*) and validation (*val*) sets. Thereby, an ensemble of models constituted the basis when selecting the best structure.

The training of data was primarily performed based on 4200 h (70%) of the time series in Fig. 6, while the rest, 1800 h, was allocated to the validation of the models. Additional information concerning the potential input/output data can be found in Appendix A.

To find a good combination of input candidates, the coefficient of determination was employed, as given as Eq. 6,

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad i = 1, \dots, n \quad (6)$$

where  $\sum(y_i - \hat{y}_i)^2$  represents the sum of the squared residuals from the regression, while  $\sum(y_i - \bar{y})^2$  denotes the sum of the squared differences from the mean of the dependent variable for  $n$  observations (Draper and Smith 1998).

It is expected for the data quality to be relatively unreliable; thus, the thresholds were set to:

$$\left. \begin{array}{l} R_{train}^2 > 0.7 \ \& \ R_{val}^2 > 0.2 \ \text{for } \hat{Y}_{int} \\ R_{train}^2 > 0.7 \ \& \ R_{val}^2 > 0.5 \ \text{for } \hat{Y}_{PULP} \end{array} \right\} \quad (7)$$

The criteria for the validation sets were deliberately set to lower values compared with the training sets.

When it comes to the estimation of the internal variables, the non-linearity effects in both temperature profile and the motor-driven soft sensor (which estimates *CFZ* and *CCD*) can be considerable. This motivates an  $R^2 = 0.2$  for the validation set, while a slightly higher value is expected (0.5) when analyzing the pulp properties.

Because the criteria in Eq. 7 was used as a first requirement to fulfill the thresholds, a selection procedure was introduced based on the weights of the obtained models in each ensemble  $i = 1, \dots, n$  to find a proper estimate. This was performed by using the obtained  $R^2$ -vector according to  $j = 1$ ; the training set,  $j = 2$ ; the validation set or  $j = 3$ ; and the hypotenuse of the two, which is given as Eq. 8:

$$\left. \begin{aligned} \Psi_{ij} &= R_{ij}^2 / \sum_i R_{ij}^2 \\ \hat{Y}_j &= \sum_i \Psi_{ij} \hat{Y}_{ij} \end{aligned} \right\} i = 1, \dots, n ; j = 1, 2, 3 \quad (8)$$

## RESULTS AND DISCUSSION

This section is divided in two subsections covering the results obtained when applying the approach described above to estimate the internal variables  $\hat{Y}_{int}$  in Eq. 4 and the pulp properties  $\hat{Y}_{PULP}$  in Eq. 5. Therefore, the set of independent input candidates interesting to use will differ from each other. Some of the most important input candidates are given as time series in Appendix A, to “assure visually” that the excitation levels in the inputs are large enough to contribute to the expected changes in the selected outputs.

### Estimation of Internal Variables in Data-Driven Soft Sensors

When the internal variables in FZ are analyzed, the external variables will be in focus as inputs, while a mix of external and internal variables will be used when analyzing the internal variables in CD as well as the pulp properties. This is motivated from a causality perspective, as all internal variables in FZ are a consequence of the external variables, while both the external and the internal variables in FZ might affect the predictions in the CD-zone, but not *vice versa*.

To analyze the internal variables in  $\hat{Y}_{int}$  in Eq. 4, the goal is to find an appropriate combination of inputs. According to the combinations in Table 1, the inputs describing FZ, will be based on external variables such as specific energy, motor load, production, the dilution water added, and the amount of sawmill chips to the refiner. As can be seen in Table 1, it is obvious from a *VIF*-perspective that Case B through Case E provides good candidates to the input matrix  $X_{ext}$ , while Case A has been included only to show the colinear effects.

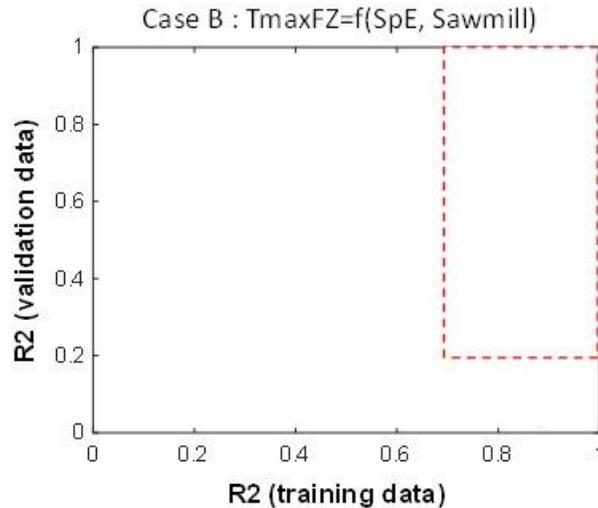
**Table 1.** Variance Inflated Factors for the Different Model Inputs

Variance Inflation Factors						
External Variables						
Case	SpE	Mload	Prod	SawMill	DilwFZ	GapFZ
A	37	49	54	1		
B	1			1		
C		2	2	1		
D		2	3	1	2	
E		2	3	1	2	1

Notes: SpE = Specific energy; Mload = Motor load, Prod = Production, Sawmill = Percentage of sawmill chips to refiner, DilwFZ = dilution water to FZ, GapFZ = Distance between refining segments in FZ

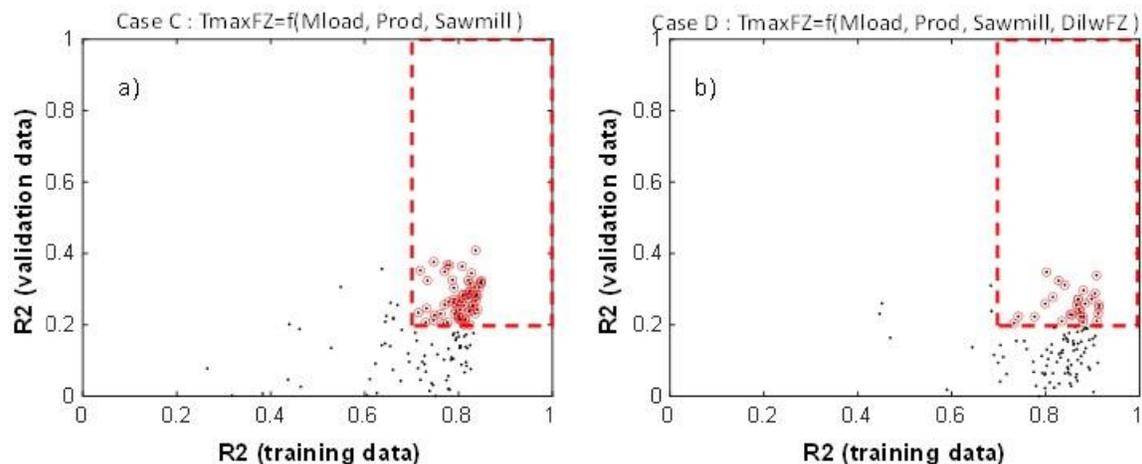
### The Maximum Temperature in FZ ( $T_{\max FZ}$ )

As shown in Fig. 7, which shows the  $R^2$  for the validation data *versus* the  $R^2$  for the training data, the ANN cannot provide good models when using the input data from Case B.



**Fig. 7.** Estimated  $R^2$  for  $T_{\max FZ}$  using the approach based on Case B in Table 1; Dashed lines: Black is required area and red is accepted according to Eq. 7

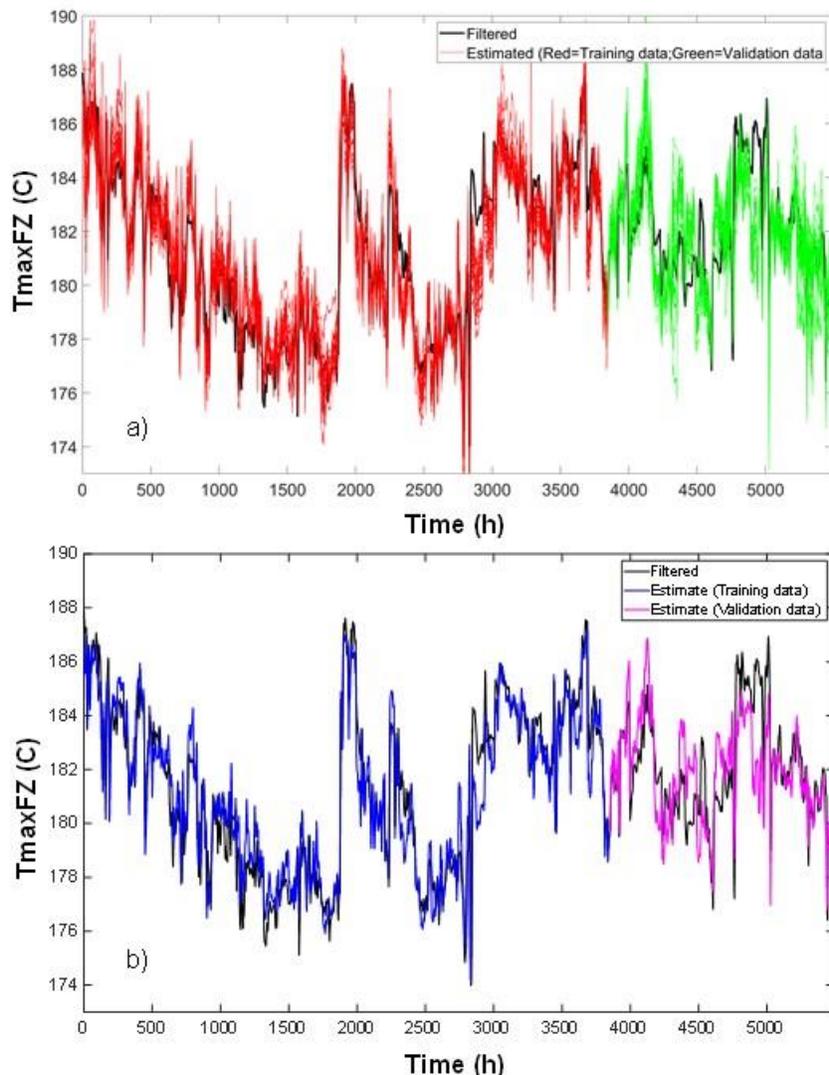
If the motor load and the production (Case C) are used as inputs instead of the specific energy, much better fits are obtained, as shown in Fig. 8. It is also shown that a slightly better estimation will be obtained if the dilution water to FZ (Case D) is included as an input. However, the  $R^2$  value for the validation set was not increased, and the question is whether the response in time domain was short enough.



**Fig. 8.** Estimated  $R^2$  for  $T_{\max FZ}$  using the approach based on a) Case C and b) Case D in Table 1; Dashed lines represent the accepted criteria in Eq. 7

Through studying the ensemble of models in the time domain, the major dynamics seem to have been covered, as shown in Fig. 9a. This is of course promising, and using the criterion  $j = 1$  in Eq. 8, the response became even more pronounced (at this stage we are

not comparing all criteria in Eq. 8, this will be done later in the text), see Fig. 9b. Hence, when estimating  $T_{\max\text{FZ}}$ , the most important result so far is that the ANN provided the best estimates if the motor load and production were included as inputs, instead of using the specific energy as a model input.

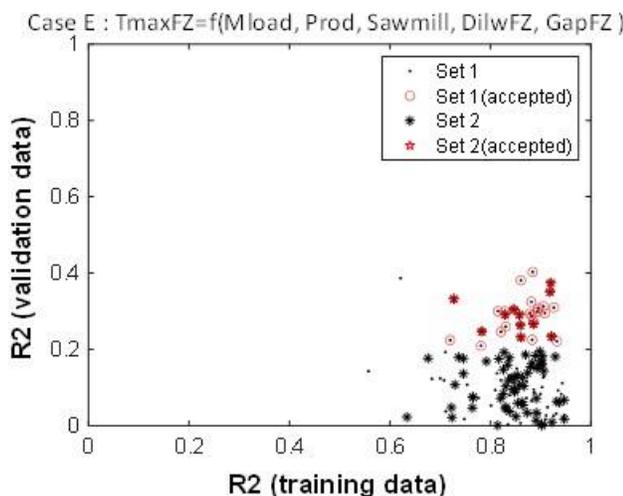


**Fig. 9.** Filtered and estimated  $T_{\max\text{FZ}}$  for Case D in Table 1: a: ensemble of models related to Fig. 8; b: Response using the weighted training set ( $j = 1$  in Eq. 8)

To complete the discussion about possible input combinations from Table 1, Case E in Fig. 10 was also studied. As can be seen, the use of the inputs from Case E provided slightly better models when adding the distance between the segments as an input. This is good news and motivates a proposal to use a set up according to Case E as a reference when estimating the maximum temperature in FZ.

To verify the robustness in the results in Fig. 10, another 200 iterations were also included to find complementary models (Set 2). As shown, the cluster of  $R^2$  will not be improved outside the original region. Instead, the ensemble is located almost in the same interval as the original set (Set 1). This indicates, to some extent, a robustness in the used methodology.

The  $R^2$  value from the validation set was, however, surprisingly low, and this undermines the use of the models derived although the time responses seem to be acceptable.



**Fig. 10.** Estimated  $R^2$  for  $T_{\max FZ}$  using the approach based on two ensemble sets for Case E in Table 1

### The Consistency Modeling in FZ (CFZ)

When it comes to the consistency modeling in the flat zone, none of the input combinations in Table 1 provided acceptable models. This could be perceived as a surprise, but the complex process conditions inside the refining zone are difficult to capture using only the external variables.

As an example, neither the external variables cover phenomena, such as the split between backflowing and forward flowing steam, nor the variable distributed shear forces, where inherent physical conditions must be considered. Hence, although the input matrix was expanded by including the maximum temperature in FZ as well, a good estimation of the consistency was not obtained.

### The Maximum Temperature in CD ( $T_{\max CD}$ )

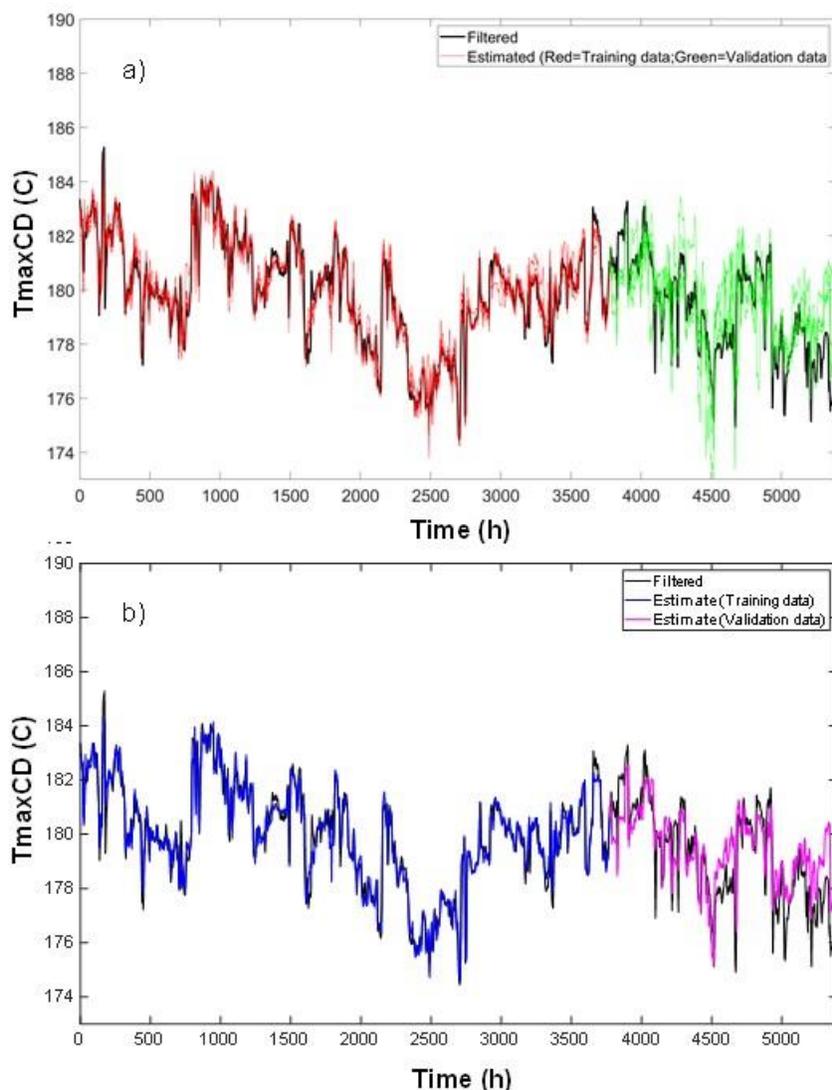
Compared with the modeling of the maximum temperature in FZ (Case E in Table 1), the models for the first temperature measurement in CD (see Fig. 1 (right figure)) need an even more complex input structure.

It is natural to expand the input structure in Case E by introducing the dilution water feed rate together with the distance between the refining segment in CD, but this was not enough, as no models fulfilled the criteria in Eq. 7.

Without going through all the details regarding different combinations and their inability to create good models, it can be concluded that the use of the maximum temperature in FZ as an additional input seems to be important to achieve acceptance according to the criteria in Eq. 7. This indicates that external variables as model inputs are not enough when estimating the maximum temperature in CD. This, of course, is a drawback from an empirical modeling perspective, as temperature measurements and most likely the model-driven soft sensor must be used when estimating internal variables in the conical zone.

Hence, when Case E in Table 1 is expanded by adding the dilution water and plate gap in CD together with the maximum temperature in FZ as inputs, a good estimate of  $T_{\max CD}$  can be provided, as shown in Fig. 11. This is a bit of a surprise, as the position for the maximum temperature in FZ is close to the periphery of the segment and thereby close to the maximum temperature in CD.

However, the correlation coefficient between them was only 0.5. The covariation was thereby unexpectedly small, which can be seen by comparing Fig. 9 and Fig. 11 as well.



**Fig. 11.** Filtered and estimated  $T_{\max CD}$  for Case E in Table 1 added with the dilution water feedrate to CD (DilwCD); the plate gap in CD (GapCD); and the maximum temperature in FZ ( $T_{\max FZ}$ ). Figure a) represents the ensemble models. Figure b) represents the weighted training set ( $j = 1$  in Eq. 8).

### The Consistency Modeling in CD (CCD)

When it comes to the consistency modeling in the conical zone, none of the input combinations mentioned above could be used, which indicates that strong process nonlinearities were present.

In summary, the maximum temperature in FZ can be estimated with quite good accuracy using external variables as inputs, whereas the model-driven soft sensor becomes vital when estimating other internal variables.

### Estimation of Pulp Property in Data-Driven Soft Sensors

When it comes to the modeling of the pulp properties in  $\hat{Y}_{PULP}$ , we have access to all internal variables from the model-driven soft sensor as well as traditional external variables. That is, the entire palette of possible inputs to the ANN models.

Here, the specific energy is used instead of the motor load and the production when estimating pulp properties. This is because pulp properties are dependent on all process conditions represented by both external and internal variables.

Because we are not using the data-driven sensor as a standalone model in this case, a strict methodology regarding the VIF does not need to be followed. However, it is still interesting to see how the collinearities are reflected by the VIF in Table 2.

**Table 2.** Variance Inflated Factors for the Different Model Inputs

Variance Inflation Factors														
Case	SpE	Dil wFZ	Dil wCD	Gap FZ	Gap CD	CFZ	CCD	Steam 5	Steam 8	$T_{\max FZ}$	TCD per	RFZ	RCD	Saw Mill
1	11	7	8	40	194	127	65	143	139	12	2	10133	10318	1
2	1	1	1	1	1									
3	6					42	15	97	122	7	2	79	62	1
4	1	2	1	1	1									1
5	2	2	1	1	1					2	1			1
6	2	2	2	1	1		2			3	1			1
7	2	2	1	1	1	2				3	1			1
8	5	2	2	1	1		2			6	1	4		1

SpE=Specific energy, Gap=Distance between refining segments in FZ and CD, CFZ and CCD=Consistency in FZ and CD, RFZ and RCD= Fiber residence time in FZ and CD, Steam5 and Steam8=Backward flowing steam in 5<sup>th</sup> and 8<sup>th</sup> position in FZ,  $T_{\max FZ}$ = Maximum temperature in FZ, TCDper=Periphery temperature in CD. Sawmill=Percentage of sawmill chips to refiner

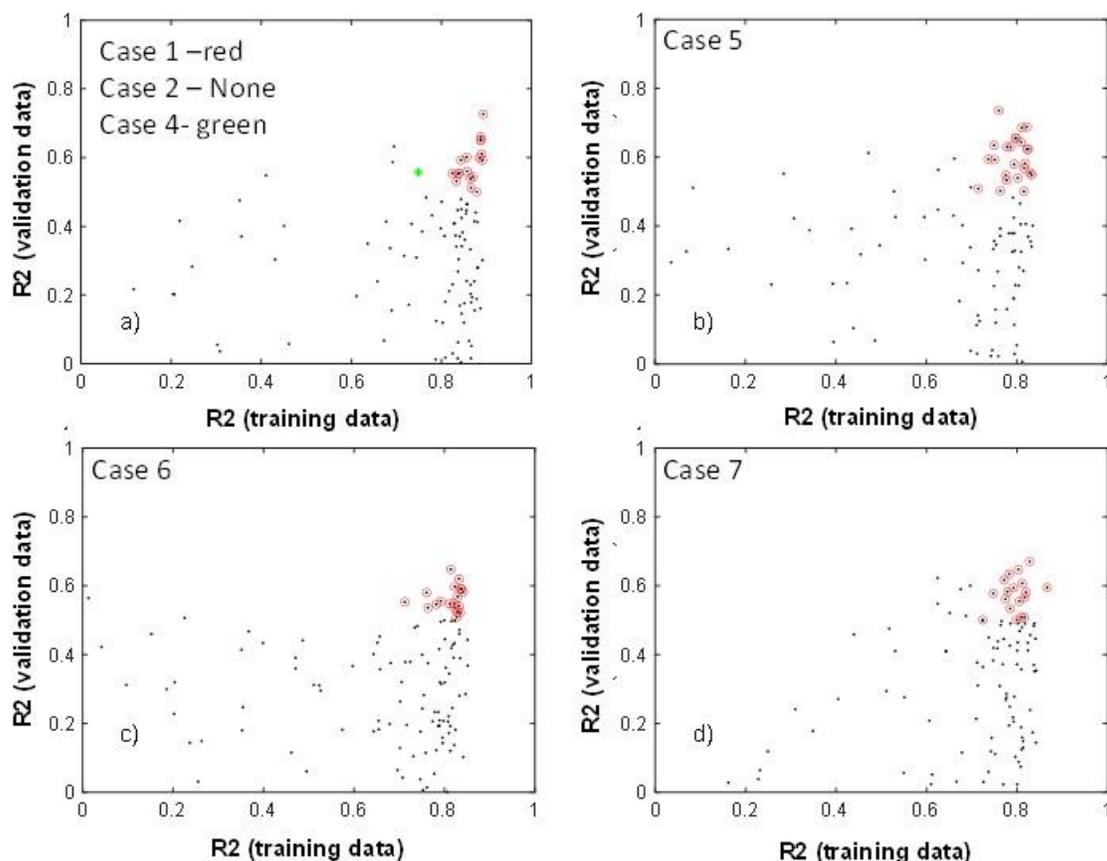
### Shives (Wide)

When estimating shives (wide) using the inputs based on Case 1 in Table 2, the criteria for pulp properties in Eq. 7 will be fulfilled for some of the models, as shown in Fig. 12 a. It is also worth mentioning that the use of just external variables in Case 2 will not result in any acceptable models even though the specific energy is split into motor load and production as inputs.

It is notable to observe that when adding the sawmill chip content as an input (Case 4), one model represented by the green dot in Fig. 12a will meet the criteria in Eq. 7. Therefore, also in this case it is indicated that the sawmill chip content is vital to include as an input in all models.

When reducing the model complexity of Case 1 in Fig. 12, it can be seen that Case 5 through Case 7 seem to provide similar performance as the ensemble of models in Case 1. A closer look at the plots in Fig. 12, shows that the best fits did not exceed an  $R^2 > 0.88$  for the training sets. Through using the procedure described in Eq. 8, the cases can be further compared. As can be seen in Table 3, the training set of Case 1 outperformed Case

5 through Case 7, even though the number of models was smaller (15 against 24). This is, to some extent, expected, as Case 1 can handle more input combinations and thereby include more process non-linearities.



**Fig. 12.** Estimated  $R^2$  for shives (wide). Figure a)– Case 1 (All input candidates included; Red accepted), Case 2 (External inputs only; No models accepted) and Case 4 (External inputs including the sawmill chips content- Green accepted). For the figure b) Case 5 through figure part c) Case 6 and part d) – Case 7, see Table 2 for corresponding input combinations.

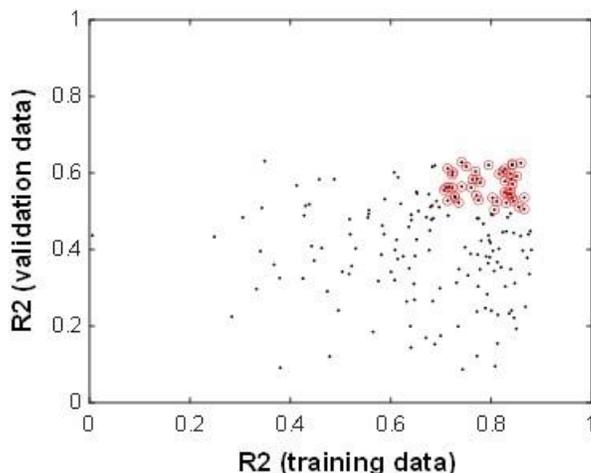
**Table 3.**  $R^2$  for the Training and the Validation Sets Based on a Normalized Weight on the Ensemble of Models from Case 1, and Case 5 Through Case 7

Models based on the ensemble of Case 1			Models based on the ensemble of Case 6		
Weighted by;	$R^2$ train	$R^2$ validation	Weighted by;	$R^2$ train	$R^2$ validation
Train set	0.88	0.71	Train set	0.84	0.66
Validation set	0.87	0.73	Validation set	0.83	0.68
Length*	0.88	0.72	Length*	0.83	0.67
Models based on the ensemble of Case 5			Models based on the ensemble of Case 7		
Weighted by;	$R^2$ train	$R^2$ validation	Weighted by;	$R^2$ train	$R^2$ validation
Train set	0.83	0.72	Train set	0.84	0.67
Validation set	0.82	0.75	Validation set	0.83	0.68
Length*	0.83	0.74	Length*	0.84	0.68

\*) Length of the hypotenuse, where training and validation ensembles are legs ( $R^2$  train and  $R^2$  val.) of the triangle

Note: See Eq. 8, Filter: Window size is 4000 samples

It is also noteworthy that an improvement in accuracy was not achieved, despite the fact that the size of the training set was increased from 70% to 90% of the time series, as shown in Fig. 13. In contrast, the number of accepted models increased drastically from 15 to 41 for Case 1 due to the smaller validation set.

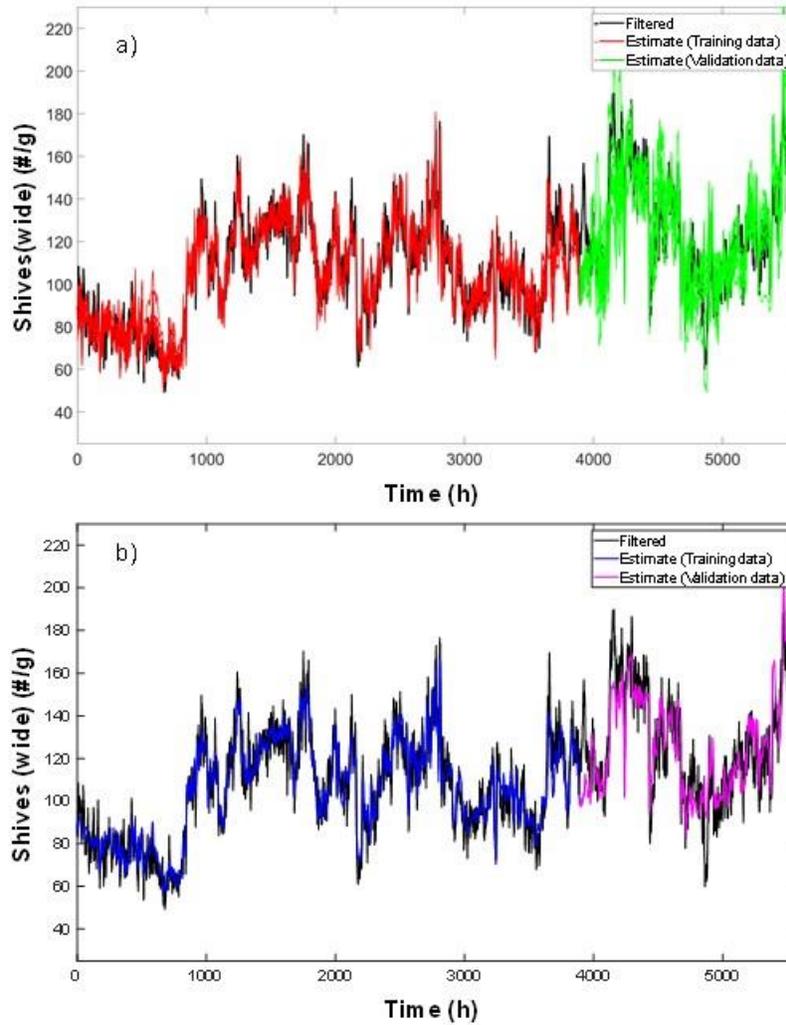


**Fig. 13.** Estimated  $R^2$  for shives (wide) for Case 1 in Table 2 when increasing the training set from 70% to 90% of the time series.

When it comes to the traditional way to interpret *VIF*, it is also worth mentioning that the large values of *VIF* for the fiber residence time in the refining zones (Case 1) seemed to have a limited impact on the robustness of the model responses in the ensembles, as shown in Fig. 14a.

Which model to choose is hard to decide on beforehand, but if the procedure in Eq. 8 is applied to the training set, a quite good response in shives (wide) is obtained, as shown in Fig. 14b. Because of the noisy signal from the pulp measurement unit, the selection of a proper filter is also important. This is shown in Table 4, where  $R^2$  could be further improved by increasing the window size of the moving average filters. However, it is always difficult to analyze proper signal to noise ratios in industrial processes. Therefore, the window size sometimes needs individual modification if other pulp properties are also studied.

The final check was to see whether the accuracy could be improved if the number of hidden layers was increased. As shown in Fig. 15, the  $R^2$  for the training set was slightly increased, while the increase in the validation set was almost negligible. Hence, an increased size of the hidden layer will not necessarily improve the final result, as shown in Table 5.

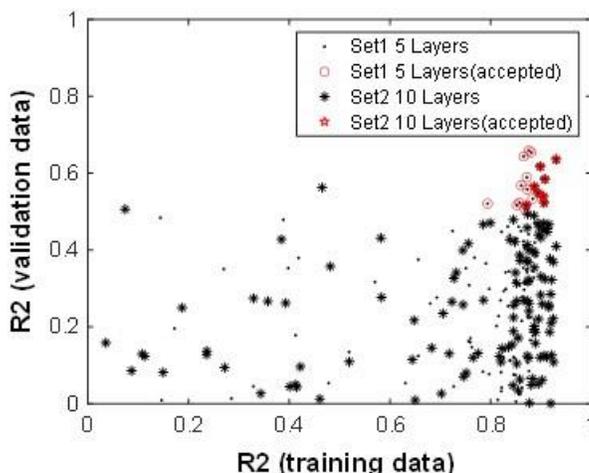


**Fig. 14.** Filtered and estimated shives (wide) for Case 1 in Table 2. a) Responses from an ensemble of the training models fulfilling the criteria in Eq. 7. b) Response using the weighted training set based on the obtained ensemble ( $j = 1$  in Eq. 8)

**Table 4.**  $R^2$  for the Training and the Validation Sets Based on a Normalized Weight on the Ensemble of Models from Case 1 for Different Window Sizes

The Ensemble of Case 1 – Filter 2500			The Ensemble of Case 1 – Filter 8000		
Weighted by;	$R^2$ train	$R^2$ validation	Weighted by;	$R^2$ train	$R^2$ validation
Train set	0.81	0.63	Train set	0.94	0.72
Validation set	0.77	0.66	Validation set	0.93	0.76
Length*	0.80	0.66	Length*	0.94	0.74
The Ensemble of Case 1 – Filter 4000			The Ensemble of Case 1 – Filter 10000		
Weighted by;	$R^2$ train	$R^2$ validation	Weighted by;	$R^2$ train	$R^2$ validation
Train set	0.88	0.71	Train set	0.95	0.74
Validation set	0.87	0.73	Validation set	0.93	0.79
Length*	0.88	0.72	Length*	0.94	0.78

\*) Length of the hypotenuse, where training and validation ensembles are legs ( $R^2$  train and  $R^2$  val.) of the triangle



**Fig. 15.** Estimated  $R^2$  for shives (wide) with two different sets of hidden layers using Case 1 in Table 2.

**Table 5.**  $R^2$  for Training and Validation Sets Based on a Normalized Weight on the Ensemble of Models

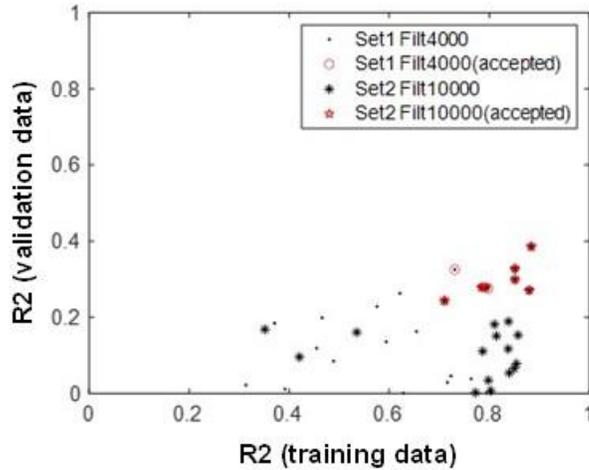
Hidden Layers	Based on Max $R^2$ in Training Set		Based on Max $R^2$ in Validation Set		Based on Max $R^2$ in Weighted Set	
	$R^2$ - Training	$R^2$ - Validation	$R^2$ - Training	$R^2$ - Validation	$R^2$ - Training	$R^2$ - Validation
2	0.78	0.61	0.77	0.60	0.78	0.62
3	0.85	0.62	0.83	0.67	0.85	0.64
4	0.87	0.69	0.85	0.71	0.87	0.71
5	0.88	0.71	0.87	0.73	0.88	0.72
6	0.90	0.73	0.90	0.74	0.90	0.75
7	0.90	0.73	0.89	0.76	0.90	0.73
8	0.90	0.70	0.89	0.72	0.90	0.71
9	0.90	0.76	0.89	0.78	0.90	0.78
10	0.90	0.78	0.89	0.75	0.90	0.74

Note: See Eq. 8 for details; Case 1 is utilized as an example where different sizes of the hidden layers have been applied.

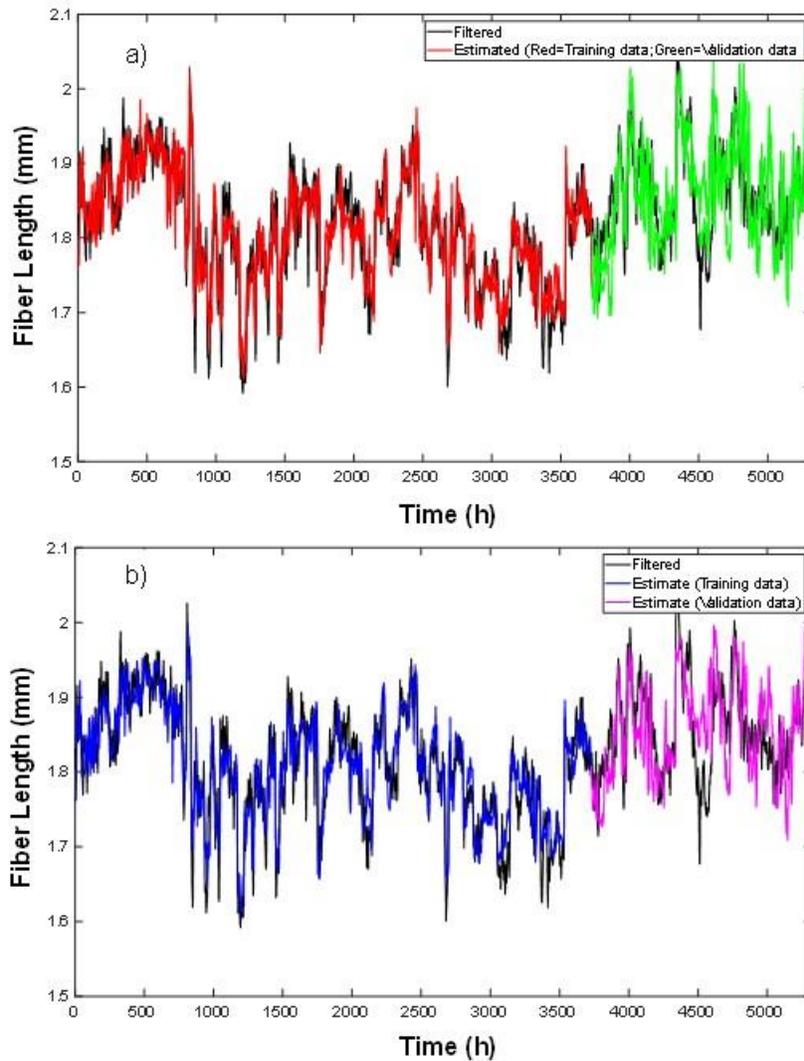
### Fiber Length

It was not possible to reach the threshold for  $R^2_{val}$  in Eq. 7, although the window size of the filter was increased. Therefore, the threshold for the validation set was reduced to 0.2. A few models were obtained that can be analyzed further, as shown in Fig. 16.

An observation worth noting in Fig. 17 is that the response was quite good, although the threshold for the validation set was not reached. This statement is subject to questions, and future studies will show how to proceed.



**Fig. 16.** Estimated  $R^2$  for fiber length using Case 1 in Table 2



**Fig. 17.** Filtered (4000 samples) and estimated fiber length for Case 1 in Table 2. a) Responses from an ensemble of the training models. b) Response using the weighted training set in Eq. 8

## Freeness

Finally, when it comes to the estimation of freeness, it can only be concluded that this measure is not reliable to use as the values of  $R^2$  (both training and validation sets) were almost zero. This is a similar result to earlier works based on an ARX-modeling procedure (Karlström *et al.* 2022).

## CONCLUSIONS

This study investigated how a parametric model structure based on an ANN (using a Levenberg-Marquardt optimization algorithm) can be applied when modeling the maximum temperatures and consistencies inside the refining zones as well as the pulp properties: shives (wide), fiber length, and freeness.

1. When estimating a pulp property, it is concluded that both external and internal variables (derived from the model-driven soft sensor) must be used as model inputs to handle the non-linearities in the process.
2. A deeper analysis of the pulp property shives (wide) showed that the full model represented by Case 1 in Table 2 will suffice as a prime ANN-model structure. With an  $R^2$  of 0.9 (0.7) for the training (validation) sets, this is certainly a candidate when implementing future advanced control concepts.
3. When predicting the pulp property fiber length, it is necessary to modify the thresholds for the validation set. As a good fit of the pulp property shives (wide) was reached, a model overfit is not expected. The backlash, related to the poor results based on the validation set, is of course a drawback. Some improvements can be achieved by introducing another window size of the filter, but a closer look at the response in the time domain provides a reasonably good prediction that might raise some questions about the selected threshold criteria used in this paper.
4. It is also concluded that freeness is not possible to estimate. This has been argued in many articles during the last decade, and the findings in this paper confirm many such statements over the years.
5. When it comes to the estimation of process conditions such as the maximum temperature in the flat zone (FZ) reliable models, based on the training set, can be obtained if the motor load and the production are used as external inputs to the model instead of the specific energy. The drawback, however, is that the  $R^2$  for the validation set was unacceptable small. This also highlights that the use of specific energy or the motor load and production alone did not give enough information about physical conditions inside the refining zones. This has been stated in previous work and can be a consequence of overfitting. It was also shown that some model improvements can be achieved if a) the sawmill chip content, b) the dilution water to FZ, and c) distance between refining segments in FZ are added as external inputs.
6. From the results it can also be concluded that the consistencies in FZ and the CD should not be estimated using an ANN. Instead, a model-driven soft sensor is preferable to use. The same situation arises when estimating the maximum temperature in CD. Most likely, this is a consequence of strong non-linearities in the refining process, making it impossible to capture without using the internal variables.

This might lead to a paradigm shift regarding the need to measure inside the refining zones to get more information about different internal variables to control.

7. Finally, to understand the control concept, it is important to control the CD-refiner using the temperature profile and certainly the maximum temperature, which can be controlled by manipulating the production. At the same time, the consistencies out from each refining zone must be controlled by manipulating the dilution water added to each zone. When it comes to the use of the estimated pulp properties, it is clear from the results in this article that shives (wide) can be controlled by manipulation the consistencies. Presently this is done in a full-scale production line in Sweden.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge the funding of the Swedish Energy Agency and StoraEnso. Special thanks go to the StoraEnso Skoghall mill for running trials and providing the excellent laboratory and process data used in this study.

## REFERENCES CITED

- Bastin, G., and Dochain, D. (1990). *On-line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam, Netherlands.
- Belsley, D. A., Kuh, E., and Welsch, R. E. (1980). *Regression Diagnostics. Identifying Influential Data and Sources of Collinearity*, John Wiley & Sons, Hoboken, NJ, USA.
- Bengtsson, F., Karlström, A., and Hill, J. (2019). “Deep learning in refining processes – A pre-study of internal and external variables impact on pulp properties,” in: *11<sup>th</sup> Fundamental Mechanical Pulp Research Seminar*, Norrköping, Sweden.
- Bengtsson, F., Karlström, A., and Wik, T. (2020). “Modeling of tensile index using uncertain data sets,” *Nord. Pulp Paper Res. J.* 35(2), 231-242. DOI: 10.1515/npprj-2019-0089
- Bengtsson, F., Karlström, A., and Wik, T. (2021). “On the modeling of pulp properties in CTMP processes,” *Nord. Pulp Paper Res. J.* 36(2), 234-248. DOI: 10.1515/npprj-2020-0084
- Berg, D., and Karlström, A. (2005). “Dynamic pressure measurements in full-scale thermomechanical pulp refiners,” in: *International Mechanical Pulping Conference*, Oslo, Norway, pp. 42-49.
- Björck, A. (1996). *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, PA, USA.
- Bohlin, T. (2006). *Practical Grey-box Process Identification - Theory and Applications, Advances in Industrial Control*, Springer-Verlag, London, United Kingdom.
- Bonnans, J. F., Gilbert, J. C., Lemarechal, C., and Sagastizabal, C. A. (2006). *Numerical Optimization: Theoretical and Practical Aspects*, 2<sup>nd</sup> ed, Springer-Verlag, Berlin Heidelberg, Germany.
- Chruy, A. (1997). “Software sensors in bioprocess engineering,” *J. Biotechnol.* 52(3), 193-199. DOI: 10.1016/S0168-1656(96)01644-6

- Demuth, H., and Beale, M. (2004). *Neural Network Toolbox, User's Guide*, The MathWorks Inc., Natick, MA, USA.
- Draper, N. R., and Smith, H. (1998). *Applied Regression Analysis*, 3<sup>rd</sup> Ed., Wiley, New York, NY, USA.
- Eriksson, K. (2005). *An Entropy-based Modeling Approach to Internally Interconnected TMP Refining Processes*, Licentiate Thesis, Chalmers University of Technology, Göteborg, Sweden.
- Eriksson, K. (2009). *Towards Improved Control of TMP Refining Processes*, Ph.D. Thesis, Chalmers University of Technology, Göteborg, Sweden.
- Ferritsius, O. (2021). *Beyond Averages - Some Aspects of How to Describe a Heterogeneous Material. Mechanical Pulp, on Particle Level*, Doctoral Thesis, Mid Sweden University, Sundsvall, Sweden.
- Ferritsius, O., Ferritsius, R., and Rundlöf, M. (2020). "Average fibre length as a measure of the amount of long fibres in mechanical pulps – Ranking of pulps may shift," *Nord. Pulp Paper Res. J.* 33(3), 468-481. DOI: 10.1515/npprj-2018-3058
- Ferritsius, R., Ferritsius, O., Hill, J., Karlström, A., and Eriksson, K. (2018). "TMP properties and process conditions in a CD82 chip refiner at different operation points Part I: Step changes of the process variables, description of the separate tests," *Nord. Pulp Paper Res. J.* 33(1), 82-94. DOI: 10.1515/npprj-2018-3003
- Ferritsius, R., Ferritsius, O., Hill, J., Karlström, A., and Ferritsius, J. (2017). "Process considerations and its demands on TMP property measurements – A study on tensile index," *Nord. Pulp Paper Res. J.* 32(1), 45-53. DOI: 10.3183/npprj-2017-32-01-p045-053
- Gavin, H. P. (2020). *The Levenberg-Marquardt Algorithm for Nonlinear Least Square Curve-Fitting Problems*, Department of Civil and Environmental Engineering, Duke University, NC, USA.
- Hagan, M. T., and Menhaj, M. (1994). "Training feed-forward networks with the Marquardt algorithm," *IEEE Transactions on Neural Networks* 5(6), 989-993. DOI: 10.1109/72.329697
- Hagan, M. T., Demuth, H. B., and Beale, M. H. (1996). *Neural Network Design*, PWS Publishing, Boston, MA, USA.
- Harkonen, E., Huusari, E. and Ravila, P. (1999). "Residence time of fibre in a single disc refiner," in: *International Mechanical Pulping Conference*, Houston, TX, USA, pp. 77-86.
- Harkonen, E., Kortelainen, J., Virtanen, J., and Vuorio, P. (2003). "Fiber development in TMP main line," in: *International Mechanical Pulping conference*, Quebec, Canada, pp. 171-178.
- Huhtanen, J. P. (2004). *Modeling of Fiber Suspension Flows in Refiner and Other Papermaking Processes by Combining Non-Newtonian Fluid Dynamics and Turbulence*, Ph.D. Thesis, Tampere University of Technology, Tampere, Finland.
- Jolliffe, I. T. (2002). *Principal Component Analysis*, 2<sup>nd</sup> Ed., Springer, New York, NY, USA.
- Jos de Assis, A. J., and Filho, R. M. (2000). "Soft sensors development for on-line bioreactor state estimation," *Comp. Chem. Eng.* 24(2), 1099-1103. DOI: 10.1016/S0098-1354(00)00489-0
- Kadlec, P., Gabrys, B., and Strandt, S. (2008). "Data-driven soft sensors in the process industry," *Computers and Chemical Engineering* 33(4), 795-814. DOI: 10.1016/j.compchemeng.2008.12.012

- Karlström, A., and Eriksson, K. (2014a). "Fiber energy efficiency. Part I: Extended entropy model," *Nord. Pulp Paper Res. J.* 29(2), 322-331. DOI: 10.3183/npprj-2014-29-02-p322-331
- Karlström, A., and Eriksson, K. (2014b). "Fiber energy efficiency. Part II: Forces acting on the refiner bars," *Nord. Pulp Paper Res. J.* 29(2), 332-343. DOI: 10.3183/npprj-2014-29-02-p332-343
- Karlström, A., and Eriksson, K. (2014c). "Refining energy efficiency. Part III: Modeling of fiber-to-bar interaction," *Nord. Pulp Paper Res. J.* 29(3), 401-408. DOI: 10.1515/npprj-2018-0019
- Karlström, A., and Eriksson, K. (2014d). "Refining energy efficiency. Part IV: Multi-scale modeling of refining processes," *Nord. Pulp Paper Res. J.* 29(3), 409-417. DOI: 10.3183/npprj-2014-29-03-p409-417
- Karlström, A., and Hill, J. (2017c). "CTMP process optimization. Part III: On the prediction of Scott-bond, z-strength and tensile index," *Nord. Pulp Pap. Res. J.* 32(1), 253-265. DOI: 10.3183/npprj-2017-32-02-p266-279
- Karlström, A., and Hill, J. (2017a). "CTMP process optimization. Part I: Internal and external variables impact on refiner conditions," *Nord. Pulp Pap. Res. J.* 32(1), 35-44. DOI: 10.3183/npprj-2017-32-01-p035-044
- Karlström, A., and Hill, J. (2017b). "CTMP process optimization. Part II: Reliability in pulp and handsheet measurements," *Nord. Pulp Pap. Res. J.* 32(1), 266-279. DOI: 10.3183/npprj-2017-32-02-p253-265
- Karlström, A., and Hill, J. (2018b). "Control strategies for refiners. Part I: Soft sensors for CD-refiner control," *Nord. Pulp Pap. Res. J.* 33(1), 28-43. DOI: 10.1515/npprj-2018-3007
- Karlström, A., Eriksson, K., and Hill, J. (2015a). "Refiner optimization and control. Part IV: Long term follow up of control performance in TMP processes," *Nord. Pulp Paper Res. J.* 30(3), 426-435. DOI: 10.3183/npprj-2015-30-03-p426-435
- Karlström, A., Eriksson, K., Sikter, D., and Gustavsson, M. (2008). "Refining models for control purposes," *Nord. Pulp Paper Res. J.* 23(1), 129-138. DOI: 10.3183/npprj-2008-23-01-p129-138
- Karlström, A., Hill, J., and Johansson, L. (2018a). "An overview of some efforts to understand CD-refiners," in: *Int. Mechanical Pulping Con.*, Trondheim, Norway.
- Karlström, A., Hill, J., and Johansson, L. (2022). "Data-driven soft sensors in refining processes – pulp property estimation and control," in: *International Mechanical Pulping Conference*, Vancouver, BC, Canada, pp. 79-91.
- Karlström, A., Hill, J., Ferritsius, R., and Ferritsius, O. (2015b). "Pulp property development Part I: Interlacing undersampled pulp properties and TMP process data using Piece-wise linear functions," *Nord. Pulp Paper Res. J.* 30(4), 599-608. DOI: 10.3183/npprj-2015-30-04-p599-608
- Karlström, A., Hill, J., Ferritsius, R., and Ferritsius, O. (2016a). "Pulp property development Part II: Process nonlinearities and its influence on pulp property development," *Nord. Pulp Paper Res. J.* 31(2), 287-299. DOI: 10.3183/npprj-2016-31-02-p287-299
- Karlström, A., Hill, J., Ferritsius, R., and Ferritsius, O. (2016b). "Pulp property development. Part III: Fiber residence time and consistency profile impact on specific energy and pulp properties," *Nord. Pulp Paper Res. J.* 31(2), 300-307. DOI: 10.3183/npprj-2016-31-02-p300-307

- Kermani, B. G., Schiffman, S. S., and Nagle, H. G. (2005). "Performance of the Levenberg–Marquardt neural network training method in electronic nose applications," *Science Direct, Sensors and Actuators B: Chemical* 110(1), 13-22. DOI: 10.1016/j.snb.2005.01.008
- Levenberg, K. (1944). "A method for the solution of certain non-linear problems in least squares," *The Quarterly of Applied Mathematics* 2(2), 164-168.
- Ljung, L. (1999). *System Identification: Theory for the User*, 2<sup>nd</sup> edition, Prentice Hall, Hoboken, NJ, USA.
- Marquardt, D. W. (1963). "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics* 11(2), 431-441. DOI: 10.1137/0111030
- Miles, K. B., and May, W. D. (1990). "The flow of pulp in chip refiners," *J. Pulp Pap. Sci.* 16(2), 63-72.
- Miles, K. B., and May, W. D. (1991). "Predicting the performance of a chip refiner: A constitutive approach," in: *International Mechanical Pulping Conference*, Minneapolis, MN, USA, pp. 295-301.
- Principe, J. C., Euliano, N. R., and Lefebvre, W. C. (2000). *Neural and Adaptive Systems*, Wiley, New York, NY, USA.
- Rajesh, K., and Ray, A. K. (2006). "Artificial neural network for solving paper industry problems: A review," *Journal of Science & Research* 65(7), 565-573.
- Sund, J., Sandberg, C., Karlström, A., Thungström, G., and Engstrand, P. (2021). "The effect of process design on refiner pulp quality control performance," *Nord. Pulp Pap. Res. J.* 36(4), 594-607. DOI: 10.1515/npprj-2021-0011
- Wold, S., Sjöström, M., and Eriksson, L. (2001). "PLS-regression: A basic tool of chemometrics," *Chemometr. Intell. Lab. Syst.* 58(2), 109-130. DOI: 10.1016/S0169-7439(01)00155-1

Article submitted: October 20, 2023; Peer review completed: November 27, 2023; Revised version received and accepted: December 5, 2023; Published: December 15, 2023.

DOI: 10.15376/biores.19.1.1030-1057

## APPENDIX A

In Fig. A1, the inputs for Case D in Table 1 are given for the entire unfiltered time series. As seen, the changes in the process values seem to be large enough for modeling purposes. An interesting observation in Fig. A2, however, is that the plate gaps in FZ and CD are noisy and not profoundly changed during the entire time series. In Fig. A3, some additional internal variables, obtained from the model-driven sensor, are given.

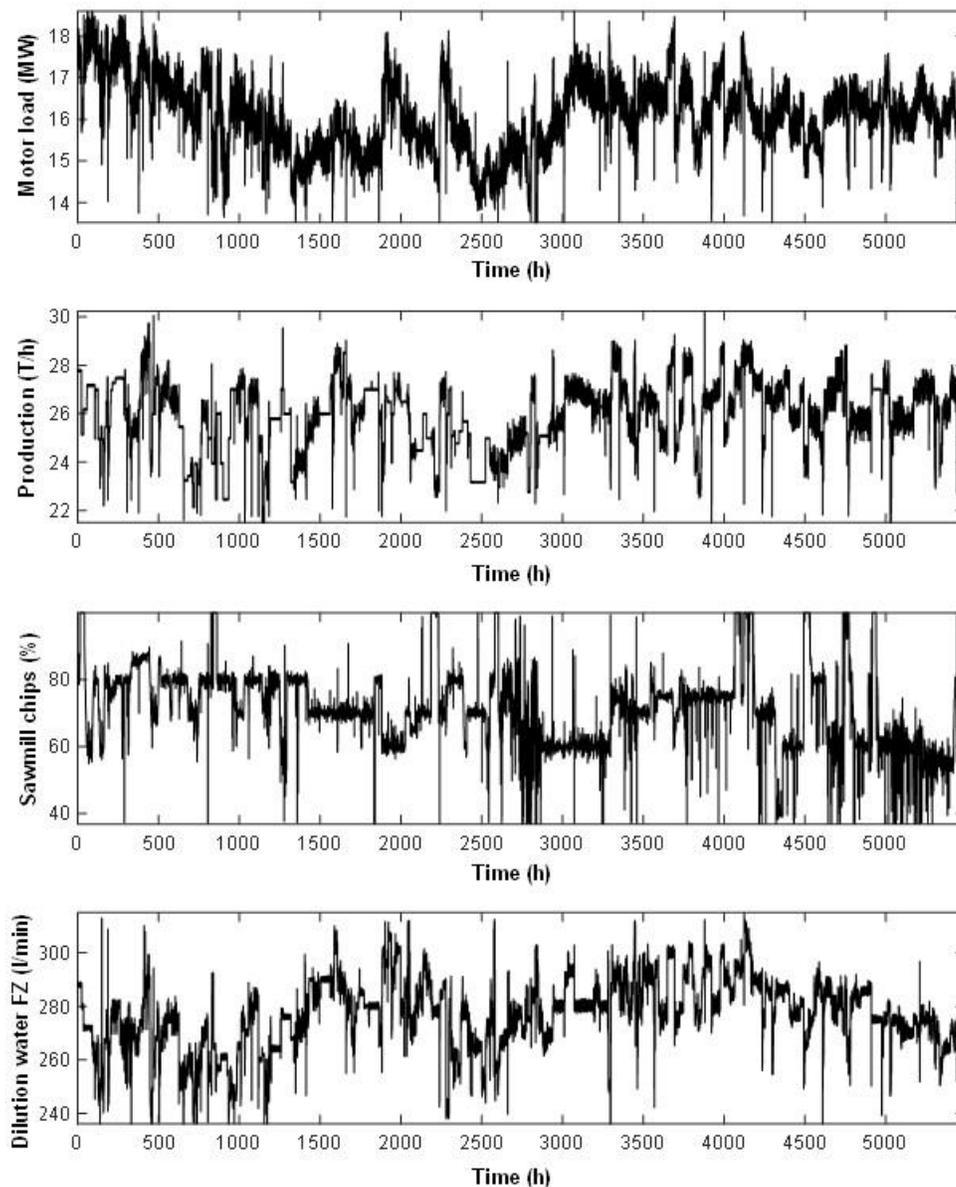
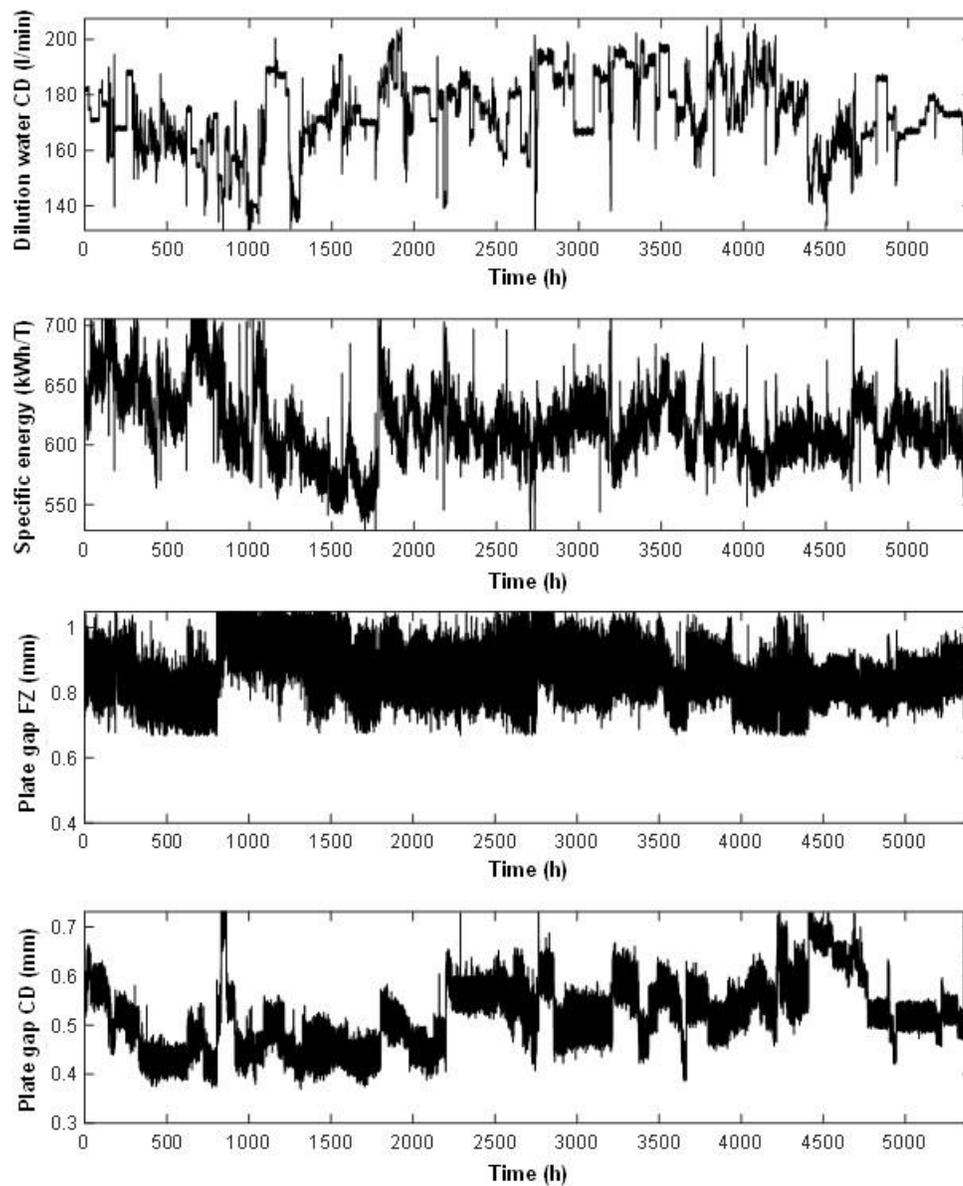
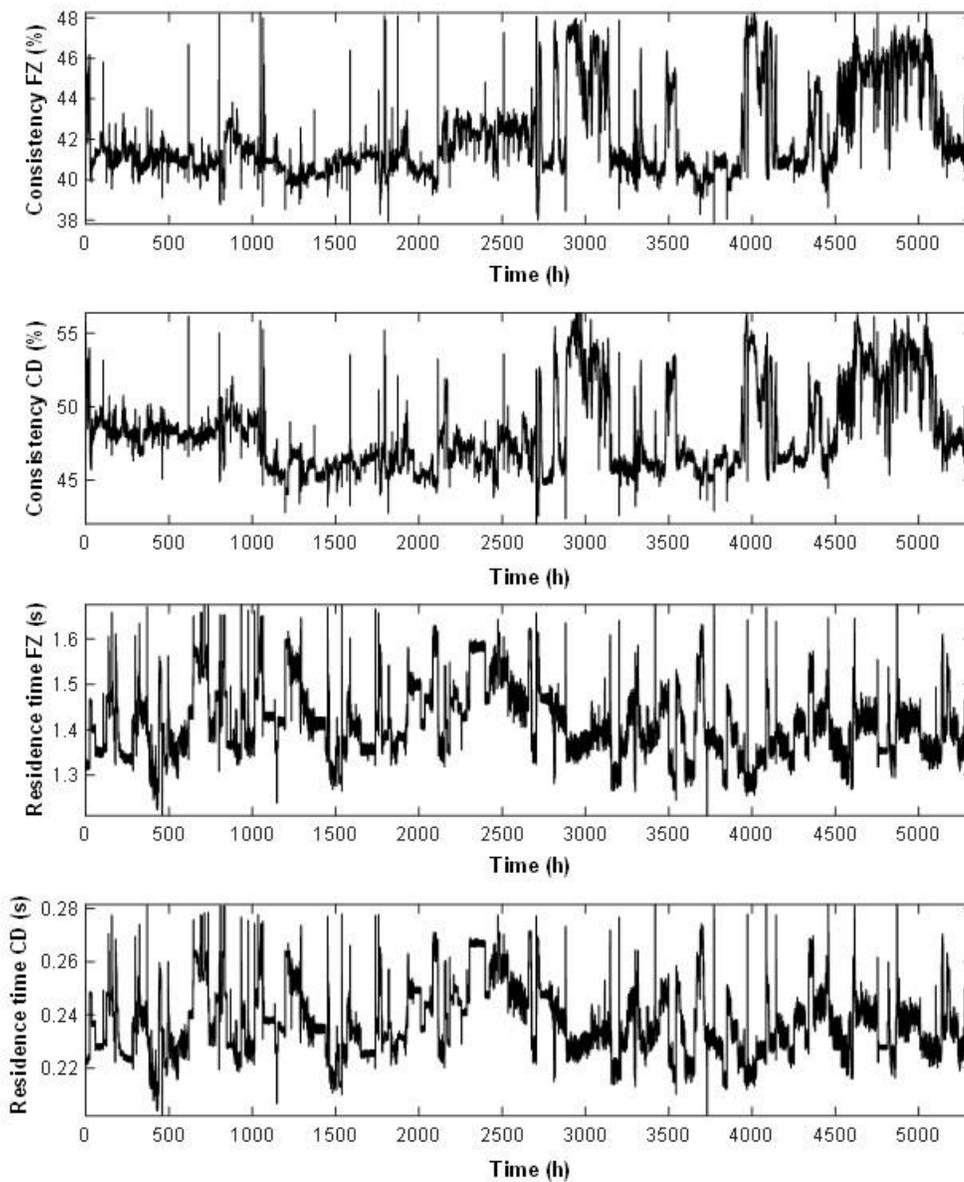


Fig. A1. Unfiltered time series for the ANN-inputs represented by Case D in Table 1



**Fig. A2.** Additional unfiltered ANN-inputs used when estimating the pulp properties



**Fig. A3.** Additional unfiltered internal variables used as inputs when estimating the pulp properties