

Multiple Linear Regression Modelling of Pulp and Handsheet Properties Based on Fiber Morphology Measurements and Process Data

Daniel Ekbåge,^{a,*} Lars Nilsson,^a Helena Håkansson,^a and Ping-I Lin^b

A multiple regression model was evaluated to predict pulp and handsheet properties including z-directional tensile strength (z-strength) and Scott bond values. One hypothesis that was central for the model evaluation was that the crill content, as measured with ultraviolet and infrared lights, would improve the statistical models. A chemi-thermomechanical pulp (CTMP) mill designed with two parallel primary refining lines and a reject refiner was the basis for this study, and all process data and pulp samples were gathered from the specific process. Pulp was extracted from the process for an extended period from a position after the latency chest (primary refined pulp) and from the pulp-stream exiting the mill to the board machine (accept pulp). The crill content was positively correlated to the z-strength of the accept pulp, explaining 55% of the variance with a linear regression model with the crill content as the sole predictor. The estimation model of the z-strength of accept pulp was based on a combination of the crill content, freeness, fibril perimeter for longer fibers, and mean kink angle, and resulted in an R^2 of 0.79. When applying cross-validation to determine the predictive model performance, the highest R^2 obtained was 0.67. This latter model included the crill content, fibril perimeter, and mean kink angle as predictors.

Keywords: CTMP; Fiber morphology; Multiple regression modelling; Handsheet; Z-strength; Scott bond; Crill

Contact information: a: Department of Engineering and Chemical Sciences, Karlstad University, SE-651 88 Karlstad, Sweden; b: Department of Health Sciences, Karlstad University, SE-651 88 Karlstad;

* Corresponding author: daniel.ekbage@kau.se

INTRODUCTION

To gain knowledge of the resulting pulp properties of refined wood chips and relate these to current specifications (required properties), it is crucial to have a sensor collecting the pulp data in an accurate and fast manner. This information can be used to control the refining process. The relationship between the pulp properties and handsheet properties is important for improvements concerning energy efficiency and product quality. For decades there has been a need to fully understand the relationship between fiber characteristics and handsheet or paper properties. Some of the listed literature below deal with bonding and fiber strengths, concepts of load transfer, fiber distributions, impact of fibrillation, and influence of hemicellulose and cellulose compositions. Incorporation of the process data has also been studied, considering multiple temperature measurements in the refining zone for modelling purposes.

Page (1969) developed theories for the tensile strength of paper and postulated that the paper strength was dependent on both the tensile strength of the fibers and the bonds

between the fibers. He developed an equation for the tensile strength of paper based on fiber and network properties. The equation declares that in the case of weakly bonded paper, the strength is controlled by the bonding strength. For papers with increasing strength, the fiber strength becomes important (Page 1969).

Another approach to predict the tensile strength of paper is the shear-lag, as outlined by Carlsson and Lindström (2005). This model is based on a concept of load transfer from the surrounding matrix to the fibers. The axial fiber stress is introduced *via* shear stresses acting at the fiber-matrix interface. If the fibers are short, then they will not reach the ultimate axial stress value before they are pulled out of the matrix. The analysis also indicates the importance of fiber length for weakly bonded sheets and also the relevance of relative bonded area for utilizing the strength potential of fibers.

Reyier Österling (2015) studied the distributions of fiber characteristics and how this information could be used to optimize the mechanical pulping process in terms of reducing the energy usage. The term bonding indicator (BIN) was introduced and defined as the bonding ability influence, which represented the predicted tensile index. If the BIN model is employed to predict the tensile index of the sheet, then it is recommended to use wall-volume-weight data. The results from Reyier Österling (2015) also revealed that the fraction of long fibers was important for the properties of the whole pulp, and that fiber wall thickness and external fibrillation were the fiber characteristics that contributed the most to tensile index of the long fiber fractions in the mechanical pulps investigated. Reyier Österling *et al.* (2012) studied the fiber dimensions impact on the tensile index and density of long fiber laboratory sheets. The results showed that the fibrillation index had a high positive influence on the long fiber tensile index and increased density, while the fiber width, fiber wall thickness, and collapse resistance were negative.

Tensile strength and Scott bond are strongly influenced by internal fibrillation and also external fibrillation (Kang and Paulapuro 2006). Karlström and Hill (2017) outlined a method that made it possible to model the Scott bond, z-strength, and tensile index for chemi-thermomechanical pulp (CTMP). Through the temperature profile measurements in the refining zone, they derived consistency and fiber residence time, and used these parameters as predictor variables in multilinear models. Models based on these types of internal variables outperformed models based on readily measurable process-variables as dilution water and plate clearance.

Through comparing various types of tests to measure the internal bond strength of papers, it has been observed that the z-directional tensile strength is the best parameter to determine the internal bond strength (Koubaa and Koran 1995). In this study the term z-strength is used to label a briefer form of this quantity.

The influence of hemicellulose and cellulose contents of spruce kraft pulp were investigated, and it was found that the Scott bond and z-strength were directly related to the sheet density and thus not influenced by the contents of hemicellulose and cellulose (Molin and Teder 2002).

The traditional method to determine the pulp characteristics involves the preparation of handsheets and the use of physical testing of the sheets. However, recently there has been a rapid development of devices for measuring the properties of fibers and pulp suspensions (Heikkurinen *et al.* 2009).

A widely-used indication of the pulp quality is freeness (Tervaskanto *et al.* 2009), and this property is a measure of the rate at which a dilute suspension may be drained according to TAPPI T227 om-17 (2017). There are also literature reports that deal with the

small particles in the pulp, such as fines and crill particles (and how this property is measured), and online measurements for strength calculations.

According to TAPPI T261 cm-94 (1994), which outlines a standard method to measure the fines content of pulp for paper manufacture, the fines content is measured using a single-screen classifier. The pulp is characterized by dividing its content into particles labelled as ‘fibers’ and ‘fines’. A separation point of these two are the particles that pass through a round hole 76 μm in diameter or a 200-mesh screen as per the TAPPI T261 cm-94 (1994) standard. According to Odabas *et al.* (2016), the fines are a key issue in paper making, and their small size is the main property generating high specific surface area. The aforementioned study proposed the potential use of fines as a control variable to obtain desired paper properties. The fines are often split into fibrils, ray cells, and flake-like fines and the fibrillar material produced in mechanical pulping contributes to higher paper strength; the fine fibrillar material also leads to poor dewatering during the sheet-forming process. The fibrillar material produced during the pulping is also the fraction demonstrating the highest swelling in water (Odabas *et al.* 2016).

The addition of fines improves paper strength properties *via* consolidation between fibers (Bäckström *et al.* 2008). Chinga-Carrasco (2013) studied microfibrillated cellulosic (MFC) materials and considered fibrillation as the disintegration of cellulosic fibers into microfibrils. Through the use of optical techniques, this quantification can potentially be implemented in online production lines for process control (Chinga-Carrasco 2013). Light transmittance is considered a measure of the degree of fibrillation (Iwamoto *et al.* 2008; Chinga-Carrasco 2013). Chinga-Carrasco (2013) exemplified this on MFC films and showed that the light transmittance decreased with increased residual fiber fraction. In other words, a higher amount of fibrillar material present induced higher light transmittance (Chinga-Carrasco 2013). This work also included flowing MFC material through a transparent tube and taking images of the dynamic suspension with a camera system, for the determination of fibrillation degree. The aforementioned experiment confirmed that the light transmittance followed an equal trend as when it was measured with an optical scanner and ultraviolet-visible spectrophotometer.

Osong *et al.* (2014) studied the mechanical properties of handsheets of chemi-thermomechanical pulp (CTMP) mixed with nanolignocellulose and nanocellulose. Two of the conclusions were that the strength properties were increased *via* the addition and that it only had a slight effect in relation to the sheet density.

According to Lundberg *et al.* (2018), there is no formal definition for crill, which relates to the crill method that was outlined as the online measurements of pulp particles in the ultraviolet and infrared spectral regions. However, these particles are supposed to be the smaller fraction of fines (Lundberg *et al.* 2018). The aforementioned study claims that conventional image analysis methods can track changes in the fiber length and width, but not of the crill attached to fibers and/or free in the suspension. Moreover, the fibrillation index measured with a camera is limited by the image resolution, and thus only includes the fibrils attached to the fibers. Within the pulp and paper industry the manufacture of MFC and nanofibrillated cellulose is rapidly growing, and the crill method has been used for characterization (Lundberg *et al.* 2018).

Williamson and Back (2014) stated that important strength properties can be calculated from online measurements, thus contributing to stable pulp quality, quicker grade changes, and reduced laboratory work. They labeled the fine material as crill and described these particles as typically 0.25 μm in width, close to a hundred times thinner

than fibers, and demonstrated that the high surface area associated with the crill is important for the strength properties of the pulp or the paper (Williamson and Back 2014). In a study made by Lindström and Aulin (2014), crill is characterized by very high anisotropy and a typical width of 0.1 μm to 1.0 μm .

In a study by Pettersson (2010), the crill was described as very thin particles that are completely or partially loosened from the fibers and being 100 times thinner than the fibers. The crill is measured *via* the comparison of two optically measured surface areas using light of various wavelengths. It reflects the degree of refining and is useful for troubleshooting, process analysis, monitoring, and control (Pettersson 2010).

The wavelength of the first light is comparable to the crill particle diameter (preferably ultraviolet light), and the second light is a wavelength longer than the first wavelength but shorter than the mean pulp diameter through the pulp suspension (Karlsson and Pettersson 1985).

With reference to the previously outlined literature, this paper adds value by including full-scale process and pulp data extracted for an extensive time period. The novelty is primarily given in the use of the statistical method for feature selection, model validation, combination of data, and measurement using ultraviolet (UV) and infrared (IR) lights for the specific pulp.

Physical Properties of Laboratory Sheets

Through measurement of the physical properties of the laboratory paper sheets, the relationship between the pulp data and the sheet properties can be investigated. In this work, the z-direction tensile strength and Scott bond were selected as responses as they are important for the final product properties. The z-direction tensile strength was obtained using a procedure for measuring the internal fiber bond strength for specimens subject to a normal separating force as per TAPPI T541 om-10 (2010) standard. The Scott bond is the most common test method for the determination of the delamination resistance of paper and board (Fellers *et al.* 2012) and measures the energy required to rapidly delaminate a specimen as per TAPPI T569 pm-00 (2000) standard.

The CTMP is often applied in the middle layer(s) in liquid packaging for providing a product structure with high stiffness, bulk, and good dimensionality (Lindholm *et al.* 2009).

The Specific CTMP Process

A specific CTMP mill designed with two parallel primary refining lines and a reject refiner was the basis for this study, and all of the process data and pulp samples were gathered from the specific process. The pulp samples were extracted from the process for an extended period, collected from a position after the latency chest, and labelled as primary refined pulp (Fig. 1); the pulp sample from the pulp-stream exiting the mill to the board machine was labelled as accept pulp (Fig. 2). The primary refined pulp was refined with a conical disc (CD) refiner followed by a fiber accelerator (driven by an electrical motor) for separation of the steam from the fibers (Fig. 1).

The raw material used for the CTMP process was spruce wood originating from a chip-line with logs and saw-mill chips.

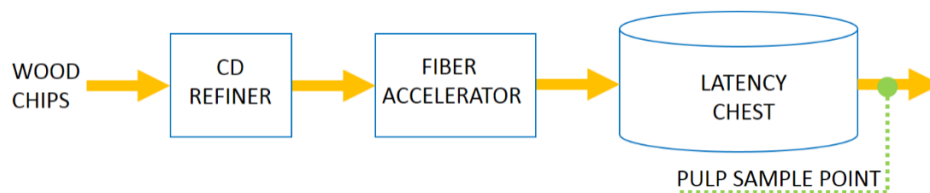


Fig. 1. Simplified process schematic displaying sample point for primary refined pulp

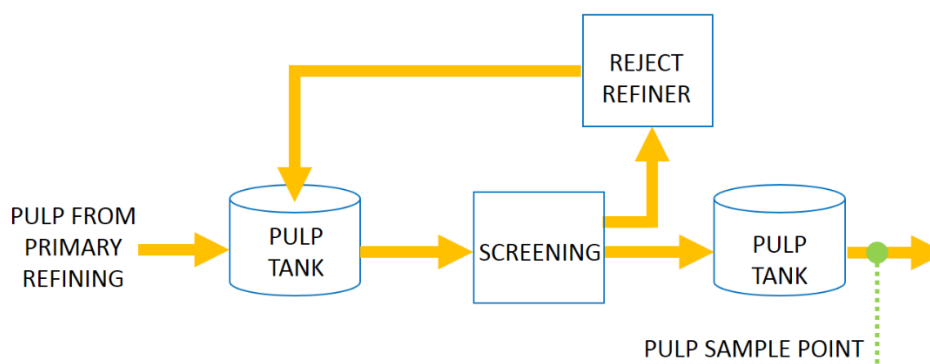


Fig. 2. Simplified process schematic displaying sample point for accept pulp

The primary aim of this study was to evaluate the applicability of a multiple regression model for predicting the z-strength and Scott bond of laboratory handsheets prepared from CTMP extracted at two individual sample points; one representing the primary refined pulp and the second sample representing the accept pulp. This study has included the primary refined pulp to evaluate prediction early in the process, and the accept pulp is important as it is the final pulp mill product delivered to the board machine.

Multiple predictor variables were extracted from the image analysis of pulp samples with the addition of UV and IR lights to include crill content. The predictor variables further include freeness as well as process data for one of the pulps. The goal was to increase the understanding of the relationship between the pulp data and the resulting physical properties of laboratory sheets to support energy efficiency and product quality improvements, with a primary aim to reduce variability.

EXPERIMENTAL

Materials

Pulp preparation and measurements

During an extended period (close to 8 weeks for the total period), the pulp samples were extracted and afterwards stored in a refrigerator. Approximately 15 samples of primary refined pulp and 20 samples of accept pulp were used because they were considered to be representative and within normal mill operating conditions. No additives were added to these samples pulp because the pulp should be representative for the normal conditions. The fiber properties were analysed using an L&W fiber tester plus (ABB AB/Lorentzen & Wettre, Kista, Sweden) equipped with an L&W crill module (hence

labelled 'plus'). The crill-value in this study is measurement of light attenuation. This is done by using a UV-light source with a wavelength of 365 nm and a IR-light source with a wavelength of 850 nm. The light is sent through the pulp suspension, and the small particles will scatter and absorb light from the UV-light to a larger extent compared to the IR light. By comparison, the larger fibers will influence the light coming from the IR-light source more. The crill-value (which is assumed to be related to the content of crill) is calculated based on the comparison of the outgoing light from the UV-light source and the IR-light source (technical documentation, provided by ABB AB/Lorentzen & Wettre). Handsheets were prepared according to a procedure aligned with the methods used at the specific mill laboratory, to mimic the conditions. Preparation of the accept pulp prior to making laboratory sheets was initiated with a disintegrator set to 10,000 revolutions using close to 1300 g of pulp that had been diluted with approximately 2.5 L of non-heated water. This pulp was then diluted with 14 L of water in a container with pneumatic fiber stirring. Approximately 1000 g of the disintegrated pulp suspension was used for preparing a test sheet to determine the dry content. A manually operated sheet former with a circular grid plate and a pneumatic stirrer (model P15490; Messmer Büchel, Gravesend, United Kingdom), for fiber distribution was used for the manufacture of handsheets. The test sheet was dried in an L&W rapid dryer (ABB AB/Lorentzen & Wettre, Kista, Sweden) to determine the dry weight that was used to calculate the sample size from the diluted pulp, with a target grammage of 150 g/m². After sheet forming, the sheets were pressed in two sequences in an L&W sheet press (type 94241; ABB AB/Lorentzen & Wettre, Kista, Sweden). A total of 5 handsheets were placed between circular steel plates and circular paper sheets. The first press sequence was set to create a pressure of 400 kPa in 5 min. The handsheet package was turned over, and new circular sheets were applied, followed by a final pressure sequence for 2 min with equal pressure. The z-strength was measured in a Zwick/Roell testing machine (type z005; Zwick/Roell, Ulm, Germany), equipped with a fixture for the z-directional tensile test of paper that conformed to TAPPI T541 om-10 (2010). Five quadratic pieces from the handsheet were applied in the z-strength fixture. For measurement of the Scott bond, an internal bond tester (model b; GCA/Precision Scientific Group, Chicago, IL, USA) was used and calibrated to operate in the low range (*i.e.*, without the extra weights on the end of the pendulum). Additionally, in this case, 5 pieces from the handsheet were applied.

Freeness was measured according to the Scandinavian Pulp Paper and Board Testing Committee (SCAN C21:65 1965) for all pulp samples. The data from the L&W Fiber Tester Plus included 49 variables describing fiber properties. For the accept pulp, these were the selected predictor variables. In all, 20 observations were made. Because the primary refined pulp samples could be synchronized with refiner process data, the predictor variables in this analysis included L&W Fiber Tester Plus data (49 parameters), process data, and freeness (18 variables) for 15 observations. The process data were synchronized with the sample time of the pulp, taking lag times into account, for example the hydraulic residence time in the latency chest and lag for changes in chip mix (ratio of saw-mill chips to the chips from logs).

Methods

Repeatability test of crill in fiber analyser

To determine the repeatability of the crill measurement, a 10-sample test was made in the L&W fiber tester plus analyser with the pulp originating from one pulp sample.

Repeatability was determined by calculating one standard deviation of the 10 pulp samples. This was calculated for both the primary refined pulp and accept pulp.

Process data

For the study of primary refined pulp, a number of refining process data variables were combined with the L&W fiber tester plus measurements, as input to the statistical analysis. The process data variables were production, plate gap in flat zone, plate gap in CD zone, water dilution to flat zone, water dilution to CD zone, water dilution temperature, water dilution pressure, refiner power, specific electricity consumption, refiner housing pressure, pre-heater pressure, temperature in flat zone, temperature in CD zone, blow line consistency, and the ratio of saw mill chips to the log chips. To enable synchronization between process and pulp data, all process variables were synchronized with the timestamp of the pulp sample.

Modelling techniques

Variable selection in statistical analysis is the key to the identification of important variables in relation to their explanatory effect on a dependent variable. There are several selection techniques available and the one used in this study was Lasso (least absolute shrinkage and selection operator).

The Lasso is a method for estimation in linear models. It minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant, providing the ability to produce coefficients that are exactly zero, which gives interpretable models (Tibshirani 1996).

This technique can be explained starting with the normal linear regression modelling, introducing d as the dependent variable, I as a matrix containing the independent variables, and β as the model coefficients. The following Eq. 1 defines the relationship in a given set of data:

$$d = I\beta \quad (1)$$

For the ordinary least square approach, Eq. 2 outlines the procedure for minimizing the sum of squares of the prediction errors:

$$\min_{\beta} \|d - I\beta\|_2 \quad (2)$$

Regularization is a method within statistical modelling to prevent overfitting. The estimation of this parameter is also known as a shrinkage technique, and is an alternative to variable selection methods (*e.g.*, stepwise regression) and data reduction techniques (*e.g.*, principal component regression) (Finch and Finch 2016).

For the Lasso method, a regularization term is added to Eq. 2,

$$\min_{\beta} \|d - I\beta\|_2 + \lambda \|\beta\|_1 \quad (3)$$

where λ is a parameter that the algorithm locates at the minimum error. As λ changes, different levels of regularization generate the models. Overall, the concept places penalty in retaining many non-zero coefficients and punishes large coefficients.

To summarize, the features characterizing Lasso are: 1) Reducing the number of predictors (by generating zero-valued coefficients); 2) Identifying important predictors; 3) Selecting amongst redundant predictors; and 4) Generating shrinkage estimates with

potentially lower errors than the ordinary least squares method (MATLAB and Statistics and Machine Learning Toolbox Release 2017b, The MathWorks, Inc., Natick, MA, USA).

In this study, the statistical modelling and analysis was performed using MATLAB R2017b (The MathWorks, Inc., Natick, MA, USA) and R version 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria).

Table 1. L&W Fiber Tester Plus Parameters

Fiber Morphology Parameters from L&W Fiber Tester Plus	Fiber Morphology Parameters from L&W Fiber Tester Plus
Mean length (> 0.2 mm)	Fibril perimeter: 0.2 mm to 0.5 mm
Fiber shape (> 0.2 mm)	Fibril perimeter: 0.5 mm to 1.5 mm
Fibril area (> 0.2 mm)	Fibril perimeter: 1.5 mm to 3.0 mm
Coarseness including fines	Fibril perimeter: 3.0 mm to 4.5 mm
Fines length (< 0.2 mm)	Fibril perimeter: 4.5 mm to 7.5 mm
FinesP, 0.1 mm to 0.2 mm, 50 µm to 100 µm, Length share	Calculated coarseness on fibers > 0.2 mm
FinesS, < 0.1 mm < 50 µm, Length share	Fiber width (> 0.2 mm)
FinesP, 0.1 mm to 0.2 mm, 50 µm to 100 µm, Volume share	Fibril perimeter (> 0.2 mm)
FinesS, < 0.1 mm < 50 µm, Volume share	Crill quota MID (Mid-infrared)/IR
FinesP, 0.1 mm to 0.2 mm, 50 µm to 100 µm, Arithmetic	Crill quota UV/IR
FinesS, < 0.1 mm < 50 µm, Arithmetic	Fraction 0.2 mm to 0.5 mm
Width: 0.2 mm to 0.5 mm	Fraction 0.5 mm to 1.5 mm
Width: 0.5 mm to 1.5 mm	Fraction 1.5 mm to 3.0 mm
Width: 1.5 mm to 3.0 mm	Fraction 3.0 mm to 4.5 mm
Width: 3.0 mm to 4.5 mm	Fraction 4.5 mm to 7.5 mm
Width: 4.5 mm to 7.5 mm	Mean kink angle
Shape: 0.2 mm to 0.5 mm	Mean kink index
Shape: 0.5 mm to 1.5 mm	Mean segment length
Shape: 1.5 mm to 3.0 mm	Number of kinks per fiber
Shape: 3.0 mm to 4.5 mm	Number of kinks per mm
Shape: 4.5 mm to 7.5 mm	Number of large kinks per fiber
Fibril area: 0.2 mm to 0.5 mm	Number of large kinks per mm
Fibril area: 0.5 mm to 1.5 mm	Million fibers (> 0.2 mm) per gram
Fibril area: 1.5 mm to 3.0 mm	
Fibril area: 3.0 mm to 4.5 mm	
Fibril area: 4.5 mm to 7.5 mm	

Multiple linear regression is a technique for the estimation of models to fit data from several variables given one response variable. A multiple regression model was defined with n variables, as I_1, I_2, \dots, I_n and the dependent (response) variable D is written as,

$$D = \beta_0 + \beta_1 I_1 + \beta_2 I_2 + \dots + \beta_n I_n + \varepsilon \quad (4)$$

where ε is the residual terms of the model. The quality of the estimation and validation models in this study were determined *via* calculation of the R^2 (coefficient of determination) and the root mean squared error (RMSE).

R^2 is defined as (Mathworks 2017a),

$$R^2 = 1 - SSE/SST \quad (5)$$

$$RMSE = \sqrt{\sum \frac{(y_{\text{pred}} - y)^2}{N}} \quad (6)$$

where R^2 is the coefficient of determination, SSE is the sum of squared error, and SST is the sum of squared total (Mathworks 2017a). The RMSE is the root mean squared error, y_{pred} is the predicted value, y is the actual value, and N is the number of observations. When Eq. 5 was used in the model validation part in MATLAB, the results for a bad model fit can be negative values, indicating model overfitting.

The statistical models were divided into estimated and validated models. An estimated model means that it has not been evaluated against test data, whereas a validated model has been generated with a training data set and then validated with test data set (which was not part of the observations used for the model generation).

Fiber Morphology Parameters from L&W Fiber Tester Plus

The L&W fiber tester plus generated a total of 49 parameters from each pulp sample, mainly describing morphology. Some of this data represent the fiber dimensions, distributions, shape, and crill (Table 1).

RESULTS AND DISCUSSION

Repeatability Test of Crill

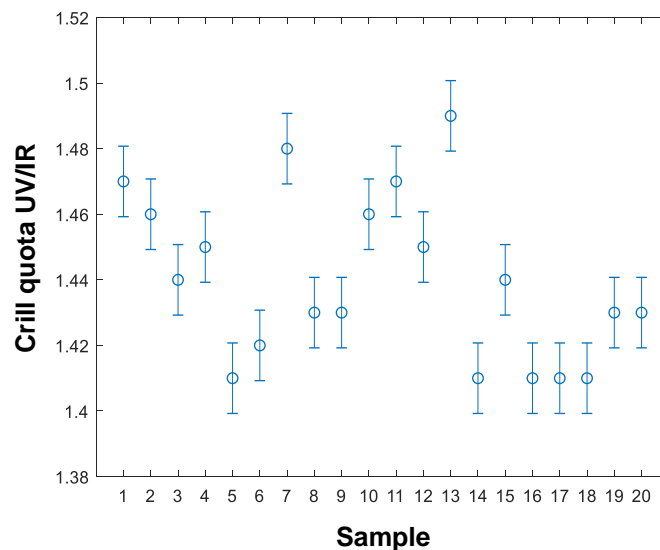


Fig. 3. The 20 observations of crill measurements with error bar calculated from the repeatability test

The standard deviations of the accept pulp and primary refined pulp were 0.011 and 0.018, respectively, revealing a better repeatability for the accept pulp.

Implementing this variability into the 20 observations for the accept pulp revealed the presence of significant differences (Fig. 3).

Physical Properties of Laboratory Handsheets

The results indicated a significant difference in the z-strength when comparing the accept pulp and the primary refined pulp, with a p-value of $3.4e-13$ at the 5% significance level, given a null hypothesis of equal means. The median was 80.2 kPa for accept pulp and 46.6 kPa for primary refined pulp (Fig. 4).

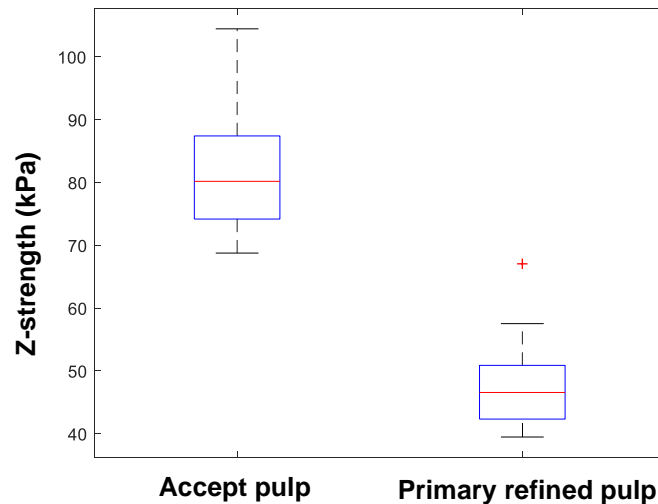


Fig. 4. Boxplots of z-strength for accept pulp and primary refined pulp

The Scott bond (Fig. 5) was also significantly different for the two pulp qualities, indicating a p-value of $1.96e-14$.

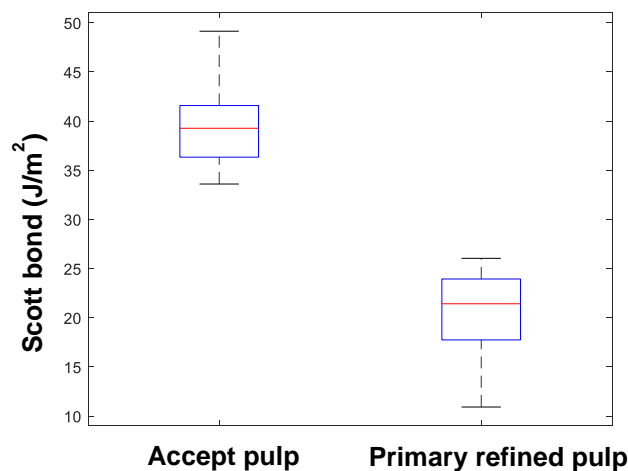


Fig. 5. Boxplot of Scott bond for accept and primary refined pulp

The series of 20 observations of the z-strength of accept pulp (Fig. 6) exposed significant changes for a large amount of the samples and had feasible characteristics for the statistical modelling phase.

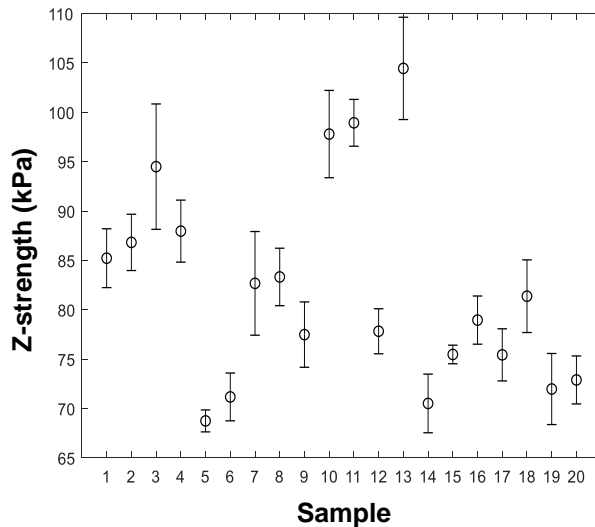


Fig. 6. Z-strength for accept pulp with error bar representing one standard deviation of five measurements

It was observed that the Scott bond measurements did not vary significantly along the series of samples, so that Scott bond was not a part of the statistical modelling procedure. The arguments for not including the Scott bond measurements were that the variations in most of the individual samples were high, which generated difficulties to identify significant process variations in the response variable. Figure 7 shows the Scott bond measurements for the primary refined pulp.

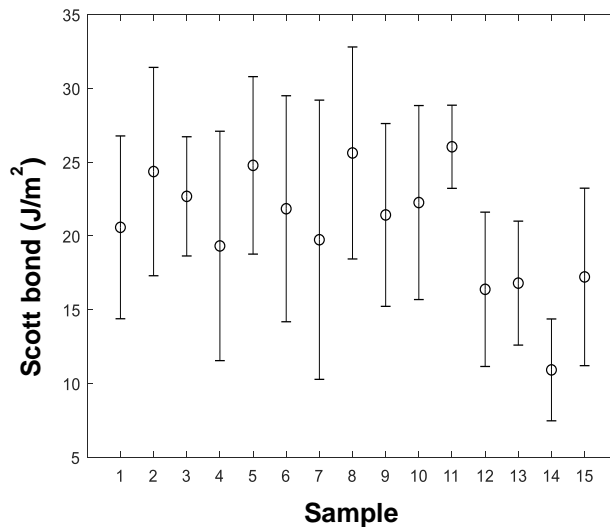


Fig. 7. Scott bond for primary refined pulp, including error bar representing one standard deviation of five measurements

Feature Selection and Model Estimation of Z-Strength for Accept Pulp

Through applying the Lasso regression to the dataset with the z-strength as the dependent variable, the following variables were selected: Canadian Standard Freeness (CSF), fibril perimeter, crill, and mean kink angle. The mean kink angle and CSF had negative correlation coefficients with the z-strength, whereas the fibril perimeter and crill were positively correlated to the z-strength (Fig. 8). CSF had a negative correlation to crill, and fibril perimeter had a positive correlation to crill. These results were found reasonable since it was expected that the amount of fines increases with increased refining, and crill increases with increased fibrillation.

The correlations are displayed with correlograms that provide a specific area with color representing the value of the correlation coefficient. Red color means negative correlation and blue indicates positive correlation, and a higher color intensity designates a higher correlation. A larger area of the square means a higher correlation coefficient was obtained.

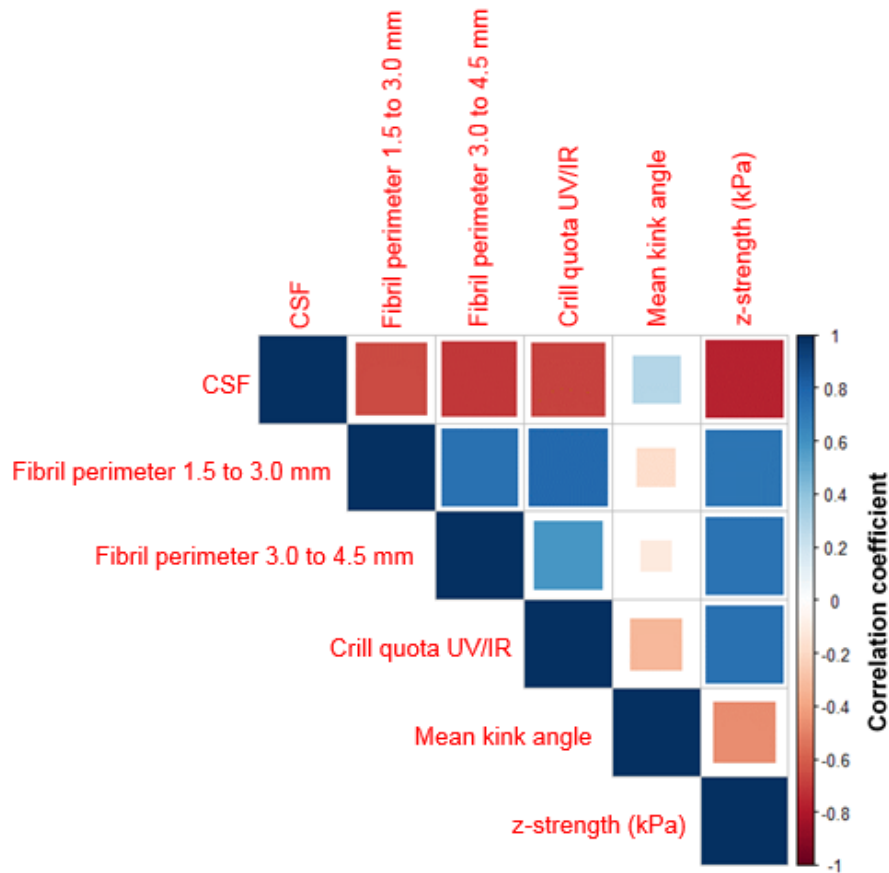
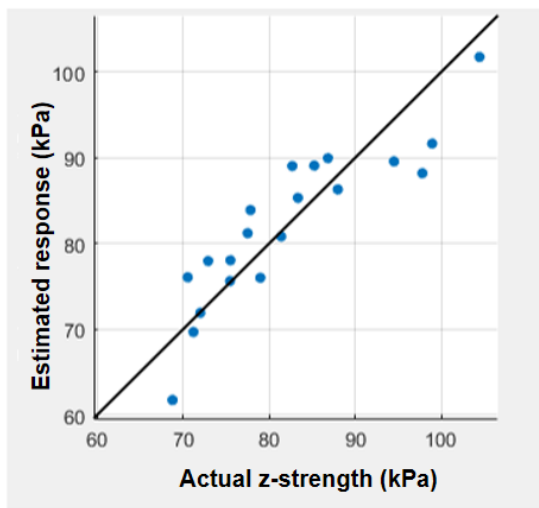


Fig. 8. Correlogram for variables selected with z-strength as dependent variable

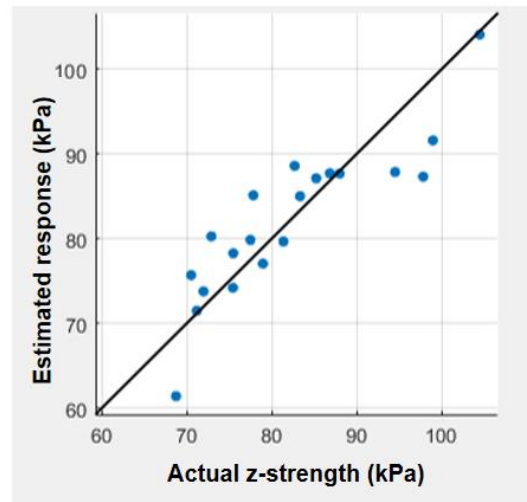
As the number of observations was limited, a multiple regression model including fewer independent variables than the entire list was studied. The model estimation revealed that a model based on fibril perimeter, crill, and kink angle gave an R^2 value of 0.77 (Table 2 and Fig. 9b), indicating that 77% of the variance in the z-strength could be explained by these four predictor variables.

Table 2. Z-strength Estimation Models for Accept Pulp

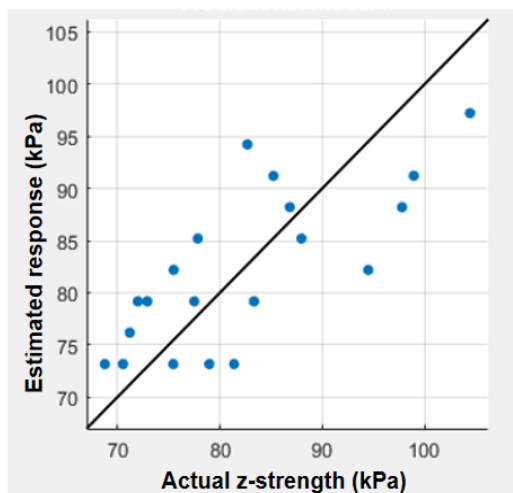
Nr.	CSF	Fibril Perimeter 1.5 to 3.0mm	Fibril Perimeter 3.0 to 4.5mm	Crill Quota UV/IR	Mean Kink Angle	R ²	RMSE
1				x		0.55	6.7
2	x					0.58	6.5
3			x			0.55	6.8
4			x	x		0.69	5.6
5	x		x	x		0.72	5.3
6		x	x	x	x	0.77	4.8
7	x	x	x	x	x	0.79	4.6



(a)



(b)



(c)

Fig. 9. Z-strength estimation models, based on data for accept pulp: (a) includes CSF, fibril perimeter, crill, mean kink angle; (b) includes fibril perimeter, crill, mean kink angle; and (c) crill

The addition of CSF to the model improved the R^2 to 0.79 (Fig. 9a). If only the crill or the CSF variable was present in the model then a $R^2 = 0.55$ (Fig. 9 c) was observed. The R^2 increased with the number of predictor variables and the model errors were decreased.

Model Validation of Z-strength for Accept Pulp

The validation process, based on 5-fold cross validation, indicated that the model including crill, fibril perimeter, and mean kink angle, had the highest R^2 of the validated models at 0.67 (Table 3 and Figs. 10a and 10b).

Table 3. Z-strength Validation Models for Accept Pulp

Nr.	CSF	Fibril Perimeter 1.5 to 3.0mm	Fibril Perimeter 3.0 to 4.5mm	Crill Quota UV/IR	Mean Kink Angle	R^2	RMSE
1				x		0.43	8.1
2	x					0.5	7.6
3			x			0.56	7.1
4			x	x		0.58	7.0
5	x		x	x		0.54	7.3
6		x	x	x	x	0.67	6.1
7	x	x	x	x	x	0.57	7.1

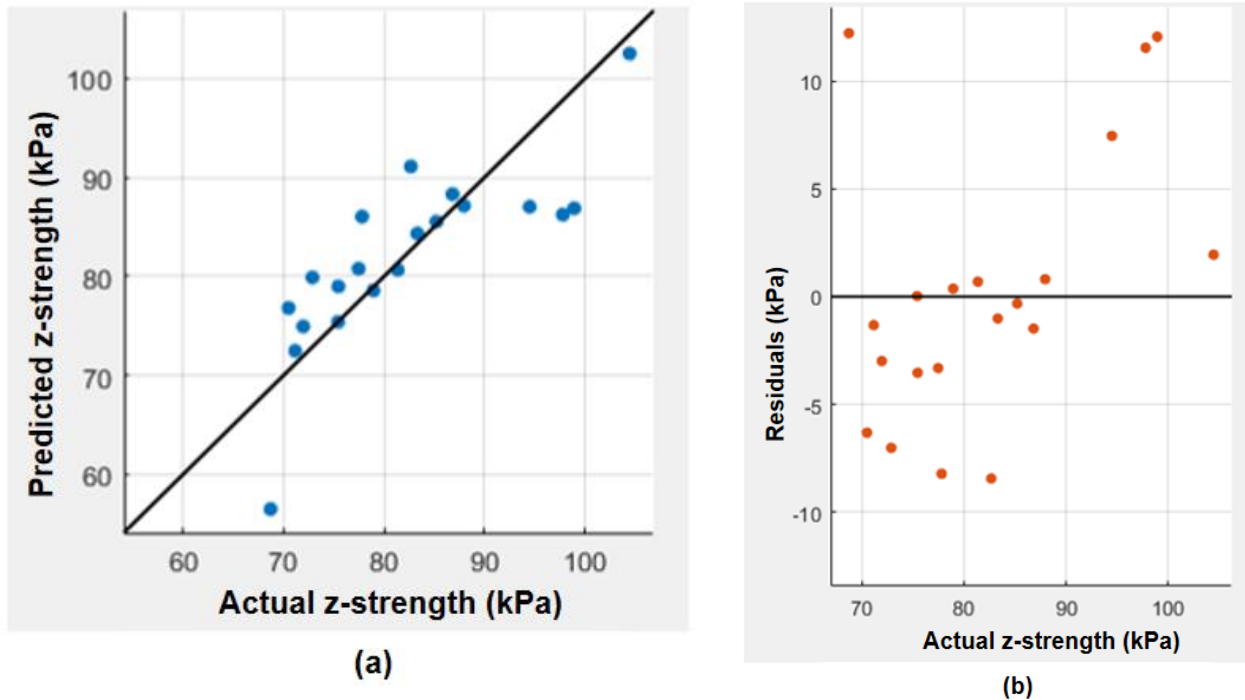


Fig. 10. Z-strength validation model based on data for accept pulp: Fibril perimeter, crill, and mean kink angle, showing the predicted vs. actual values (a) and residuals (b)

If the model contained only the crill variable, 43% of the variance was explained, which can be compared to the estimation model at 55%. It was observed that the addition of CSF to the fiber and crill variables reduced the R^2 value from 0.67 to 0.57.

Based on the tabulated results of the estimation and validation models, it was observed that the RMSE developed differently with the model complexity (number of variables present in the statistical model). The RMSE of the estimation model decreased with increased model complexity over the entire range of predicting variables used, whereas the validation model indicated a minimum RMSE at a model complexity of 4 (Fig. 11), which in this case represented fibril perimeter, crill, and mean kink angle.

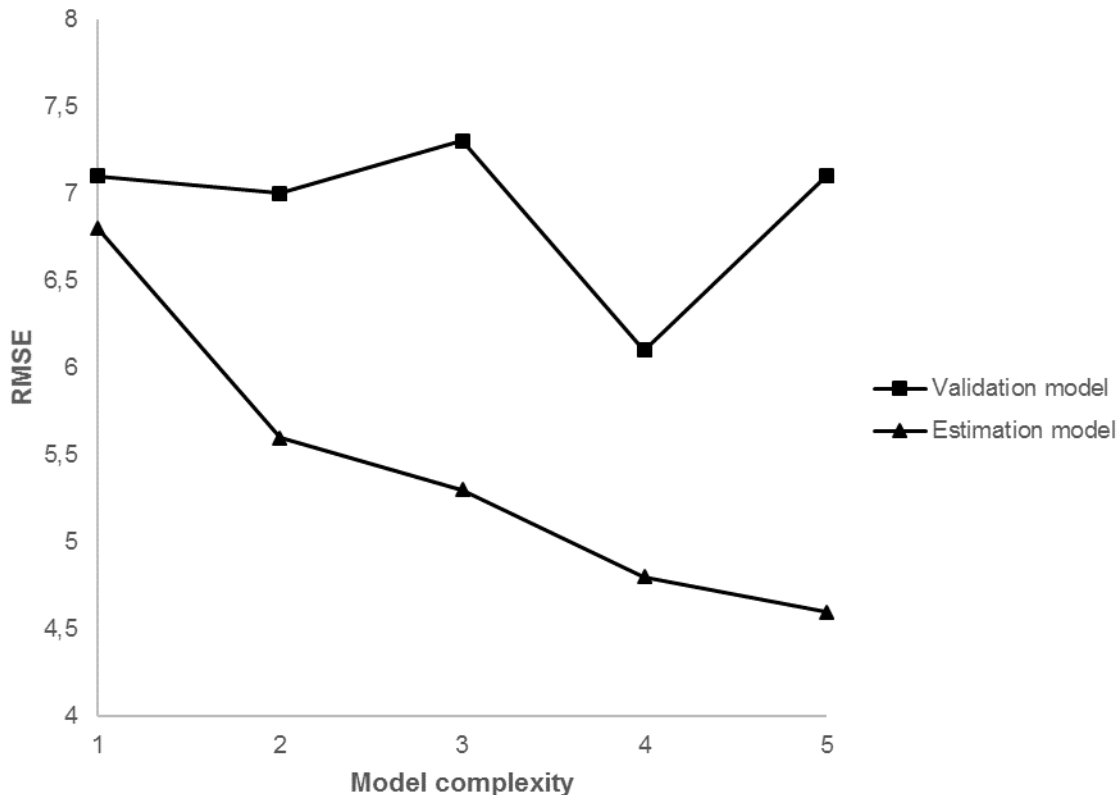


Fig. 11. Model complexity and RMSE for estimation and validation models

Model Estimation and Feature Selection of Z-Strength of Primary Refined Pulp

For the z-strength of primary refined pulp as the dependent variable, Lasso selected plate gaps, blow line consistency, CSF, coarseness, fines, shape, and smaller and larger fractions (Fig. 12). The highest negative correlation coefficient was observed for CSF, and the highest positive correlation was with fraction 0.5 mm to 1.5 mm. The single variable that explained most of the variance was CSF (Table 4), with an R^2 of 0.77. The fraction at 0.5 mm to 1.5 mm was second highest at $R^2 = 0.7$. The process variables only reached an R^2 value of 0.61 and through the addition of CSF this value was increased to 0.91. The R^2 of fiber morphology only reached 0.92. Including all variables in the linear regression model resulted in an R^2 of 0.95 (Table 4).

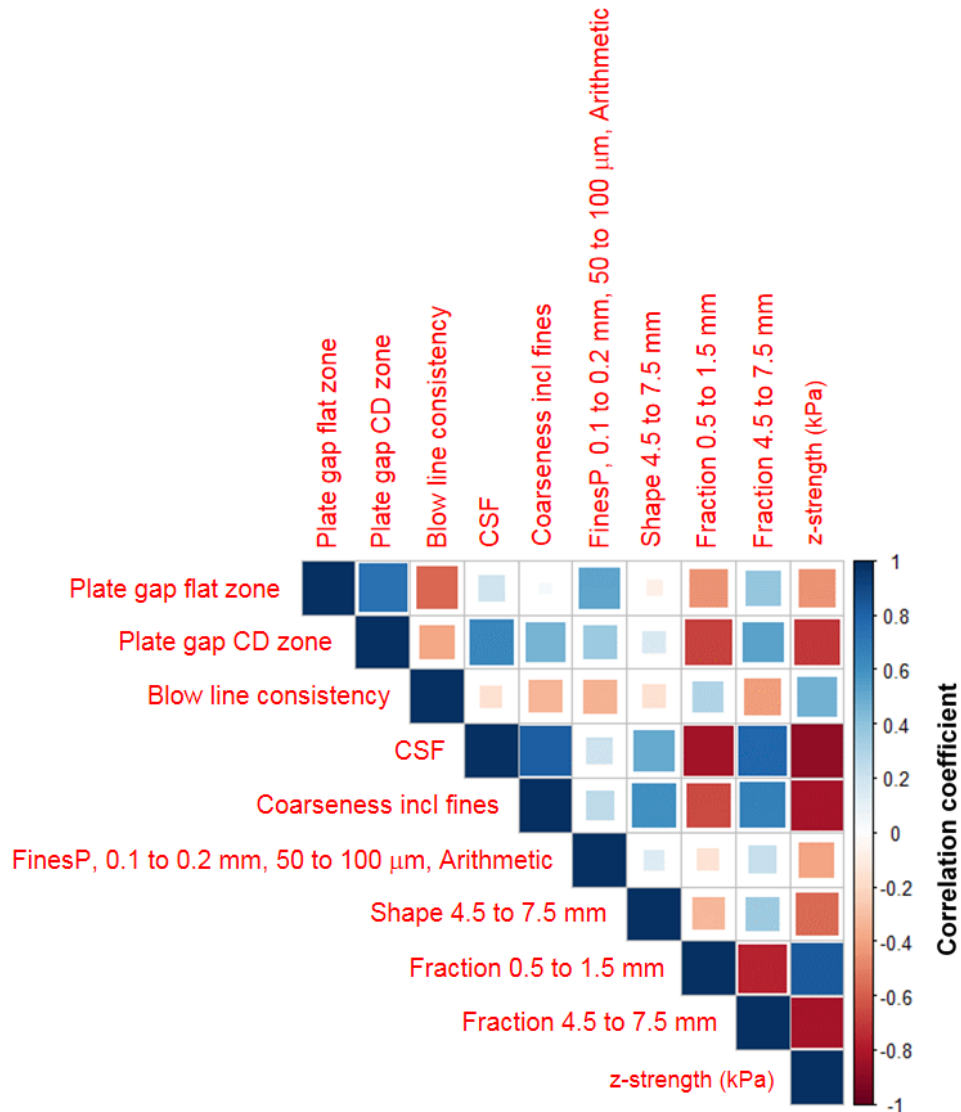


Fig. 12. Correlogram for variables selected having z-strength as dependent variable

Correlation Analysis of Process and Fiber Morphology Data Based on 27 Observations of the Primary Refined Pulp

The correlation study of process and L&W fiber tester plus data based on 27 observations of the primary refined pulp is displayed in Fig. 13. The correlogram indicates that crill content had a negative correlation with the refiner gap and the quota of saw mill chips. One of the highest positive correlations was obtained with refiner power. Dilution water in the CD zone indicated a significant negative correlation to the fiber shape and width, whereas the flow in the flat zone was positively correlated to the fiber shape. The mean length was significantly positively correlated to the plate gap in the CD zone and negative to the refiner power and refiner pressure.

Table 4. Z-strength Estimation Models for Primary Refined Pulp

Nr.	Plate Gap Flat Zone	Plate Gap CD Zone	Blow line Consistency	CSF	Coarseness Incl. Fines	FinesP, 0.1 mm to 0.2 mm, 50 µm to 100 µm, Arithmetic	Shape 4.5-7.5 mm	Fraction 0.5-1.5 mm	Fraction 4.5-7.5 mm	R ²	RMSE
1				x						0.77	3.3
2			x							0.23	6.2
3	x									0.2	6.3
4		x								0.5	5.0
5								x		0.7	3.9
6						x				0.16	6.5
7	x	x	x							0.61	4.4
8	x	x	x	x						0.91	2.2
9					x	x	x	x	x	0.92	2.04
10	x	x	x	x	x	x	x	x	x	0.95	1.6

Model Validation of Z-strength of Primary Refined Pulp

These results varied a lot depending on the predictor variables selected (Table 5). For the validated models, the one with the highest R² included only the fiber morphology and no process variables and had an R² of 0.64. Including the CSF variable only improved the R² to 0.66. Most of the models had very low or negative R² values.

Table 5. Z-Strength Validation Models for Primary Refined Pulp

Nr.	Plate Gap Flat Zone	Plate Gap CD Zone	Blow line Consistency	CSF	Coarseness Incl. fines	FinesP, 0.1 mm to 0.2 mm, 50 µm to 100 µm, Arithmetic	Shape 4.5 to 7.5mm	Fraction 0.5 to 1.5 mm	Fraction 4.5 to 7.5 mm	R ²	RMSE
1				x						0.66	4.6
2			x							0.17	7.1
3	x									-0.41	9.3
4		x								0.1	7.2
5								x		0.48	5.6
6						x				-0.1	8.2
7	x	x	x							-0.01	7.9
8	x	x	x	x						-1.26	11.8
9					x	x	x	x	x	0.64	4.7
10	x	x	x	x	x	x	x	x	x	-3.01	15.7

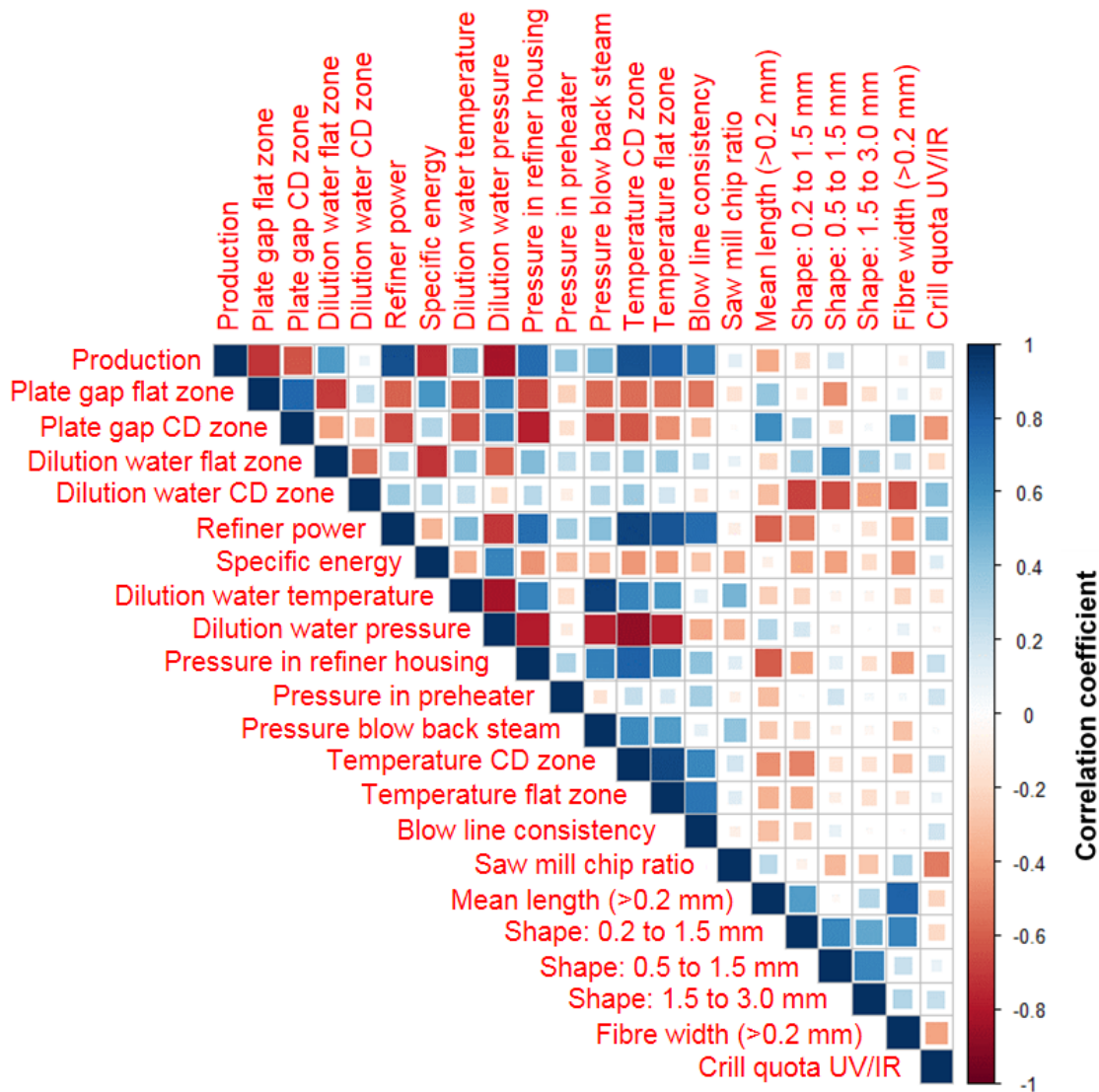


Fig. 13. Correlogram over process and fiber data

Limitations of the study

During the manufacture of the handsheets made from primary refined pulp, some smaller bundles of fibers remained attached to the grid plate. This kind of fiber loss increased the probability of structural variations of the handsheet.

No retention aid was added to the pulp samples after they had been collected at the mill. Fines retention during handsheet forming depend on mesh size. This study used a standard mesh size for sheet forming but the retention was not measured. Tovar, Fischer, Eckhart, and Bauer (2015) observed that fines retention was close to 78% when forming one handsheet using a standard mesh.

One of the limitations in the statistical analysis in this work has been the relatively low number of observations (15 for the primary refined pulp). If this figure could be higher, it would reduce the level of uncertainty in the determination of the quality of the models.

The statistical results in this study depend on specific process and material conditions and should not be interpreted as a generalized outcome for mechanical pulping processes.

When dealing with several independent variables, some of them might be correlated with each other, and this can increase the uncertainty when interpreting the coefficient of determination for the model. The cross-validation procedure took this limitation into account, making use of all data available, which could be compared to hold-out validations that reduce the data available for the model estimation. An observation was that some randomness occurred in the cross-validation procedure, which meant that the model results varied when repeating the analysis. For some models, the statistical estimation model generated a reasonably high coefficient of determination, whilst the cross-validated model of the same predicting variables gave a very low coefficient. This indicated the importance of validating the models to a test-set of data to study the predictive performance.

One of the variables that was correlated with the internal strength of the accept pulp was the fibril perimeter for the longer fiber fractions. It was important to notice that even if a statistical significance was observed, the size of the specific fraction was relevant. In this study the average fraction for the fibers in the interval of 3.0 mm to 4.5 mm was 17%. For the estimation of internal strength from L&W fiber tester plus data, fibrillation and small particles both attached and un-attached to the fibers are important, and all of these contribute to the specific surface area. However, small particles and fibrillation were not the only the properties that were relevant because the fiber geometry was considered through the mean kink angle.

CONCLUSIONS

1. The crill variable (assumed to be related to crill content) was positively correlated to the z-directional tensile strength (z-strength) of accept pulp, which explained 55% of the variance with a linear regression model.
2. Through application of a 5-fold cross-validation procedure, the same predicting variable reached an R^2 (coefficient of determination) of 0.43.
3. The estimation model of the z-strength of accept pulp based on a combination of crill, freeness, fibril perimeter, and mean kink angle resulted in an R^2 of 0.79. Using the same predicting variables with a cross-validation procedure reduced the R^2 to 0.57, but this value was increased to 0.67 through reducing the number of variables by excluding freeness, keeping only fiber morphology variables and crill.
4. Z-strength cross-validation models of primary refined pulp showed large variations in R^2 depending on the predicting variables that were present in the model. Parameters from L&W fiber tester plus gave a better validation model compared to using only refining process data.

ACKNOWLEDGMENTS

This study involved personnel from multiple organisations. The measurements and the research were accomplished in collaboration with the mill, R&D organisation, Karlstad University, and suppliers of measurement systems. Maria Björk and Rickard Wadsborn were company supervisors, and the Board Technology Development team at Stora Enso was involved in constructive discussions. Robert Westberg, Klas Norborg, and the production operators were helpful at the mill. Lisa Edqvist Holt, Pia Eriksson, and Birgitta Gustafsson were helpful in the laboratory. Per-Ivar Fransson at ABB provided valuable support in the fiber analysis work and the writing of this paper. The authors would like to express their appreciation for their valuable support and to acknowledge their contributions.

The study was performed as part of the Industrial Graduate School VIPP (Values Created in Fiber-Based Processes and Products) at Karlstad University, with financial support from the Knowledge Foundation, Sweden, and by Stora Enso.

REFERENCES CITED

- Bäckström, M., Kolar, M.-C., and Htun, M. (2008). "Characterisation of fines from unbleached kraft pulps and their impact on sheet properties," *Holzforschung* 62(5), 546-552. DOI: 10.1515/HF.2008.081
- Carlsson, L. A., and Lindstrom, T. (2005). "A shear-lag approach to the tensile strength of paper," *Composites Science and Technology* 65(2), 183-189. DOI: 10.1016/j.compscitech.2004.06.012
- Chinga-Carrasco, G. (2013). "Optical methods for the quantification of the fibrillation degree of bleached MFC materials," *Micron* 48, 42-48. DOI: 10.1016/j.micron.2013.02.005
- Fellers, C., Östlund, S., and Mäkelä, P. (2012). "Evaluation of the Scott bond test method," *Nordic Pulp and Paper Research Journal* 27(2), 231-236. DOI: 10.3183/NPPRJ-2012-27-02-p231-236
- Finch, W. H., and Finch, M. E. H. (2016). "Regularization methods for fitting linear models with small sample sizes: Fitting the lasso estimator using R," *Practical Assessment, Research and Evaluation* 21(7), 1-13.
- Heikkurinen, A., Leskelä, L., Heinermann, S., and Vehniäinen, A. (2009). "The character and properties of mechanical pulps," in *Mechanical Pulping*, 2nd edition, B. Lönnberg (ed.), Helsinki: Paper Engineers' Association, Helsinki, Finland, pp. 456-514.
- Iwamoto, S., Abe, K., and Yano, H. (2008). "The effect of hemicelluloses on wood pulp nanofibrillation and nanofiber network characteristics," *Biomacromolecules* 9(3), 1022-1026. DOI: 10.1021/bm701157n
- Kang, T., and Paulapuro, H. (2006). "Effect of external fibrillation on paper strength," *Pulp and Paper Canada* (7-8), 51-54.
- Karlsson, H. I., and Pettersson, J. G. T. (1985). "Method of measuring fines in pulp suspensions," U.S. Patent No. 4514257.
- Karlström, A., and Hill, J. (2017). "CTMP process optimization part III: On the prediction of scott-bond, z-strength and tensile index," *Nordic Pulp and Paper Research Journal* 32(2), 266-279. DOI: 10.3183/npprj-2017-32-02-p266-279

- Koubaa, A., and Koran, Z. (1995). "Measure of the internal bond strength of paper/board," *TAPPI Journal* 78(3), 103-111.
- Lindholm, C.-A., Nickull, O., and Pitkänen, M. (2009). "Chemimechanical pulping," in *Mechanical Pulping*, 2nd edition, B. Lönnberg (ed.), Helsinki: Paper Engineers' Association, Helsinki, Finland, pp. 247-281.
- Lindström, T., and Aulin, C. (2014). "Market and technical challenges and opportunities in the area of innovative new materials and composites based on nanocellulosics," *Scandinavian Journal of Forest Research* 29(4), 345-351. DOI: 10.1080/02827581.2014.928365
- Lundberg, M., Norgren, M., and Edlund, H. (2018). "Validation of crill measurements in a high-yield pulp refining process for improved fines material control," *Nordic Pulp and Paper Research Journal* 33(2), 200-209. DOI: 10.1515/npprj-2018-3019.
- Molin, U., and Teder, A. (2002). "Importance of cellulose/hemicellulose-ratio for pulp strength," *Nordic Pulp And Paper Research Journal* 17(1), 14-19. DOI: 10.3183/npprj-2002-17-01-p014-019
- Odabas, N., Henniges, U., Potthast, A., and Rosenau, T. (2016). "Cellulosic fines: Properties and effects," *Progress in Materials Science* 83, 574-594. DOI: 10.1016/j.pmatsci.2016.07.006
- Osong, S. H., Norgren, S., and Engstrand, P. (2014). "Paper strength improvement by inclusion of nano-ligno-cellulose to chemi-thermomechanical pulp," *Nordic Pulp and Paper Research Journal* 29(2), 309-316. DOI: 10.3183/NPPRJ-2014-29-02-p309-316
- Page, D. H. (1969). "A theory for the tensile strength of paper," *TAPPI Journal* 52(4), 674-684.
- Pettersson, T. (2010). "Crill measurements for improved control of refining," in: *Innventia Research Seminar*, Stockholm, Sweden.
- Tovar, R. G., Fischer, W. J., Eckhart, R., and Bauer, W. (2015). "White water recirculation method as a means to evaluate the influence of fines on the properties of handsheets," *BioResources* 10(4), 7242-7251.
- Reyier Österling, S. (2015). *Distributions of Fiber Characteristics as a Tool to Evaluate Mechanical Pulps*, Doctoral Thesis, Mid Sweden University, Sundsvall, Sweden.
- Reyier Österling, S., Ferritsius, O., and Ferritsius, R. (2012). "The influence of fiber dimensions on mechanical pulp long fiber tensile index and density," *Nordic Pulp and Paper Research Journal* 27(5), 844-859. DOI: 10.3183/NPPRJ-2012-27-05-p844-859
- SCAN C21:65 (1965). "Dräneringsförmåga hos massa, bestämd med Canadian Standard Freeness Tester [Drainability of pulp, determined with Canadian Standard Freeness Tester]," Scandinavian Pulp Paper and Board Testing Committee, Stockholm, Sweden.
- TAPPI T227 om-17 (2017). "Freeness of pulp (Canadian standard method)," TAPPI Press, Atlanta, GA, USA.
- TAPPI T261 cm-94 (1994). "Fines fraction of paper stock by wet screening," TAPPI Press, Atlanta, GA, USA.
- TAPPI T541 om-10 (2010). "Internal bond strength of paperboard (z-direction tensile)," TAPPI Press, Atlanta, GA, USA.
- TAPPI T569 pm-00 (2000). "Internal bond strength (Scott type)," TAPPI Press, Atlanta, GA, USA.

- Tervaskanto, M., Ikonen, E., and Ahvenlampi, T. (2009). "Refiner quality control in a ctmp plant," in: *2009 IEEE International Conference on Control and Automation*, Christchurch, New Zealand, pp. 1266-1271. DOI: 10.1109/ICCA.2009.5410443
- Tibshirani, R. (1996). "Regression shrinkage and selection via the Lasso," *Journal of the Royal Statistical Society - Series B (Methodological)* 58(1), 267-288.
- Williamson, M., and Back, S. (2014). "Hairy fibers make strong paper," *Pulp and Paper International* 56(7), 37-39.

Article submitted: April 18, 2019; Peer review completed: September 17, 2019; Revised version received: October 11, 2017; Accepted: October 12, 2019; Published: December 4, 2019.

DOI: 10.15376/biores.15.1.654-676