Talebjedi, B.; Laukkanen, T.; Holmberg, H.; Vakkilainen, E.; Syri, S.

Advanced energy-saving optimization strategy in thermo-mechanical pulping by machine learning approach

Published in:
Nordic Pulp & Paper Research Journal

DOI:
10.1515/npprj-2022-0013

Published: 03/09/2022

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
Talebjedi, B., Laukkanen, T., Holmberg, H., Vakkilainen, E., & Syri, S. (2022). Advanced energy-saving optimization strategy in thermo-mechanical pulping by machine learning approach. *Nordic Pulp & Paper Research Journal*, *37*(3), 434-452. https://doi.org/10.1515/npprj-2022-0013

Mechanical pulping

B. Talebjedi*, T. Laukkanen, H. Holmberg, E. Vakkilainen and S. Syri

Advanced energy-saving optimization strategy in thermo-mechanical pulping by machine learning approach

https://doi.org/10.1515/npprj-2022-0013
Received February 15, 2022; accepted June 13, 2022; previously published online June 22, 2022

Abstract: Thermo-mechanical Pulping (TMP) is one of the most energy-intensive industries where most of the electrical energy is consumed in the refining process. This paper proposes the energy-saving refining optimization strategy by integrating the machine learning algorithm and heuristic optimization method. First, refining specific energy consumption (RSEC) and pulp quality identification models are developed using Artificial Neural Networks. In the second step, the developed identification models are incorporated with the Genetic algorithm to minimize the total refining specific energy consumption while maintaining the same pulp quality. Simulation results prove that a deep multilayer perceptron neural network is a powerful tool for creating refining energy and quality identification models with the model correlation coefficients of 0.97, 0.94, 0.92, and 0.67 for the first-stage RSEC, second-stage RSEC, final pulp fiber length, and freeness prediction, respectively. Findings confirm that the average total RSEC reduction of 14 % is achievable by utilizing the proposed optimization method.

Keywords: artificial neural network; data analysis; forest industry; machine learning; refining energy simulation; thermo-mechanical pulping.

***Corresponding author: B. Talebjedi,** Department of Mechanical Engineering, School of Engineering, Aalto University, Espoo, Finland, e-mail: behnam.talebjedi@aalto.fi
T. Laukkanen, H. Holmberg, S. Syri, Department of Mechanical Engineering, School of Engineering, Aalto University, Espoo, Finland, e-mails: timo.laukkanen@aalto.fi, henrik.holmberg@aalto.fi, sanna.syri@aalto.fi
E. Vakkilainen, Department of Energy, Lappeenranta University of Technology, Lappeenranta, Finland, e-mail: Esa.Vakkilainen@lut.fi

Introduction

Energy efficiency is one of the most critical issues in the energy-intensive industries (Talebjedi and Behbahaninia 2021), such as the pulp and paper sector, which stands as the fourth largest industrial energy user worldwide (International Energy Agency 2007, Oliveira and Almada-lobo 2012). Pulp mills are trying to achieve a more sustainable production process by increasing energy efficiency and reducing energy costs to remain competitive (Hong et al. 2011). 20 to 25 % of the world pulp production is mechanical pulp, and this figure is increasing because of the high yield ratio of the mechanical pulping process and limitation in fiber resources (Bajpai 2018). Refining is the main process in the thermo-mechanical pulp (TMP) mill in which the fiber is treated to achieve the desired quality for the stocks sent to the paper mill. Different approaches to increase refining energy efficiency include optimizing the chip pre-heating condition, improving heat integration, using more appropriate refining plates by optimizing plate patterns geometry, advanced process optimization, and control techniques (Sandberg et al. 2021). This research aims to develop an advanced optimization technique to improve refining energy efficiency and achieve sustainable production. The neural network concept has been employed to create the refining energy and pulp quality identification models to introduce the nonlinear refining behavior. The optimization objective function and constraints are formulated by refining energy and final pulp quality identification models. The optimization is governed by the Genetic algorithm (GA), a subdivision of the heuristic optimization algorithms. The generated refining identification models can be further used to develop an optimal refining control strategy where the optimal setpoints can be achieved using this research's proposed refining energy-saving optimization strategy. Using neural networks to generate a modern process identification model can significantly improve the performance of the control models since an accurate system identification model highly affects the control systems (Perrusquía and Yu 2021).

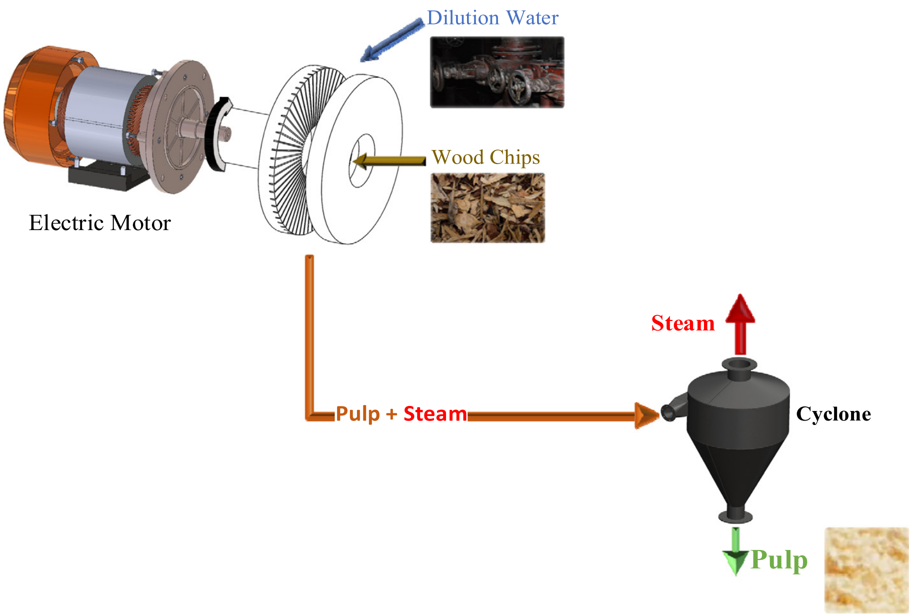

Figure 1: Mechanical refining principle.

All pulping processes aim to separate fibers bonded together by a natural glue called lignin. The fiber separation prepares fibers for the paper-making process. The fibers can be separated either by the chemical or mechanical pulping process. In chemical pulping, fibers are separated by dissolving lignin, where fibers are not being severely damaged. In the mechanical pulping process, fibers are softened and mechanically treated (refined, beaten) by mechanical forces. The typical mechanical pulping process has a yield ratio of 97 % for Norway spruce, while the mentioned ratio is 45 %–50 % in chemical pulping (Bajpai 2016). A higher yield ratio implies that more paper could be produced out of limited raw material resources, which is beneficial from the ecological and national economy point of view. In addition, mechanical pulping is attractive to pulp producers due to its low investment cost and simplicity compared to chemical pulping. However, mechanical pulping also has drawbacks. The mechanical pulping process requires high-quality raw wood input materials, and electricity consumption per unit of pulp production is high. Refining is the most energy-intensive process in thermo-mechanical pulping, which consumes almost 80 % of the mill's electricity. Due to the process's nonlinearity, complex dynamics, and the fact that its operation is affected by many factors, developing an advanced refining optimization strategy to achieve the desired pulp quality with the minimum possible specific energy consumption is challenging. The main factor in developing an optimization strategy is constructing the refining identification model. Since it is not clear how the energy transforms to the pulp in the chip refiner, generating an accurate refining identification model is challenging. For example, it is demanding to estimate the impact of refining disturbance variables such as plate condition and the quality of the input wood chips on the refining process's performance. The quality of the wood chips, including the chip species, dramatically affects the pulping process and end-product characteristics (Li et al. 2011). A number of studies address the seasonal variations in wood chips properties and their effect on final pulp quality and energy-saving opportunities (Browne et al. 2004, Fuhr et al. 1998, Persson and Berntsson 2009). Based on the research conducted by Talebjedi et al. (2021), the effect of disturbance variables in the refining identification model could be considered using artificial neural networks while adding time as a predictor variable. Their research proves a remarkable correlation between the main disturbance variables and time. They suggest considering the time as a predictor variable for the identification model to improve the prediction accuracy. It is clear that the higher the refining identification model's accuracy, the easier it will be to adopt an energy-saving optimization strategy.

Figure 1 shows the refining principle by considering a refiner having a single rotating disc and one stationary disc. Wood chips are fed to the refiner stator and are beaten by plate segments and patterned grooves. The refiner output is pulp mixed by generated steam from moisture content in the wood chips and evaporation of refining dilution water. The pulp is separated from steam by the following cyclone. The generated steam is often utilized to supply the paper machine heating demand in the drying section.

Several statistical regression approaches create refining energy and pulp quality identification models from measured refining data. The general objective of these methods, such as linear regression models, nonlinear regression models, and discriminant analysis (which primarily work by least-squares error or maximum likelihood to measure the quality of an estimator) is to find a mathematical relationship (such as linear or polynomial) between the predictor variables and model outputs. However, these models' efficiency decreases significantly with the increasing number of variables and the complexity of the problem. Harinath et al. (2011) used MATLAB two-stage thermo-mechanical pulping toolbox as their control system identification model. Their prediction model for refining motor load consists of the nonlinear regression model where the deficiencies of this model as refining energy identification model are addressed in Talebjedi et al. (2021). Their research objective is to minimize the refining specific energy consumption while respecting the model constraints, which is the pulp quality limit. Their results show a 12 % reduction in the process-specific energy consumption by integrating economic objectives to nonlinear model predictive control (NMPC) construction. In the other research, Talebjedi et al. (2021) developed a refining identification model by further development of MATLAB Thermo-mechanical pulping simulation toolbox model. Their model is based on theoretical (mechanistic methods) and empirical (resulting from linear and nonlinear regression models) equations.

The artificial neural network (ANN) concept is an alternative statistical approach for developing process identification models and pattern recognition in variable relationships when the hidden pattern is complex and hard to present. Several researches (Wang and Elhag 2007, Jang and Topal 2013, Talebjedi et al. 2020) address the superiority of the neural network to the regression models. These models can be utilized to develop accurate refining identification models in thermo-mechanical pulping. To the authors' best knowledge, Talebjedi et al. (1996) are the only people who analyzed developing refining identification models in thermo-mechanical pulp mill based on the artificial neural network. They conducted a comprehensive study to develop a refining energy identification model in a thermo-mechanical pulp mill using six different machine learning approaches. Their findings prove the high efficiency of the machine learning methods to establish refining energy identification models despite the increased complexity of the refining phenomena. They used the nonlinear regression model discussed in Schwartz et al. (1996) to validate the measured data from a thermo-mechanical pulp mill. Their research proves the machine learning algorithm's superiority to the regression-based models where the integration of the adaptive neuro-fuzzy inference system (ANFIS) and the particle swarm optimization (PSO) algorithm gives the most efficient combination. However, their research could be improved by performing a parametric study and evaluating the effect of the different refining manipulated variables on the refining energy simulation. Similar studies regarding the application of the machine learning approach in chemical pulping could be found in Simula and Alhoniemi (2006), Ciesielski and Olejnik (2014), Musavi et al. (1995), Sainlez and Heyen (2013).

The most critical changes in the morphology of wood chips occur in the refining stages, making refining the most energy-intensive process in the TMP mill. This research provides a refining optimization strategy in two-phase which can be further used to generate optimal set-points for refining control strategy. In the first phase, the refining identification models are developed utilizing a deep multilayer perceptron neural network. Refining identification models simulate the first and second stages of the refining specific energy consumptions, final pulp fiber length, and final pulp freeness based on the first and second stages refining variables collected from the pulp mill. Canadian Standard of Freeness (CSF) is considered a measure to evaluate the final pulp freeness. In phase 2, the Genetic optimization algorithm has been employed to minimize the total refining specific energy consumption by manipulating the first and second stage refining plate gap and dilution water. The developed refining identification models in phase 1 are utilized to construct the optimization objective function and constraints. The optimization model is subjected to keep the pulp quality the same as the current system operating condition to evaluate the performance of the refining optimization strategy. CSF and fiber length are considered as main variables to determine the final pulp quality. Finally, the optimization results are compared with the current operating conditions of the process to assess the capability of the proposed energy-saving refining optimization method. Paper novelty and contributions are summarized as follows:

- An advanced energy-saving refining optimization strategy is developed based on the integration of Artificial Neural Network (ANN) and Genetic algorithm (GA).
- The performance of the proposed energy-saving optimization strategy has been increased by considering the effect of refining disturbance variables on the system operation.

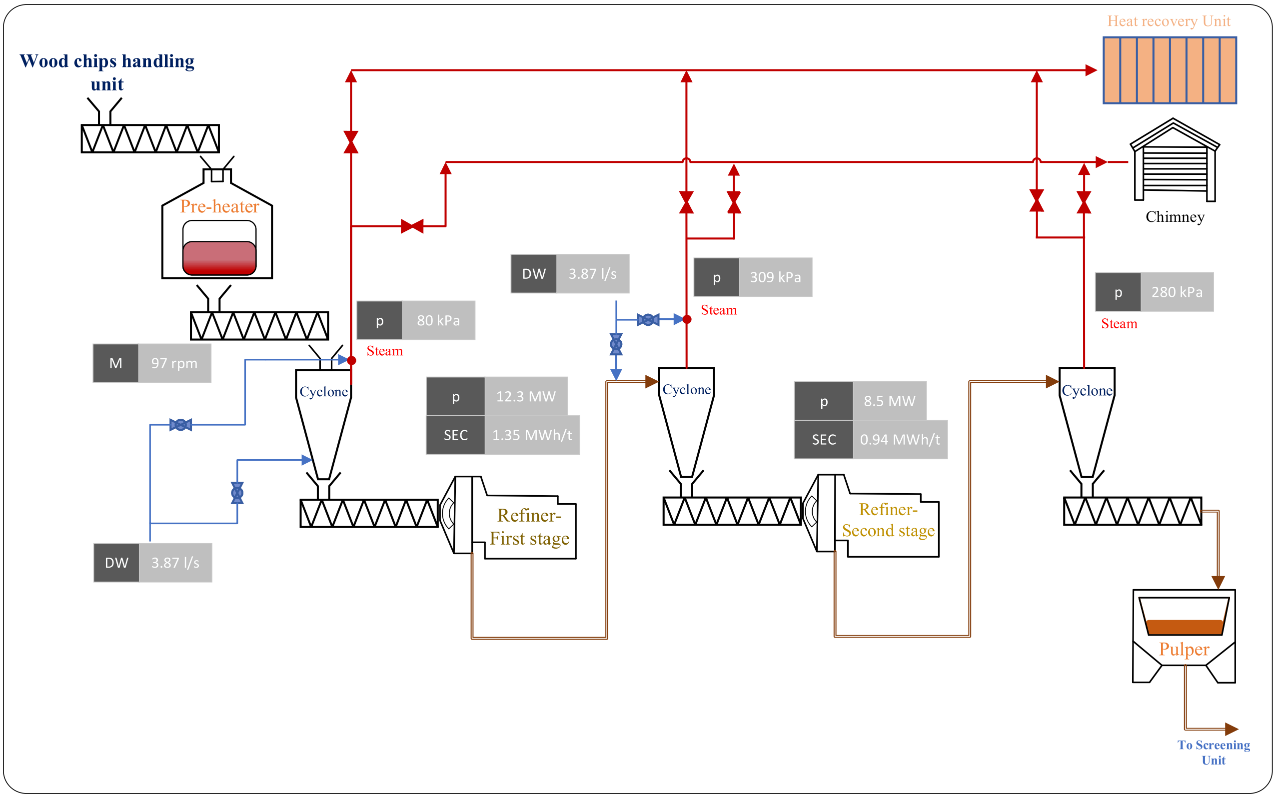

Figure 2: Schematic of the main refining line with sample measured data.

- The developed energy-saving optimization strategy has not been used before in the pulp and paper industry and is applicable to other complicated industrial processes such as chemical pulping.
- The proposed optimization strategy is compared with the current system operating condition to evaluate the efficiency of the suggested method.

Case study

Collected raw data for this study is measured data from a thermo-mechanical pulp mill refining line in a Nordic country. The studied mill has multiple parallel refining lines with the same technical details. This research considers the first refining line (main refining line) as the case study. The main refining line's schematic and representative measured process data are presented in Figure 2. Process measured data are refining plate gap, refining dilution water, refining motor load for the first and second refining stages, feeder screw speed, final pulp freeness (CSF), and pulp fiber length. Wood chips from the wood handling unit are pre-heated and fed to the first stage refiner. Cyclone in the following of each refiner separates the pulp from steam. The first cyclone in the refining line separates backflow steam (which is undesirable). The final pulp is directed to the screening unit after pulper for post refining treatment and probable third stage refining in the reject refiner.

The data used in this research are measured data from the TMP mill. Due to the contract made with the mill to maintain data security, the raw data has not been published, and the normalized version has been reported. Therefore, data in Figures 7–11, Figures 13–14, and Figure 17 are normalized to keep the data security. Those figures which report the percentage of changes in variables (such as Figure 12, Figures 15–16) are based on the actual data, not the normalized one.

Pulp quality control tests

Pulp freeness

The measurement of the water portion in the pulp suspension (Freeness measurement) is a common method to ob-

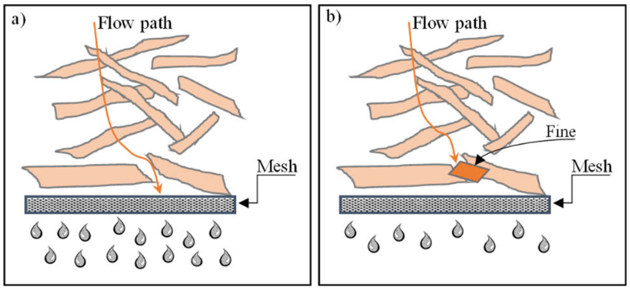

Figure 3: Pulp fibers characteristic and drainage ability a) before and b) after refining.

Table 1: Typical ranges of the pulp freeness and energy consumption in mechanical pulping.

Mechanical pulping subdivisions	Canadian Standard Freeness (ml)	Specific energy consumption (kWh/ton of pulp)
Thermomechanical (TMP)	400–30	800–3600
Chemi-thermomechanical (CTMP)	700–30	1000–4300
Refiner mechanical (RMP)	350–30	1600–3000
Groundwood (GW)	350–30	1100–2200
Pressure groundwood (PGW)	350–30	1110–2300

serve stock preparation corrections. Freeness value could represent the refining impacts on pulp properties since it has a massive correlation with fines formation and fiber fibrillation. Therefore, freeness could indicate the pulp and fiber characteristics (Bhardwaj et al. 2007).

The Canadian Standard Freeness (CSF-test) is mainly used to measure the drainability of a dilute pulp suspension under specific conditions according to ISO 5267-2 (suspension of 3 g of pulp in 1 L of water), especially in the laboratory scale. It is proven that some pulp properties such as the distribution of fiber length in the short and medium fractions, conditions of the surface, and fiber compression affect the pulp freeness. Drainability or freeness composes a helpful index of the intensity of mechanical treatment to which the pulp has been subjected (Browne et al. 2004). CSF decreases with refining and is sensitives to the quality of water and fines. CSF is one of the most well-known indicators to determine the pulp's quality and the refining level. CSF is utilized in this study to define the quality of the final produced pulp.

To better understand the effect of refining on freeness value, the theory of drainage ability in paper manufacturing could be helpful. Refining makes the fibers more flexible and softer by swelling the fibers with mechanical forces. In such a case, firmly entangled fibers make a web in the drainage test. Moreover, secondary fines are generated during the refining process external fibrillation or fiber shortening. These fines are not bound to the fibers and can easily move through the pulp suspension while eventually getting stuck between fibers pores. This phenomenon decreases the drainage ability since fines block the water flow path and increase the drainage time. Therefore, the drainage ability is deteriorated by the refining, which is not desirable in the paper-making process (Gao et al. 2009, Hubbe and Heitmann 2007, Paradis et al. 2002). Figure 3 gives more information about the fibers and fines formation before and after refining (Gharehkhani et al. 2015).

Table 1 shows the typical ranges of the energy consumption and fiber length for different mechanical pulping subdivisions such as thermo-mechanical pulping (TMP), chemi-thermomechanical pulping (CTMP), refiner mechanical pulping (RMP), groundwood (GW), and pressure groundwood (PGW) pulping (Gharehkhani et al. 2015).

Fiber length

Fiber shortening is an unwanted change in the fiber quality due to fiber beating (refining) (Kerekes and Olson 2003). Fiber breaking could occur by sufficient stress on the fiber or when fibers get chopped through shearing forces during the passages of the refiner bars or in case of being pulled from other fibers (Kerekes 1990). Also, fines are generated during the refining as a consequence of extreme force throughout external fibrillation. Therefore, there is a clear correlation between fines generation and fiber cutting (Batchelor et al. 1994). More information regarding the effect of fiber length on paper characteristics has been provided and investigated in Pulkkinen (2010), Seth and Page (1988). Table 2 demonstrates the pulping properties of US Softwoods and Hardwoods (Bajpai 2018).

Theory

Artificial Neural Network (ANN): the deep learning method

ANN models are becoming more and more popular in recent decades because of their capabilities to learn and discover nonlinear patterns, adapt to environmental disturbances, and tolerate errors in datasets and measurements. Multilayer Perceptron Neural Networks (MLPNNs) are the

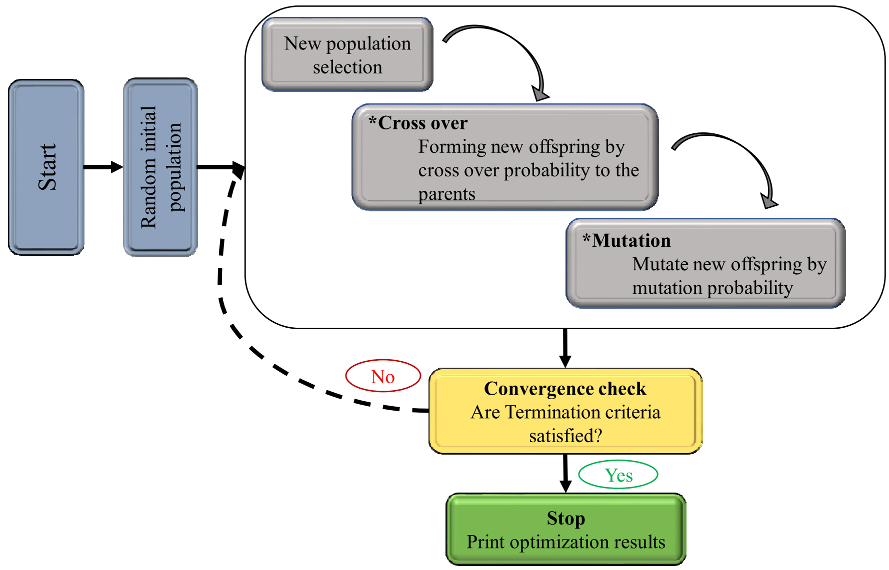

Figure 4: The design structure of the genetic algorithm.

Table 2: Pulping Properties of the US Softwoods and Hardwoods.

Species	Scientific Name	Average Fiber Length (mm)
Softwoods		
Jack (pine)	Pinus banksiana	3.5
Loblolly (pine)	Pinus taeda	4.00
Lodgepole (pine)	Pinus contorta	3.5
Monterey (pine)	Pinus radiata	2.60
Black (spruce)	Picea mariana	3.5
Blue (spruce)	Picea pungens	2.8
Hardwoods		
Ailanthus	Ailanthus, altissima	1.20
Ash, green	Fraxinus pennsylvanica	1.05
Maple, silver	Acer saccharinum	1.75

prominent and popular feedforward artificial neural networks. The MLPNN models are easier to implement, function, and have a high level of training capability, even with small datasets. Deep learning implemented with Deep Neural Networks (DNNs) are models with more profound neural network architecture where input data experience more transformation to form model output. Recent studies show that deep learning approaches emerged promising in pattern recognition and simulating complicated industrial processes such as mechanical and chemical pulping. Although these models require higher execution time due to the more complex structure of deep learning models, deep learning models are very efficient in detecting nonlinear hidden relationships between variables (Schmidhuber 2015).

Genetic Algorithm (GA)

The Genetic algorithm (GA) is an optimization technique used to solve nonlinear or non-differentiable optimization problems. GA optimization is a heuristic searching method based on the natural evolution of Charles Darwin's theory. The name Genetic algorithm originates from the fact that it mimics the evolutionary biology techniques but in a number domain. The process begins with a random initial generation of the candidate's solution tested against the objective function. There is competition between population members to get reproduction right. Those who are chosen for the next generation have a better performance in fitness function reduction. Therefore, the fitter individual has more chances to be selected for the next generation. At the end of each iteration, the selected individuals are licensed to produce progeny populations for the next iteration if the stop condition is not met. The optimization flowchart is summarized in Figure 4. After some iterations, the optimum solution is obtained, where the stopping criteria are satisfied, and the program converges.

Refining identification models and feature selection

Refining energy and pulp quality identification models

The refining identification model describes the hidden correlation between refining variables. A mathematical model is required to simulate refining specific energy con-

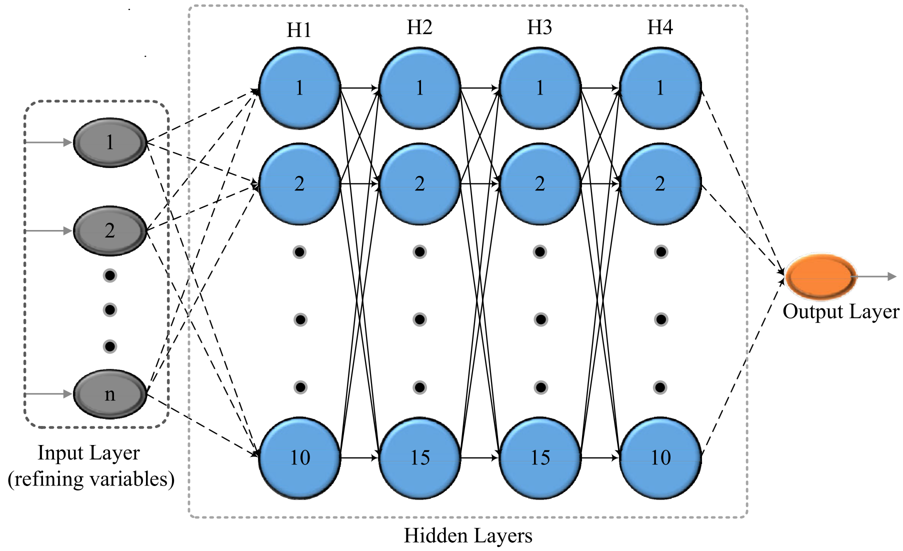

Figure 5: The architecture representation of the optimum ANN consists of four hidden layers.

sumption and pulp quality (fiber length and CSF). The mathematical model in this study is generated using an Artificial Neural Network. Four neural networks are established and trained to predict the first and second stages of the refining specific energy consumption, final pulp fiber length, and final pulp freeness (CSF value). Model input features and outputs for each neural network are explained in Refining variables and ANN feature selection section.

Obtaining optimal hyperparameters of an MLP usually includes a trial-and-error approach, and there is no available predefined or analytical method. Hyper-parameters of the MLP model, such as the number of neurons and hidden layers, are calculated based on the random search method, which has a lower computational load than the other trial-and-error approaches such as grid search (Bergstra and Bengio 2012). For the random search purpose, the examined neural networks include up to 4 hidden layers and 20 neurons in each layer. Optimal results show better performance of a Multilayer Perceptron Neural Network with four hidden layers, ten neurons in the first and fourth hidden layers, and fifteen neurons in the second and third hidden layers. Figure 5 depicts the topology of the employed multilayer perceptron neural network.

Identification model data sets consist of 3800 measured data from the TMP mill, randomly divided into three parts; 70 % as training data set, 15 % for validation, and the remaining 15 % are dedicated for testing data set. The training subset contributes to update the network biases, weights and calculate gradients. The validation data set evaluates the model fit on the training data set while tuning the model parameters. Validation error determines the running stopping point to avoid overfitting to the training data set. When the model is entirely trained, the test data set is used to evaluate the model's accuracy by the performance function of mean squared error (MSE). In order to build a feedforward neural network, 'trainlm' network training function is utilized in MATLAB software. The network bias and weights are being updated according to Levenberg-Marquardt optimization. Although 'trainlm' requires more memory to run, it is still considered the best option for supervised learning as it is the fastest backpropagation algorithm in the MATLAB toolbox. However, this model has some limitations. The 'trainlm' function employs the Jacobian for computations, which supposes that performance is a mean or sum of squared errors (MSE). Accordingly, networks with this training function must use MSE as a performance function.

The levenberg-Marquardt algorithm is developed to increase the training speed without the need to compute the Hessian matrix, similar to the Gauss-Newton technique. The Hessian matrix can be approximated by Equation 1 for the ordinary feedforward network performance function with a sum of squares formation.

$$H = J^T J \tag{1}$$

and the gradient can be computed as:

$$g = J^T e \tag{2}$$

Where e is a vector of network errors and J is the Jacobian matrix including the first derivatives of the network errors with respect to the biases and weights. The

Table 3: Details of each neural network data sets.

Input data combinations	Refining plate gap – 1st stage	Refining dilution water – 1st stage	Refining plate gap – 2nd stage	Refining dilution water – 2nd stage	Feeder screw speed	Time
Data set 1	✓	✓	✗	✗	✓	✓
Data set 2	✗	✗	✓	✓	✓	✓
Data set 3	✓	✓	✓	✓	✓	✓

Jacobian matrix can be calculated via a standard backpropagation approach which induces less computational load and complexity than calculating the Hessian matrix. Equation 3 is the Marquardt-Levenberg modification to the Gauss-Newton method:

$$x_{k+1} = x_k - [J^T J + \mu I] J^T e \tag{3}$$

I is the identity matrix. The algorithm is as same as Gauss-Newton's method for small μ. For the large μ, the algorithm becomes gradient descent with a low step length. Newton's method has a higher speed around the minimum error, so it is determined to move towards Newton's method as fast as possible. The μ decreases after each successful step that reduces the performance function, whereas if an unconfirmed step increases the performance function, it causes to increase μ. Thus, the algorithm is designed in such a way that the performance function is always reduced along with it. A comprehensive description of the Levenberg-Marquardt algorithm is provided in Marquardt (1963). In addition, the application of the Levenberg-Marquardt algorithm to the neural network concept is discussed in Hagan and Menhaj (1994). This algorithm has been shown to have the highest speed in training medium-sized neural networks that can cover up to several hundredweights. Since MATLAB has an internal function for matrix equation solution, this algorithm shows high performance in MATLAB software, and its features are more prominent in MATLAB.

Refining variables and ANN feature selection

Refining variables could be categorized into refining manipulated, disturbance, and operating variables. Manipulated variables are independent variables that influence refining operating variables. Refining operating variables such as production rate, refining motor load, pulp freeness, and fiber length can be predicted by the combination selection of the manipulated variables. The main two-stage refining process manipulated variables are chip transfer screw speed, primary and secondary refining plate gap, and dilution water (Tian et al. 2020). Disturbance variables are basically uncontrollable and disrupt the optimal control of the process by changing the system operating conditions. The main refining disturbance variables are seasonal changes in input wood chips quality and refiners plate condition. Changes in the quality of wood chips, such as moisture content, affect the pulp consistency and production rate. It should be noted that there is no accurate sensor for measuring pulp consistency, and the consistency measurement includes huge undesirable errors (in the studied mill, the error could be up to 20 %).

On the other hand, the exact rate of pulp production must be calculated based on pulp consistency, but due to errors in measuring the pulp consistency, pulp mills usually calculate the production rate according to the empirical equations based on screw feeder speed. This calculation method also has an error, but according to the experience, it seems more reliable than other consistency-based calculation methods. On the other hand, with the operation of refiner plates over time, these plates wear out due to the harsh conditions inside the refining zone, which causes changes in the system's operating condition. Li et al. (2011) reviews the effect of main variables of wood chips qualities and refining process on the final quality of TMP product. To apply the effect of disturbance variables that affect the performance of the refining process, Talebjedi et al. (2021) suggests adding the time variable as one of the predictor variables of the ANN refining identification model.

Three data sets are prepared to predict refining energy and quality operating variables by training four neural network models. Each dataset includes selected features combination as ANN model input variables for the particular refining operating variable prediction. Data set 1 is responsible for creating the first-stage refining specific energy consumption (RSEC) identification model. The second data set (data set 2) is used as the set of selected model features to simulate the second stage refining specific energy consumption, while the third data set (data set 3) is dedicated for final pulp freeness (CSF) and fiber length prediction. The total RSEC forms the objective function of the optimization model. Table 3 presents the data

Table 4: Refining neural networks identification models input (predictor) and output (target) variables.

ANN^1 (Refining energy identification model)		ANN^2 (Refining energy identification model)		ANN^3 (Pulp quality identification model)		ANN^4 (Pulp quality identification model)	
Predictor variables	Target variable	Predictor variables	Target variable	Predictor variables	Target variable	Predictor variables	Target variable
Data set 1	RSEC – First stage	Data set 2	RSEC – Second stage	Data set 3	Fiber length	Data set 3	Freeness (CSF)

sets details which construct the neural networks input feature combinations. Table 4 illustrates the different refining neural networks identification models input (predictor) and output (target) variables.

Refining energy-saving two-phase optimization plan

The simplified construction of the two-phase optimization plan is presented in Figure 6. The optimization process starts with phase 1, where generating refining energy and pulp quality identification models (neural networks 1 to 4). As mentioned earlier, four datasets have been used to model refining energy consumption and final refined pulp quality, the details of which are given in Tables 3 and 4. After the training phase (phase 1), neural networks are used to form the GA optimization constraints and objective function in phase 2. The optimization method is limited to keep the paper's quality (CSF value and fiber length) at the expected level (current operating condition) while minimizing the model objective function, which is the summation of the first and second stage refining specific energy consumption (RSEC). Other optimization constraints define the boundaries of the decision variables based on the real system operating condition and the component characteristic.

The optimization decision variables are the first and second stages refining dilution water and plate gap. Once the optimization is completed, the model reports the optimal values of the primary and secondary refining dilution water, plate gap, and refining specific energy consumptions as the final results. The obtained total refining specific energy consumption from the optimization plan is compared to the measured values of the total refining specific energy consumption from the mill (current system operating condition) to evaluate the performance of the proposed optimization method to increase the refining energy efficiency. The following gives more information about the optimization objective function and constraints.

Optimization model objective function

The optimization objective function given in Equation 4 is the sum of the primary and secondary refining specific energy consumptions. Refining is responsible for the most changes in the fiber morphology and is the most energy-intensive process in thermo-mechanical pulping. In our case study, the average motor load for the secondary refiner (pulp refiner) is 70 % of the primary refiner (wood refiner), making the first stage of refining more energy-intensive.

$$ObjFcn\colon RSEC^1 + RSEC^2 \tag{4}$$

Optimization model constraints

The optimization model constraints are about the limits on the upper and lower boundaries of the decision variables as well as the expected quality of the paper, including the pulp freeness value and fiber length. Quality constraints are defined as inequality constraints to give more flexibility to the model to avoid causing the probable infeasible optimization model. For this reason, the average neural network error is considered as an acceptable interval at each data point that the CSF value could settle to satisfy the model pulp quality constraint. This fact is also true for the fiber length as a measure of pulp quality control. The following equations show the optimization model constraints:

$$FL^{m,j} - FL^{m,j} \times MAPE^{ANN^k} \le FL^j \le FL^{m,j} + FL^{m,j} \times MAPE^{ANN^k,j} \quad j = 1:N;\ K = 3 \tag{5}$$

$$CSF^{m,j} - CSF^{m,j} \times MAPE^{ANN^k} \le CSF^j \le CSF^{m,j} + CSF^{m,j} \times MAPE^{ANN^k,j} \quad j = 1:N;\ K = 4 \tag{6}$$

$$DW^{min,i,j} \le DW^{i,j} \le DW^{max,i,j} \quad j = 1:N;\ i = 1 \tag{7}$$

$$PG^{min,i,j} \le PG^{i,j} \le PG^{max,i,j} \quad j = 1:N;\ i = 1 \tag{8}$$

$$DW^{min,i,j} \le DW^{i,j} \le DW^{max,i,j} \quad j = 1:N;\ i = 2 \tag{9}$$

$$PG^{min,i,j} \le PG^{i,j} \le PG^{max,i,j} \quad j = 1:N;\ i = 2 \tag{10}$$

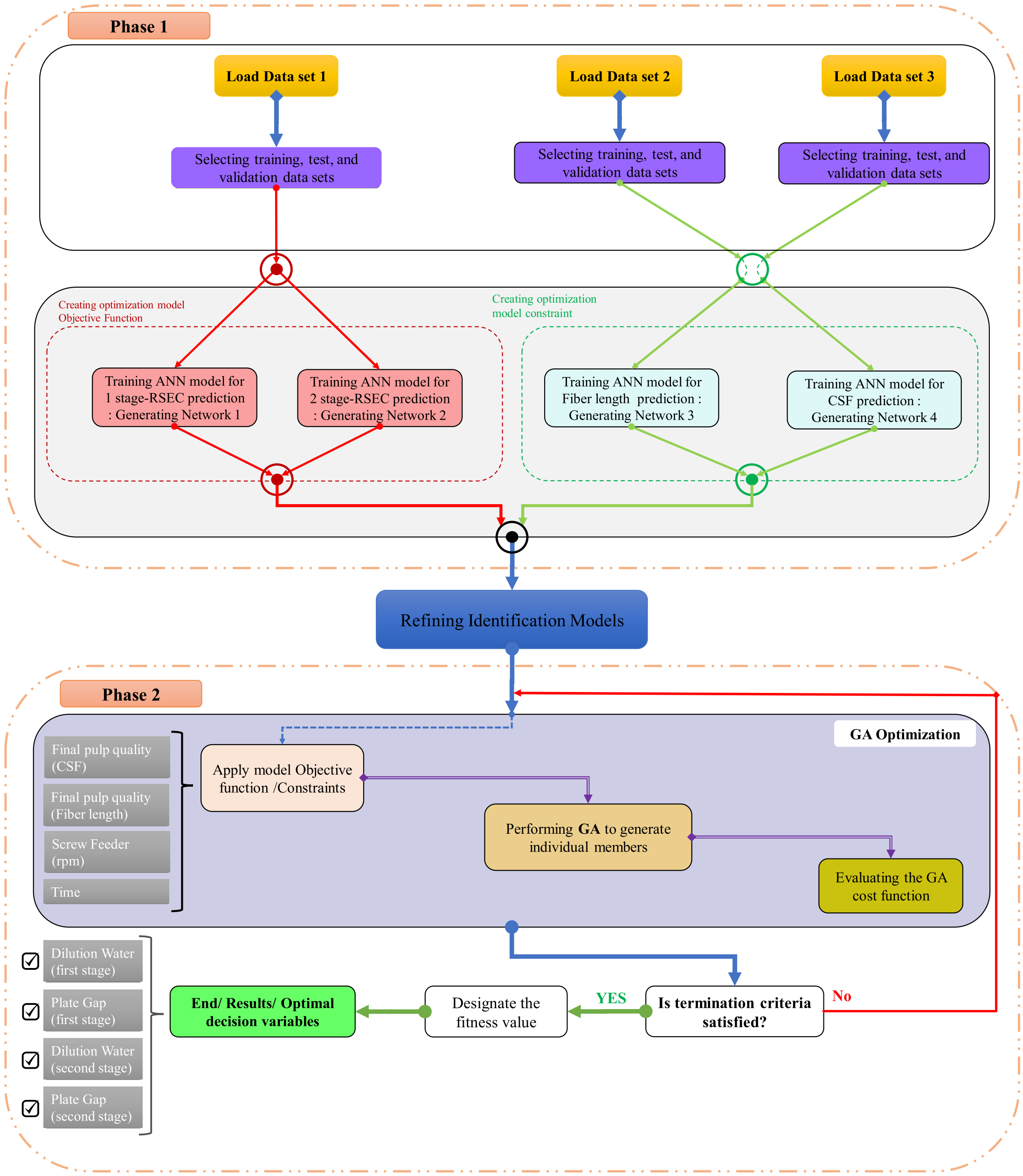

Figure 6: Schematic representation of the proposed advanced energy-saving refining optimization strategy.

Where i indicates the refining stage ($i = 1, 2$) and j is the index for each data point. k specifies the neural network number, and m stands for the measured value. N is the total number of data points.

Identification models evaluation criteria

Mean square error (MSE) is the neural network's performance function. It measures the network's performance according to the mean of squared errors. A lower value for MSE designates a higher accuracy of the model. MSE is a scale-dependent accuracy evaluation criterion that could be used to explain the error in the original scales. Whereas for comparing different models with different datasets, other metrics such as Mean absolute percentage error (MAPE) are required. Equation 11 and Equation 12 provide the mathematical terms for MSE and MAPE calculations, respectively.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_j - y_j)^2 \tag{11}$$

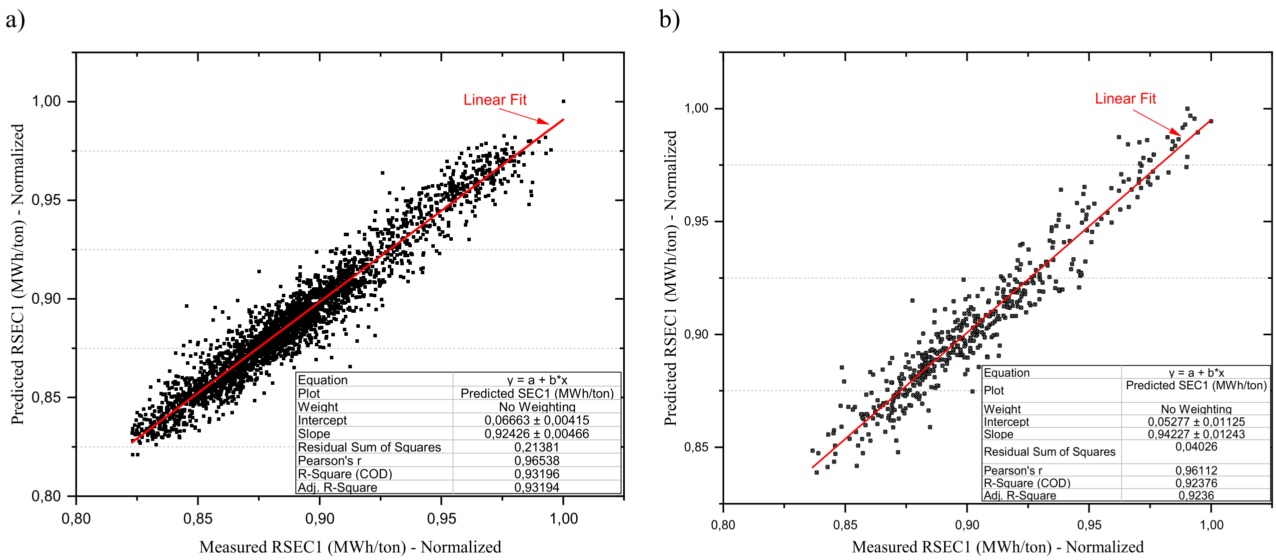

Figure 7: Refining SEC^1 ANN model regression plots for the (a) train and (b) test data sets (model: ANN1).

$$MAPE = \frac{1}{n}\sum_{j=1}^{n}\frac{|x_j - y_j|}{|x_j|} \tag{12}$$

The model correlation coefficient assesses the extent of correlation between target and model output variables. The correlation of zero ($R = 0$) indicates that there is no relationship between model output and target variables, while $R = 1$ indicates the perfect linear relation between the mentioned variables. The model correlation coefficient is calculated using Equation 13.

$$R = \frac{\sum_{j=1}^{n}(x_j - \overline{x})(y_j - \overline{y})}{\sqrt{\sum_{j=1}^{n}(x_j - \overline{x})^2 \sum_{j=1}^{n}(y_j - \overline{y})^2}} \tag{13}$$

Where x_j, y_j, x, y, and n are measured value, predicted value, mean of measured data, mean of predicted data, and the number of data, respectively.

Table 5: Neural Networks prediction performances on training and testing data sets.

Neural network Model	Metrics	Train	Test
ANN^1 (First stage-RSEC prediction)	R	0.98	0.97
	MSE	0.0112	0.0124
	$MAPE$(%)	0.79	0.85
ANN^2 (First stage-RSEC prediction)	R	0.95	0.94
	MSE	0.0122	0.0136
	$MAPE$(%)	1.06	1.18
ANN^3 (CSF prediction)	R	0.72	0.67
	MSE	3.112	4.313
	$MAPE$(%)	1.69	1.95
ANN^4 (Fiber length prediction)	R	0.95	0.92
	MSE	0.0110	0.0117
	$MAPE$(%)	0.46	0.55

Results and discussion

Identification models

For the optimization plan, mathematical models (identification models) are required to model the refining specific energy consumption and final pulp quality based on system manipulated variables. The models for the refining (first and second stages) energy simulation form the optimization objective function, and the models for the final pulp quality simulation contribute to construct the optimization constraints. In this research, the concept of artificial neural networks has been employed for refining identification models generation. The identification models are four neural networks, each trained for a particular purpose. ANN^1 to ANN^4 are generated to predict the first and second stages of refining specific energy consumption, final pulp fiber length, and final pulp freeness value. The accuracy of the obtained models is examined by the three variables listed in Table 5 and more discussed in Identification models evaluation criteria section. In addition, data are surveyed to check the relationship between the measured and predicted variables. Figures 7–10 show the regression plot between predicted and measured data for refining SEC, final pulp CSF, and final pulp fiber length simulation. A higher value for the model correlation coefficient

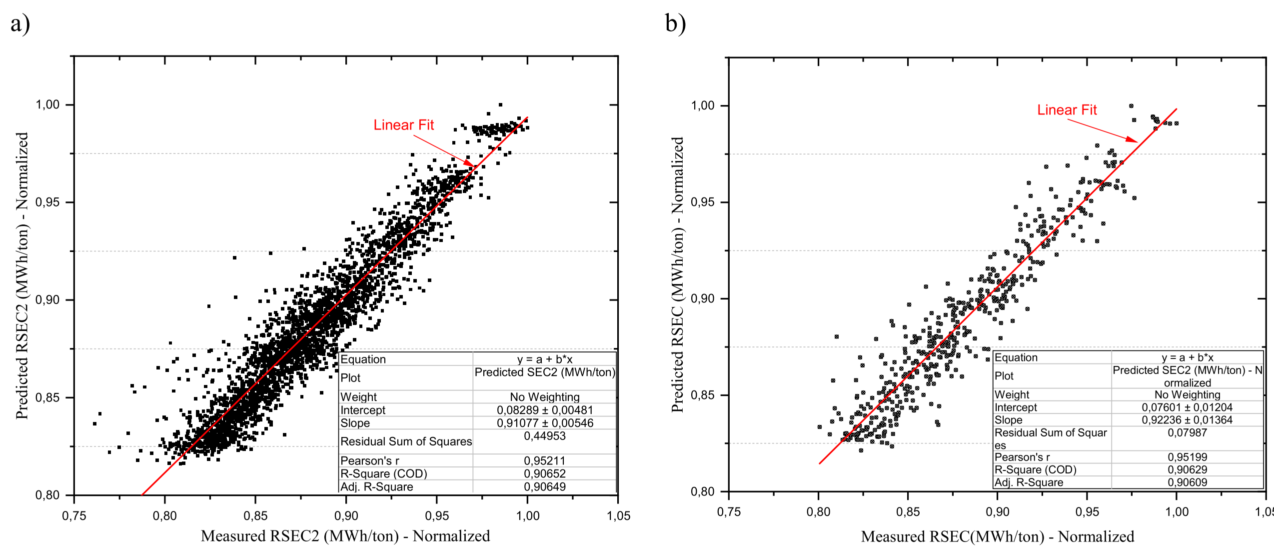

Figure 8: Refining SEC^2 ANN model regression plots for the (a) train and (b) test data sets (model: ANN2).

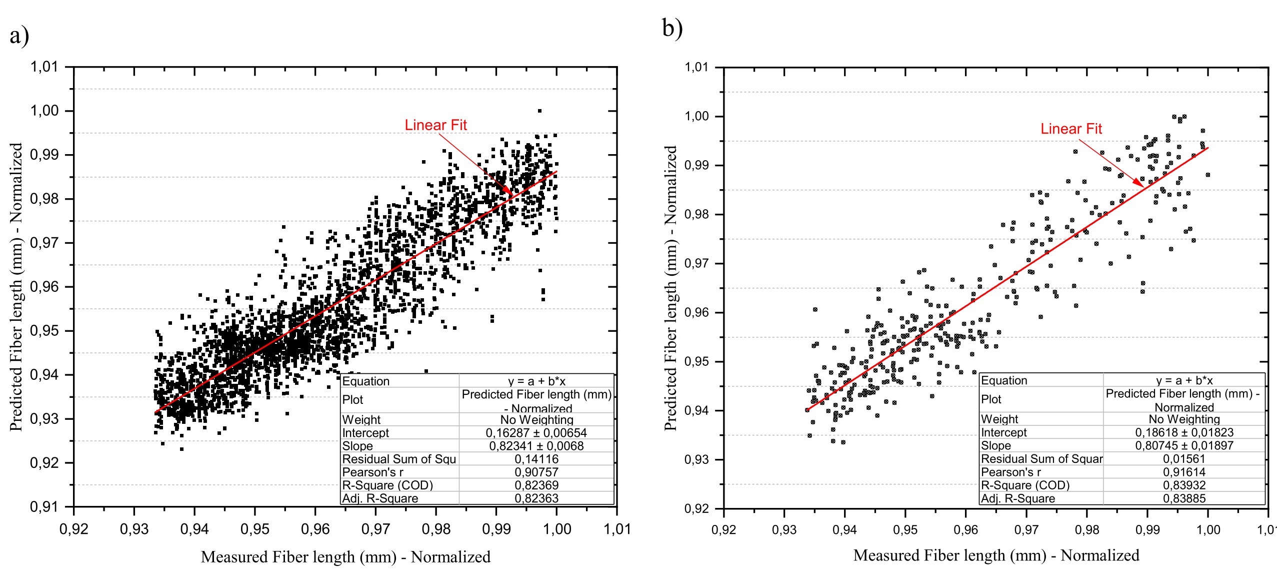

Figure 9: Final pulp Fiber length ANN model regression plots for the (a) train and (b) test data sets (model: ANN3).

demonstrates a better connection between the target and estimations of response.

According to Table 5, the evaluation criteria prove the high efficiency of neural network models in predicting refining operating variables. The obtained mathematical models are reliable to use for building optimization model objective function and constraints. Results confirm lesser accuracy for CSF prediction than other refining operating variables, which could be because of ignoring the effect of some other refining variables affecting CSF. Investigating the cause is beyond the scope of this article and is intended as a research question only. From the performance perspective, the optimum ANN structure for CSF prediction gives a correlation of 0,72 for the train and 0,67 for the test data set, whereas for fiber length prediction, the ANN model correlation coefficient is 0.95 for the train and 0.92 for the test data set. The best ANN practice for the first and second stage RSEC correlation coefficient is 0.98, and 0.95 for the train and 0.97, and 0.94 for test data sets, respectively. On the other hand, the obtained mean square error (*MSE*) for the best practice is 0.0124, 0.0136, 4.313, and 0.0117 for the first stage RSEC, second stage RSEC, final pulp CSF, and final pulp fiber length (FL) simulation, respectively.

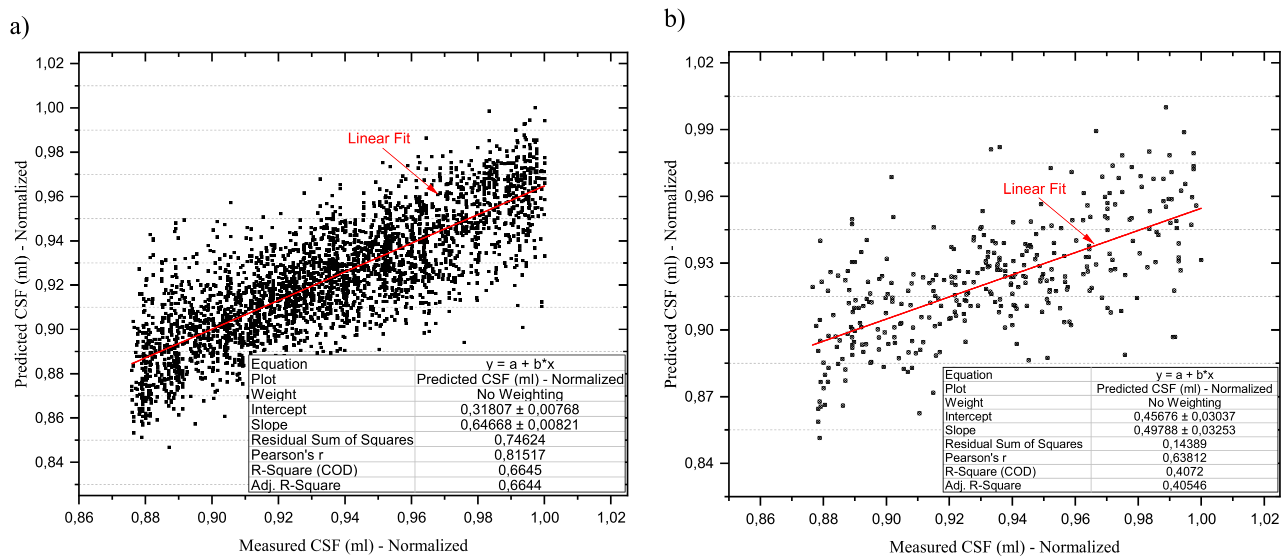

Figure 10: Final pulp freeness ANN model regression plots for the (a) train and (b) test data sets (model: ANN4).

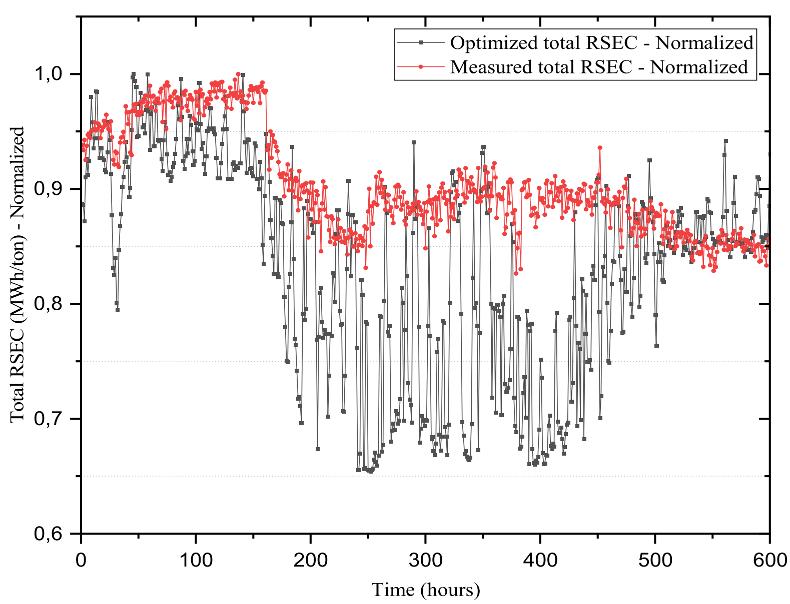

Figure 11: Representative measured and optimized data for the total RSEC.

Optimization result

Objective function

The optimization aims to minimize the total refining specific energy consumption at the same pulp quality as the current system operating condition by manipulating the refining manipulated variables. The integration of artificial neural networks (to develop refining identification model) and Genetic algorithm (as an optimization tool) has been proposed for the optimization strategy. The reliability of the proposed method for energy efficiency improvement in the refining process of thermo-mechanical pulp mill has been examined by comparing the optimization results to the measured values of the total RSEC. Figure 11 demonstrates measured and optimized total RSEC values of the first 600 representative data points. Measured values that have been pointed out by the red dots are the mill's collected data under the current operating condition where the refining process is handled utilizing the model predictive controller (MPC). Black dots represent the optimized values of total RSEC obtained from the

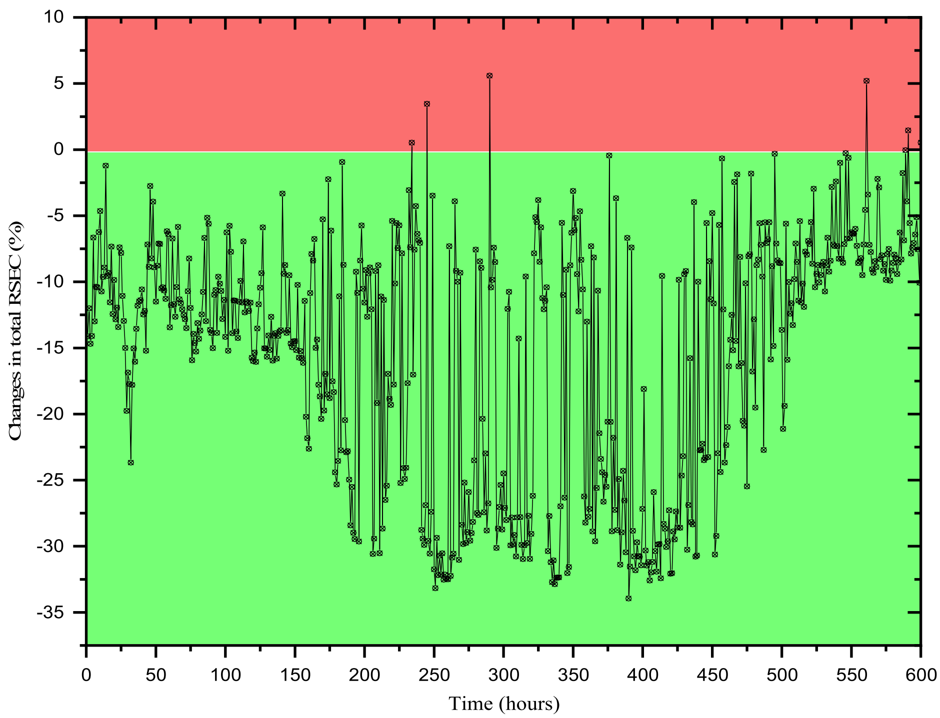

Figure 12: Representative changes in total RSEC while utilizing the proposed optimization plan compared to the current system operating condition (the calculation term is explained in Equation 14).

optimization plan proposed in this research. As mentioned earlier, these values are normalized to keep our promise to the mill regarding maintaining data security. However, the percentage of changes shown in Figure 12 is for actual data, not normalized. It is clear from Figure 11 that almost in all of the data points, black dots that indicate the optimized values are below the red dots, which are the measured ones; meaning the optimization method has reduced the refining specific energy consumption, which implies the fact that refining energy efficiency has been increased. Figure 12 illustrates the percentage of total refining specific energy consumption changes using the proposed optimization strategy compared to the current system operating condition. Findings confirm that a specific energy consumption reduction of up to 35 % could be possible in some system operating conditions by employing the proposed method, while the average reduction is 14 %. Equation 14 shows the expression for the percentage of change in the refining variables.

$$Change(\%) = 100 \times \frac{Optimized\ value - measured\ value}{measured\ value} \tag{14}$$

Decision variables (manipulated refining variables)

The optimum combination of refining plate gap and dilution water plays an essential role in constructing the refining optimization strategy from the energy and quality point of view (Harinath et al. 2013, Elahimehr et al. 2018). The optimization model decision variables in the studied refining process are the refining plate gap and dilution water for the first and second refining stages. These variables are considered the most essential manipulated refining variables to optimize the refining energy consumption and control final pulp quality. Figures 13 and 14 show the optimum values of the first and second-stage dilution water flow rate and plate gap obtained by the proposed optimization method as well as the measured values under the current system operating condition for 600 hours' time period.

Figures 15 and 16 show the percentage of changes in refining manipulated variables from the measured values to the optimum values. The expression to calculate the percentage of changes in the refining variables from the optimized values by the proposed optimization strategy and the current system operating is provided in Equation 14. It is clear from Figures 15 and 16 that in the majority of the data points, the optimum value of the first-stage refining plate gap and second-stage dilution water by the proposed optimization strategy is higher than the current system operating condition with an average change of 11.56 % and 3.50 %, respectively. Meanwhile, the optimum first-stage dilution water and second-stage plate gap are less than the

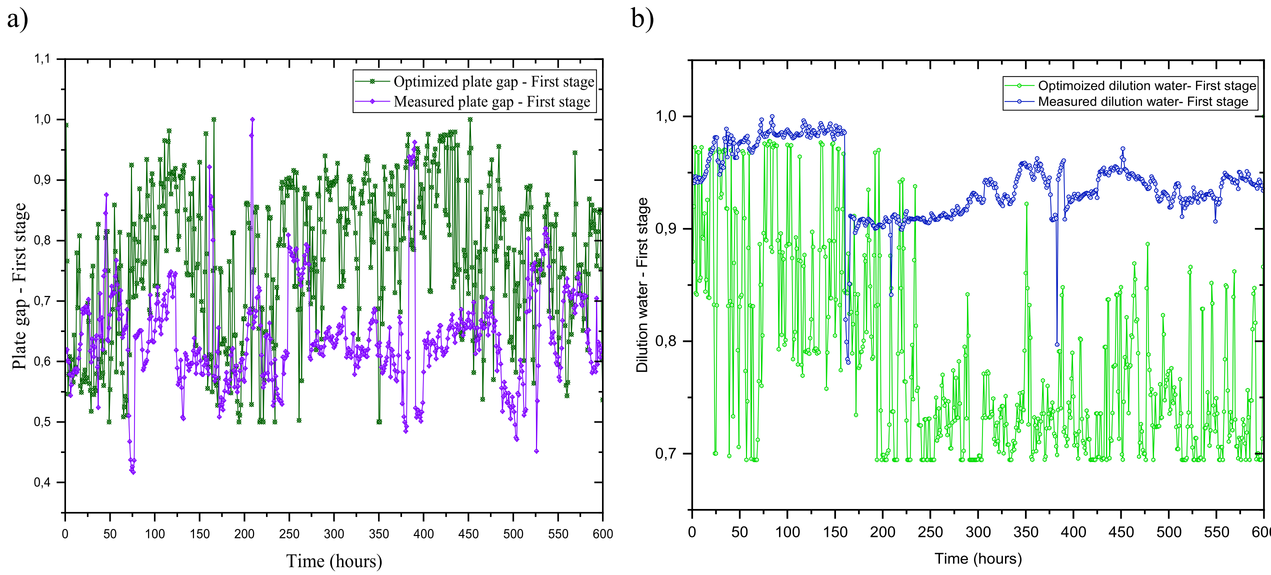

Figure 13: Representative normalized optimum and measured values for first-stage a) refining plate gap, and b) refining dilution water.

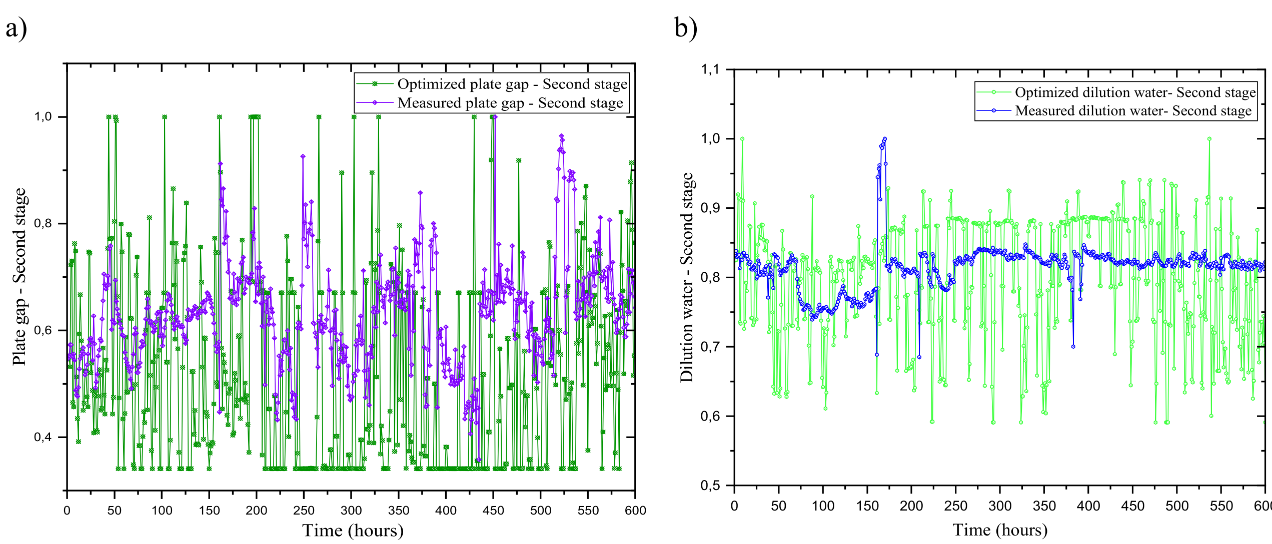

Figure 14: Representative normalized optimum and measured values for second-stage a) refining plate gap, and b) refining dilution water.

measured values. The average changes in the first-stage dilution water and second-stage plate gap are −7.6 % and −27.52 %, respectively.

Optimization constraints: final pulp quality limit

Energy-saving optimization plan is limited to keep the final pulp quality within a certain range around the current pulp quality. Therefore, process optimization should be designed in a way that the pulp quality is maintained within the desired range, which is close to the current system operating condition. In order to prevent the infeasibility of the optimization model and increase the processing speed, some flexibility is applied to the model constraints. The degree of flexibility is due to the error in the final pulp quality identification model, which is discussed in detail in Refining energy-saving two-phase optimization plan section. The ultimate goal of the proposed optimization strategy is to achieve a certain pulp quality following the current system operating conditions with the least total refining specific energy consumption. Therefore, the closer the estimated variables of pulp quality after energy optimization and the measured variables (current system

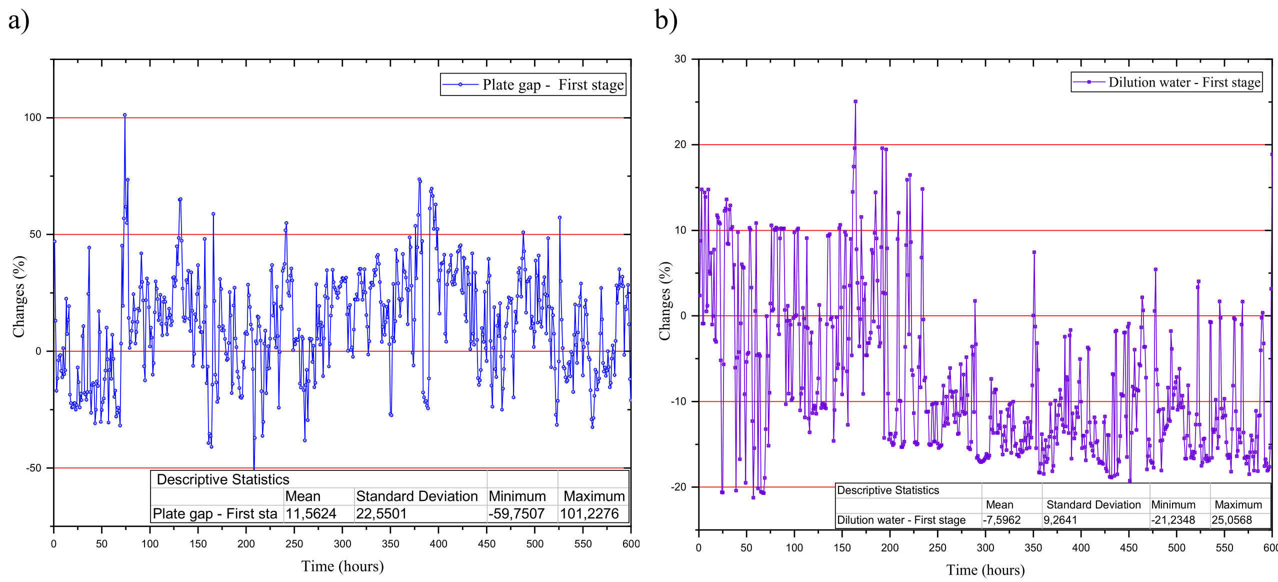

Figure 15: Representative percentages of change in the first-stage a) plate gap and b) dilution water from the measured value to the optimized value (the calculation term is explained in Equation 14).

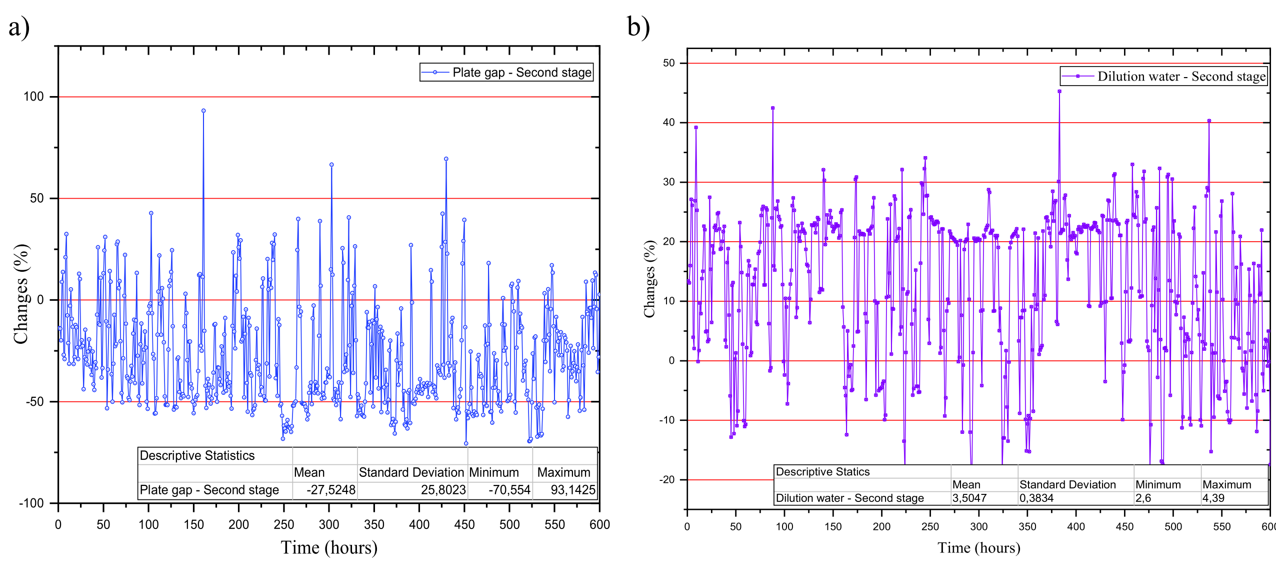

Figure 16: Representative percentages of change in the second-stage a) plate gap and b) dilution water from the measured value to the optimized value (the calculation term is explained in Equation 14).

operating condition) are, the easier it is to evaluate the performance of the proposed optimization strategy. Figure 17 shows the measured and optimization strategy final pulp fiber length and freeness (CSF). The proposed energy-saving optimization strategy aims to minimize the percentage of changes in paper quality and keep it as close to the measured values because the reduction in refining specific energy consumption must be such that the quality of the paper is maintained. Fewer changes in the pulp quality indicate that the proposed optimization algorithm has a higher power in maintaining pulp quality while reducing energy consumption. The absolute average changes in the final pulp FL and CSF are 1.33 % and 1.53 %, respectively.

Conclusions

This paper proposes an advanced energy-saving optimization strategy in TMP mill by integration of machine learn-

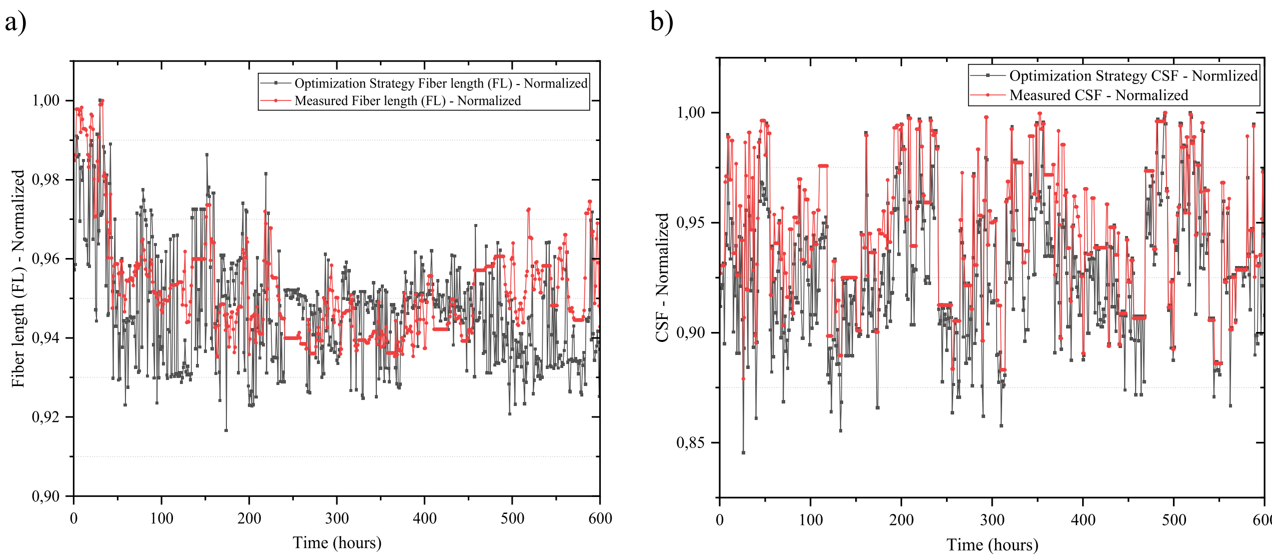

Figure 17: Final pulp quality before (red dots) and after (black dots) applying the proposed optimization strategy: a) Final pulp FL, b) Final pulp CSF.

ing algorithm (artificial neural networks) and a heuristic optimization approach (Genetic algorithm). In the beginning, the refining identification models are created using artificial neural networks to simulate the two-stage high consistency refining process from the energy and pulp quality point of view. The results indicate the high accuracy of artificial neural networks in modeling the refining process, confirming the license to use these models as refining identification models to develop the refining optimization strategy. While the generated neural networks are utilized as system identification models, the Genetic Algorithm (GA) is employed as an optimization solver to minimize the total refining specific energy consumption. The optimization objective function is the sum of the first and second stage refining specific energy consumption where the optimization is subjected to keep the final pulp quality at the current system operating condition. Canadian Standard of Freeness (CSF) and pulp fiber length (FL) are considered as a measure to evaluate the final pulp quality. Optimization decision variables are the first and second-stage refining dilution water and refining plate gap. The proposed refining optimization strategy results are compared with the current system operating condition to evaluate the energy-saving potential. Findings confirm that an energy consumption reduction of up to 35 % could be possible in some system operating conditions by employing the proposed method, while the average reduction is 14 %, equal to 305 kWh per ton of production.

Funding: The authors appreciatively acknowledge the funding of the Academy of Finland, grant number 315020.

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

Nomenclature

ML^i	Refining Motor load (i^{th} stage)
PG^i	Refining plate gap (i^{th} stage)
DW^i	Dilution water (i^{th} stage)
$RSEC^i$	Refining specific energy consumption (i^{th} stage)
CSF	Canadian Standard Freeness
J	Jacobian matrix
e	Vector of network errors
R	Correlation coefficient
MAPE	Mean absolute percentage error
MSE	Mean squared error
MPC	Model predictive control
AI	Artificial intelligence
ANN	Artificial Neural Network
MLP	Multi-Layer Perceptron
ANFIS	Adaptive neural fuzzy inference system
GA	Genetic Algorithm
PSO	Particle swarm optimization
TMP	Thermo-mechanical pulping

Indices

i	refining stage (1,2)
j	index for data points
k	iteration step
N	total number of data points
m	measured value
max	maximum limit value for the variable
min	minimum limit value for the variable

References

Bajpai, P. Pulp and Paper Industry: Energy Conversion. Elsevier, 2016.

Bajpai, P. Biermann's Handbook of Pulp and Paper: Raw Material and Pulp Making. Third edition. Elsevier, Kanpur, India, 2018.

Batchelor, W., Kjell-Arve, K., Ouellet, D. (1994) Refining and the development of fibre properties. Nord. Pulp Pap. Res. J. 14(4):285–291.

Bergstra, J., Bengio, Y. (2012) Random search for hyper-parameter optimization. J. Mach. Learn. Res. 13:281–305.

Bhardwaj, N.K., Hoang, V., Nguyen, K.L. (2007) A comparative study of the effect of refining on physical and electrokinetic properties of various cellulosic fibres. Bioresour. Technol. 98(8):1647–1654.

Browne, B.T., Miles, K., Mcdonald, D., Wood, J., Browne, T. (2004) Multivariate analysis of seasonal pulp quality variations in a TMP mill. Pulp Pap. Can. 105(10):35–39.

Ciesielski, K., Olejnik, K. (2014) Application of Neural Networks for Estimation of Paper Properties Based on Refined Pulp Properties. Fibres Text. East. Eur. 5(107):126–132.

Elahimehr, A., Olson, J.A., Martinez, D.M. (2018) Understanding LC refining: The effect of plate pattern and refiner operation. Nord. Pulp Pap. Res. J. 28(3):386–391.

Fuhr, B.J., Henry, D., Leary, G., Smith, G. (1998) Seasonal variations at a mechanical newsprint mill. Pulp Pap. Can. 99(2):45–49.

Gao, Y., Huang, F., Rajbhandari, V., Li, K., Zhou, Y. (2009) Effect of separate refining and co-refining of bctmp/kp on paper properties. Pulp Pap. Can. 110(6):28–33.

Gharehkhani, S., et al. (2015) Basic effects of pulp refining on fiber properties – A review. Carbohydr. Polym. 115:785–803.

Hagan, M.T., Menhaj, M.B. (1994) Training feedforward networks with the Marquardt algorithm. IEEE Trans. Neural Netw. 5(6):2–6.

Harinath, E., Biegler, L., Dumont, G. (2013) Predictive optimal control for thermo-mechanical pulping processes with multi-stage low consistency refining. J. Process Control 23(7):1001–1011.

Harinath, E., Biegler, L.T., Dumont, G.A. (2011) Control and optimization strategies for thermo-mechanical pulping processes: Nonlinear model predictive control. J. Process Control 21(4):519–528.

Hong, G., Ma, C., Chen, H., Chuang, K., Chang, C., Su, T. (2011) Energy flow analysis in pulp and paper industry. Energy 36(5):3063–3068.

Hubbe, M.A., Heitmann, J.A. (2007) Review of factors affecting the release of water from cellulosic fibers during paper manufacture. BioResources 2(3):500–533.

International Energy Agency. Tracking industrial energy efficiency and CO2 emissions. OECD/IEA, Paris, 2007.

Jang, H., Topal, E. (2013) Optimizing overbreak prediction based on geological parameters comparing multiple regression analysis and artificial neural network. Tunn. Undergr. Space Technol. 38:161–169.

Kerekes, R. (1990) Characterization of pulp refiners by a C-factor. Nord. Pulp Pap. Res. J. 5(1):3–8.

Kerekes, R., Olson, J. (2003) Perspectives on fibre length reduction in refining. In: Proceedings of 7th PIRA international refining conference & exhibition.

Li, B., Li, H., Zha, Q., Bandekar, R., Alsaggaf, A., Ni, Y. (2011) Review: Effects of wood quality and refining process on TMP pulp and paper quality. BioResources 6(3):3569–3584.

Marquardt, D. (1963) An algorithm for least-squares estimation of nonlinear parameters. SIAM J. Appl. Math. 11(2):431–441.

Musavi, M.T., Coughlin, D.R., Qiao, M. (1995) Prediction of wood pulp K with radial basis function neural network. In: Proceedings of ISCAS'95 – International Symposium on Circuits and Systems. pp. 1716–1719.

Oliveira, M., Almada-lobo, B. (2012) Integrated pulp and paper mill planning and scheduling. Comput. Ind. Eng. 63(1):1–12.

Paradis, M.A., Genco, J.M., Bousfield, D.W., Hassler, J.C. (2002) Determination of drainage resistance coefficients under known shear rate. Tappi J. 1(8):12–18.

Perrusquía, A., Yu, W. (2021) Neurocomputing Identification and optimal control of nonlinear systems using recurrent neural networks and reinforcement learning: An overview. Neurocomputing 438:145–154.

Persson, J., Berntsson, T. (2009) Influence of seasonal variations on energy-saving opportunities in a pulp mill. Energy 34(10):1705–1714.

Pulkkinen, I. From eucalypt fiber distributions to technical properties of paper. Aalto University, Espoo, 2010.

Sainlez, M., Heyen, G. (2013) Comparison of supervised learning techniques for atmospheric pollutant monitoring in a Kraft pulp mill. J. Comput. Appl. Math. 246:329–334.

Sandberg, C., Ferritsius, O., Ferritsius, R. (2021) Mechanical pulping Energy efficiency in mechanical pulping – definitions and considerations. Nord. Pulp Pap. Res. J. 36(3):425–434.

Schmidhuber, J. (2015) Deep learning in neural networks: An overview. Neural Netw. 61:85–117.

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

Seth, R., Page, D. (1988) Fiber properties and tearing resistance. Tappi J. 71(2):103–107.

Simula, O., Alhoniemi, E. (2006) SOM based analysis of pulping process data. In: International Work-Conference on Artificial Neural Networks. pp. 567–577.

Talebjedi, B., Behbahaninia, A. (2021) Availability analysis of an Energy Hub with CCHP system for economical design in terms of Energy Hub operator. J. Build. Eng. 33:101564.

Talebjedi, B., Khosravi, A., Laukkanen, T., Holmberg, H., Vakkilainen, E., Syri, S., (2020) Energy Modeling of a Refiner in Thermo-Mechanical Pulping Process Using ANFIS Method.

Energies 13(19):5113. https://doi.org/10.3390/en13195113.
Talebjedi, B., Laukkanen, T., Holmberg, H., Vakkilainen, E., Syri, S., (2021) Energy Efficiency Analysis of the Refining Unit in Thermo-Mechanical Pulp Mill. Energies 14(6):1664. https://doi.org/10.3390/en14061664.
Talebjedi, B., Laukkanen, T., Holmberg, H., Vakkilainen, E., Syri, S. (2021) Energy simulation and variable analysis of refining process in thermo-mechanical pulp mill using machine learning approach. Math. Comput. Model. Dyn. Syst. 27(1):562–585.
Tian, H., Prakash, J., Zavala, V.M., Olson, J.A., Gopaluni, R.B. (2020) A tractable approximation for stochastic MPC and application to mechanical pulping processes. Comput. Chem. Eng. 141:106977.
Wang, Y., Elhag, T. (2007) A comparison of neural network, evidential reasoning and multiple regression analysis in modelling bridge risks. Expert Syst. Appl. 32(2):336–348.