Proceedings of the 14th International Symposium on Process Systems Engineering – PSE 2021+
June 19-23, 2022, Kyoto, Japan

Moisture Estimation in Woodchips Using IIoT Wi-Fi and Machine Learning Techniques

Kerul Suthar[a], Q. Peter He[a*]

[a]*Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA*

qhe@auburn.edu

Abstract

For the pulping process in a pulp & paper plant that uses woodchips as raw material, the moisture content (MC) of the woodchips is a major process disturbance that affects product quality and consumption of energy, water, and chemicals. Existing woodchip MC sensing technologies have not been widely adopted by the industry due to unreliable performance and/or high maintenance requirements that can hardly be met in a manufacturing environment. To address these limitations, we propose a non-destructive, economic, and robust woodchip MC sensing approach utilizing channel state information (CSI) from industrial Internet-of-Things (IIoT) based Wi-Fi. While these IIoT devices are small, low-cost, and rugged to stand for harsh environment, they do have their limitations such as the raw CSI data are often very noisy and sensitive to woodchip packing. Thus, direct application of machine learning (ML) algorithms leads to poor performance. To address this, statistics pattern analysis (SPA) is utilized to extract physically and statistically meaningful features from the raw CSI data, which are sensitive to woodchip MC but not to packing. The SPA features are then used for developing multiclass classification models as well as regression models using various linear and nonlinear ML techniques to provide potential solutions to woodchip MC estimation for the pulp and paper industry.

Keywords: systems engineering, machine learning, feature engineering, channel state information, IIoT sensors.

1. Introduction

The US pulp and paper industry ranks the third in energy consumption among US industries. The pulping process, which converts woodchips into pulp by displacing lignin from cellulose fibers, is one of the most energy intensive processes and has been identified as a major opportunity to improve energy productivity and efficiency of the industry (Brueske et al., 2015). Currently, vast majority of the US pulp is produced by chemical pulping processes and most of them utilize continuous Kamyr digesters. For Kamyr digesters, the incoming woodchip moisture content (MC) is a major disturbance that affects the cooking performance.

Currently, the woodchip MC is not measured in real-time due to the lack of affordable, reliable, and easy-to-maintain sensors. As a result, the performance of existing control solutions is often unsatisfactory and process engineers often overcook the woodchips to ensure pulp quality, which results in significant loss of pulp yield, overuse of heat/energy and chemicals. Chemical overuse also adds burdens to the downstream processes, such as washing and evaporation, and results in increased energy and chemical usages for downstream processes as well. To address this need, this work proposes a non-destructive, economic, and robust approach based on 5 GHz IIoT short-range Wi-Fi and use channel state information (CSI) to estimate MC in woodchips. Both classification and regression techniques are studied for MC estimation. For classification, we investigate linear discriminant analysis (LDA), support vector machine (SVM), artificial neural network (ANN), bagging with LDA, and ensemble boosting XGBoost. For regression, we study ANN, k-nearest neighbor regression (KNNR), Gaussian process regression (GPR), and support vector regression (SVR) with radial basis function (RBF) kernel.

The remainder of this work is organized as follows: Section 2 describes the experimental setup and software tools used in this study, as well as the features proposed and the modeling techniques utilized in this work. Section 3 presents results and discussions of this work, and Section 4 draws conclusions.

2. Data collection and feature engineering

2.1. Channel state information for moisture estimation

Using Wi-Fi cards such as IWL5300, it is convenient to collect CSI measurements that record the channel variation during propagation of wireless signals. After being transmitted from a source, the wireless signal is expected to experience impairments caused by obstacles before the signal reaches the receiver. CSI can reflect indoor channel characteristics such as multipath effect, shadowing, fading, and delay. In this work, we collect CSI using CSItool, which is built on IWL5300 NIC using a custom modified firmware and open-source Linux wireless drivers. The channel response of the i^{th} subcarrier can be represented as:

$$CSI_i = |CSI_i| \exp\{\angle CSI_i\} \quad (1)$$

where $|CSI_i|$ is the amplitude and $\angle CSI_i$ is the phase response of the i^{th} subcarrier.

2.2. Data description

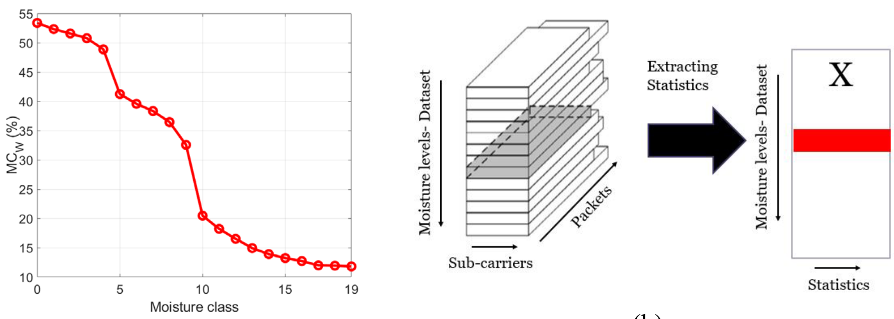
In this work, data are collected for 20 different MC classes or levels ranging from 53.39% to 11.81% on the wet basis (see Eqn (2)). A single antenna is used on the transmitter side which is configured in injection mode to send CSI and 3 antennas are used on the receiving side to take advantage of diversity. Woodchips are places in an airtight container between the transmitter and receiver to collect data. 10,000 packets are sent from the transmitter to the receivers for each sample collection. Total mass (m_T) is measured during each experiment and oven drying method was performed after all experiments were conducted to determine the oven dry weight (m_D). m_T and m_D are then used to determine the mass of water (m_W) and MC as the following.

$$MC = \frac{m_W}{m_T} \times 100\% = \frac{m_W}{m_W + m_D} \times 100\% \quad (2)$$

The 20 different MC levels are plotted in Figure 1(a), which shows that MC levels are narrowly separated at the high MC region and even more so at the low MC region. The minimum difference between MC levels is 0.05%, which is more than sufficient for pulping process optimization and control.

2.3. Methodology and feature engineering

To address the shortcoming of raw CSI features that lead to poor classification and prediction performance, in this work, statistics pattern analysis (SPA) is utilized to generate more robust and predictive features. In SPA, the statistics of the process variables, instead of process variables themselves, are used for modeling. This is based on the hypothesis that these statistics are sufficient and even better in capturing process characteristics than original process variables. This hypothesis has been supported in various applications, including fault detection (He et al., 2019; He & Wang, 2011, 2018; Wang & He, 2010), fault diagnosis (He & Wang, 2018), and virtual metrology or soft sensor (Shah et al., 2019, 2020; Suthar et al., 2019). SPA is selected in this work to extract robust and predictive features from raw CSI data. It is worth noting that SPA does not require preprocessing of the CSI data (e.g., outlier detection and handling, noise removal/reduction) that has been required in previous studies (Hu et al., 2019; Yang et al., 2018). A schematic for SPA based feature engineering is shown in Figure 1 (b). After a deeper exploration of candidate features and statistics, mean difference of consecutive subcarrier in CSI amplitude are chosen which leads to 87 features considering all 3 antennas on the receiving side.



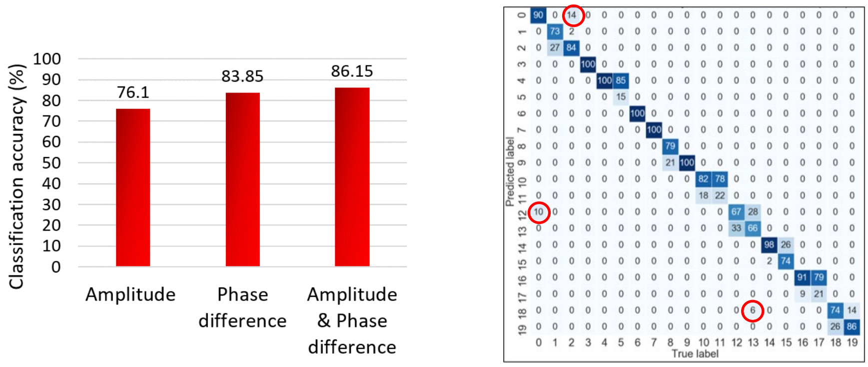
(a) (b)

Figure 1 (a) 20 different moisture levels tested in this work; (b) SPA based feature engineering for MC estimation

3. Results and discussion

In this work, we conduct investigations from three perspectives: (1) comparing raw CSI data vs engineered features; (2) comparing the performance of different classification approaches; and (3) comparing the performance of different regression approaches. For each model, 9 samples are randomly selected as training samples from 10 shuffled samples at the same MC level for each of the 20 MC levels, which results in 180 training samples. The remaining shuffled sample for each of the MC levels is used for testing. In this work we use Monte Carlo validation and testing (MCVT) procedure 100 times for performance comparison. To assess various classification approaches, the mean and standard deviation of classification accuracy of the 100 MCVT simulations are reported. For regression approaches, the mean and standard deviation of root mean square error (RMSE) for the 100 MCVT simulations are reported.

First, raw CSI data are used for MC level classification. The results are similar across different classification techniques. Due to limited space, only results from LDA are discussed here. Figure 2 (a) shows the overall classification accuracy of all classes when the raw CSI data were used. The comparison indicates that LDA classifier using both amplitude and phase difference performs the best with 86.15% classification accuracy, followed by LDA classifier using phase difference with 83.85% classification accuracy, while the LDA classifier using amplitude alone results in the lowest classification



(a) (b)

Figure 2 (a) Overall classification accuracy using different raw CSI data with LDA classifier based on 100 Monte Carlo runs. (b) Classification confusion matrix of 100 MCVT when both amplitude and phase difference are used. The far-off misclassifications (*i.e.*, the predicted class differs from the true class by more than one MC level) are highlighted by red circles.

accuracy of 76.10%. Figure 2 (b) plots the confusion matrix for the LDA classifier using both CSI amplitude and phase difference, which allows us to dig deeper into the classification results. As can be seen from Figure 2 (b), classification accuracy of individual classes ranges from 15% to 100%. It can also be seen that classification accuracy alone is not a good performance indicator. For example, the far-off misclassifications (i.e., the predicted class of a sample is off its true class by more than one level) will have worse consequences than the nearest-neighbor misclassifications (i.e., the predicted class is off true class by one level, either above or below) if they were used to control the white liquor usage or digester temperature. It can be seen from Figure 2 (b) that the classification results using raw CSI data are poor as there are samples misclassified far off their true classes. There are totally 478 misclassified samples, of which 30 are far-off misclassifications (highlighted by red circles in Figure 2 (b)). Also, the overall classification accuracy is not satisfactory.

Next the 87 rationally engineered features (i.e., the mean difference of consecutive subcarrier in CSI amplitude) are used for MC level classification and the results are summarized in Table 1. The classification accuracies shown in Table 1 indicate that all methods perform well with higher than 95% classification accuracy. The significantly improved performance compared to that of the raw CSI data demonstrates that the engineered features are more informative and characterize the MC in woodchips far better than the raw CSI data. Among all classification methods studied in this work, the bagging LDA performs the best with 98.75% average classification accuracy. The standard deviation of its classification accuracy is the lowest of 2.29%, indicating the

Table 1 Classification accuracy using engineered features

Method	Classification Accuracy	
	Mean	Std. dev.
SVM	95.50	3.79
ANN	95.85	4.15
XGBoost	96.40	3.70
LDA	97.55	2.89
Bagging (LDA)	98.75	2.29

bagging LDA is also the most robust or consistent classifier among all methods studied in this work.

Finally, we study different regression methods for MC estimation. When raw CSI data are used, all regression methods perform poorly, similar to the classification results when the raw CSI data are used. Due to limited space, they are not shown here. When the same 87 engineered features are used for regression-based MC estimation, a well-tuned ANN with two hidden layers outperforms other regression-based approaches as shown in Table 2. KNNR performs comparable to ANN while GPR and SVR with RBF kernel have relatively higher average RMSE's for 100 MCVT simulations.

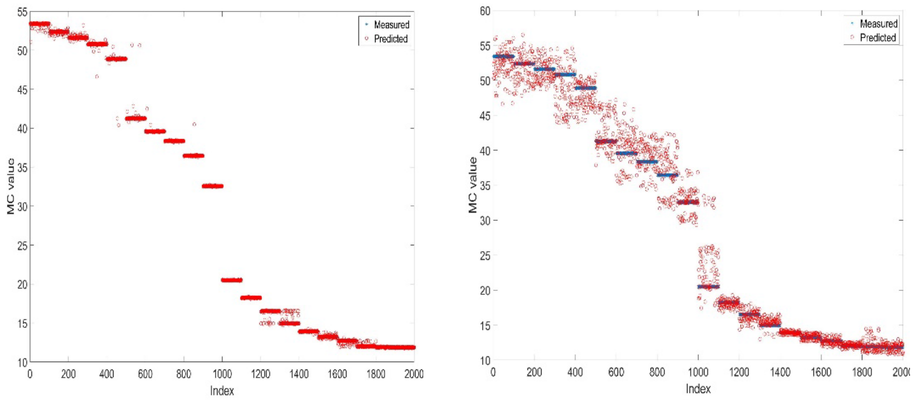
Table 2 Regression for MC estimation using engineered features

Method	RMSE	
	Mean	Std. dev.
ANN	0.51	0.3921
KNNR	0.6573	0.5055
GPR	1.9223	0.5714
SVR(RBF)	2.0179	0.523

Figure 3 shows the measured vs predicted MC values for ANN and SVR(RBF). It can be seen from Figure 3(a) that the ANN predicted MC values agree very well with the actual or measured MC values. In comparison, while SVR captures the MC trend, its predictions have much higher standard deviation compared to ANN. It is worth noting for all the above-mentioned results, the models and their hyperparameters were tuned using random search followed by Bayesian optimization (Bergstra & Bengio, 2012).

4. Conclusions

In this work, we investigate the potential of an IIoT short-range Wi-Fi based woodchip MC sensing technology to overcome some limitations of the existing technologies. The proposed technology takes the advantages of IIoT devices (e.g., toughness, connectivity,



(a) (b)

Figure 3 Measured vs predicted MC by (a) ANN and (b) SVR(RBF) low-cost, small-size, etc.), while overcoming their shortcomings (e.g., the machine learning challenges of messy big data) through SPA-based feature engineering. We investigate the use various classification and regression approaches for the estimation of 20 different moisture levels. We demonstrate that with SPA-based features, all classification approaches studied in this work can successfully classify 20 different MC levels, some of which are separated by small margins. We also investigate the use of different regression approaches for continuous MC estimation. While SVR and GPR capture the trend of measured MC values but with relatively high RMSE's, methods including ANN and KNNR predict the moisture levels accurately. The relationship between the CSI and woodchip MC is very complex, which requires further work to get a better understanding of this relationship for further improvement of this work.

References

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*.

Brueske, S., Kramer, C., & Fisher, A. (2015). *Bandwidth Study on Energy Use and Potential Energy Saving Opportunities in US Pulp and Paper Manufacturing*. Energetics.

He, Q. P. P., & Wang, J. (2018). Statistical process monitoring as a big data analytics tool for smart manufacturing. *Journal of Process Control*, *67*, 35–43.

He, Q. P., & Wang, J. (2011). Statistics pattern analysis: A new process monitoring framework and its application to semiconductor batch processes. *AIChE Journal*, *57*(1), 107–121.

He, Q. P., & Wang, J. (2018). Statistical process monitoring as a big data analytics tool for smart manufacturing. *Journal of Process Control*, *67*, 35–43.

He, Q. P., Wang, J., & Shah, D. (2019). Feature Space Monitoring for Smart Manufacturing via Statistics Pattern Analysis. *Computers & Chemical Engineering*, *126*, 321–331.

Hu, P., Yang, W., Wang, X., & Mao, S. (2019). MiFi: Device-free wheat mildew detection using off-the-shelf wifi devices. *2019 IEEE Global Communications Conference, GLOBECOM 2019 - Proceedings*.

Shah, D., Wang, J., & He, Q. P. (2019). A feature-based soft sensor for spectroscopic data analysis. *Journal of Process Control*, *78*, 98–107.

Shah, D., Wang, J., & He, Q. P. (2020). Feature Engineering in Big Data Analytics for IoT-Enabled Smart Manufacturing–Comparison between Deep Learning and Statistical Learning. *Computers & Chemical Engineering*, 106970.

Suthar, K., Shah, D., Wang, J., & He, Q. P. (2019). Next-generation virtual metrology for semiconductor manufacturing: A feature-based framework. *Computers and Chemical Engineering*, *127*, 140–149.

Wang, J., & He, Q. P. (2010). Multivariate Statistical Process Monitoring Based on Statistics Pattern Analysis. *Industrial & Engineering Chemistry Research*.

Yang, W., Wang, X., Cao, S., Wang, H., & Mao, S. (2018). Multi-class wheat moisture detection with 5GHz Wi-Fi: A deep LSTM approach. *Proceedings - International Conference on Computer Communications and Networks, ICCCN*.