Creating Adaptive Predictions for Packaging Critical Quality Parameters Using Advanced Analytics and Machine Learning

Cydney Rechtin, Anthony Lewis and Beth Ann Zarko, *Solenis*
Chitta Ranjan, *ProcessMiner*

PaperCon 2019 | May 2019

Creating Adaptive Predictions for Packaging Critical Quality Parameters Using Advanced Analytics and Machine Learning

Cydney Rechtin (Solenis)**, Chitta Ranjan** (ProcessMiner)**, Anthony Lewis** (Solenis)**, Beth Ann Zarko** (Solenis)

ABSTRACT

Packaging manufacturers are challenged to achieve consistent strength targets and maximize production while reducing costs through smarter fiber utilization, chemical optimization, energy reduction, and more. With innovative instrumentation readily accessible, mills are collecting vast amounts of data that provide them with ever increasing visibility into their processes. Turning this visibility into actionable insight is key to successfully exceeding customer expectations and reducing costs.

Predictive analytics supported by machine learning can provide real-time quality measures that remain robust and accurate in the face of changing machine conditions. These adaptive quality "soft sensors" allow for more informed, on-the-fly, process changes, fast change detection, and process control optimization without requiring periodic model tuning.

The use of predictive modeling in the paper industry has increased in recent years, however, little attention has been given to packaging finished quality. The use of machine learning to maintain prediction relevancy under ever-changing machine conditions is novel. In this paper, we demonstrate the process of establishing real-time, adaptive quality predictions in an industry focused on reel-to-reel quality control and we discuss the value created through the availability and use of real-time critical quality.

INTRODUCTION

The most common method for determining finished product quality is periodic lab testing. Even if a lab is capable of testing every paper reel produced, the time between test results is significant because of the time it takes to produce a reel of paper and the laboriousness of sample preparation, testing, and data recording. Additionally, making machine control decisions to meet quality standards based on periodic lab tests has many inherent limitations, such as low frequency of data, lack of MD (machine direction) data, and the existence of both measurement error and human error in lab tests. Acquiring quality data after the production of a reel of paper is too infrequent for the papermaker to proactively make process changes, leading to increased off-quality product and material and energy waste. Testing the quality of a single MD location of the paper reel is an incomplete representation of the product and can lead to downstream issues, such as low converting efficiency and increased customer complaints. Using a lab measurement with both high measurement and human error to make machine control decisions induces unwanted variability into the papermaking process.

Real-time predictions provide high-frequency data every 15–30 seconds, thereby creating a machine direction quality profile. However, a prediction built without consideration of the unique process it is simulating may result in high prediction error. Using an erroneous prediction to make control decisions can induce more process variability than using the lab test. Thus, a valuable prediction must have no more error than the lab test it is simulating. A useful prediction must use methodologies that reflect the process being simulated.

In this paper, we demonstrate the process for establishing accurate, real-time, and adaptive paper finished quality predictions for strength parameters such as Mullen, Ring Crush, and STFI. We discuss the structure of paper machine quality and process data, highlight appropriate methods for creating quality predictions (mentioning novel techniques when appropriate), define simple metrics to measure prediction accuracy, describe the importance of applying machine learning to maintain prediction accuracy, and exemplify the value that can be created through the availability and use of real-time strength predictions in packaging manufacturing.

PROCESS FOR CREATING ADAPTIVE AND REAL-TIME PREDICTIONS

Data Collection

The process of establishing accurate, real-time, and adaptive paper quality predictions begins by collecting historical quality and process data. To build a valuable predictive model, finished quality data and meaningful process data must be collected for a period of 6–12 months. The recent rise of innovative instrumentation on paper machines has allowed for vast data collection; however, not all available information is useful for predicting paper quality. Figure 1 depicts important data sources for building paper strength predictions.

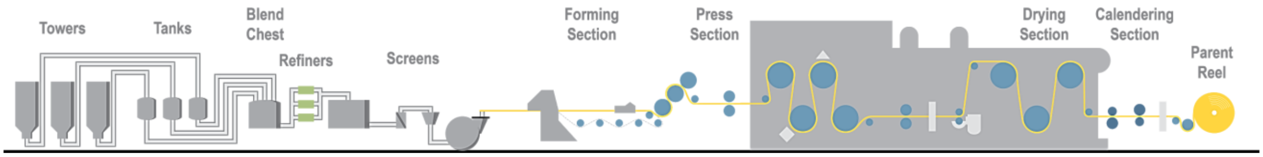

Figure 1. Important data sources for paper strength predictions

Data Survey

Surveying paper quality and process data consists of analyzing the structure and type of data, characterizing the quality variable, called the response, and the process data, called the predictors, and discovering the nature of the relationships between the response and the predictors.

The type of data collected from a paper manufacturing operation is called time series cross-sectional (TSCS). TSCS data is characterized by encompassing many variables collected together over time. The sequential nature of TSCS data requires special techniques that account for the dynamics of the data. TSCS data often suffers from autocorrelation, a correlation between elements in the time series [1]. Ignoring autocorrelation in the data leads to uninterpretable and ineffective predictive models. Thus, techniques that remove the effects of autocorrelation should be used when analyzing paper machine quality and process data.

The response variable must exhibit near-normal behavior for proper interpretation of prediction confidence intervals. Because a single paper machine produces many grades with various specifications, data surveying on a grade-by-grade basis may be required. Normal or near-normal behavior can be defined as approximately following a normal distribution. As Figure 2 shows, a normal distribution is symmetric, unimodal, asymptotic, and the mean, median, and mode are equal [2]. Meaningful predictors must exhibit moderate variability. Additionally, a healthy predictor is monitored no less frequently than the desired real-time prediction frequency and has limited blank, zero, and outlier values.

The response-predictor relationships in papermaking can be either linear or nonlinear, including polynomial, exponential, power, or binomial relationships, among others. It is important to understand the true response-predictor relationships to avoid creating an oversimplified predictive model. For example, it is well-known that refining energy and strength are positively related up to a point where, as Figure 3 shows, increasing refining energy decreases strength [3]. This relationship is nonlinear and should be treated as such in predictive modeling. Using graphs such as the one in Figure 3 can help to determine response-predictor relationships.

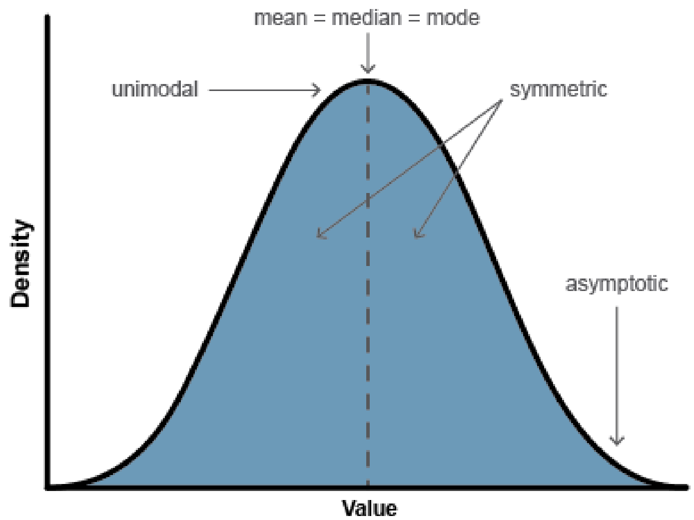

Figure 2. A normal distribution: symmetric, unimodal, asymptotic, and equal mean, median, and mode

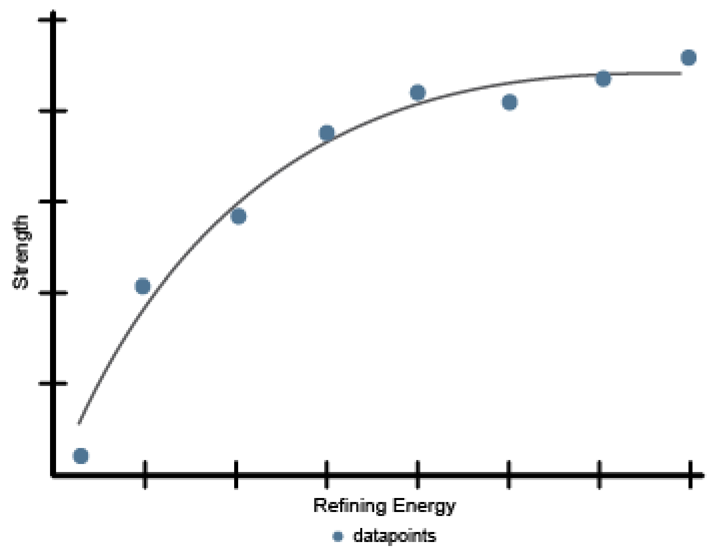

Figure 3. Refining energy vs. strength relationship exhibiting nonlinear behavior

Data Cleaning

Once properly surveyed, historical data is cleaned to support predictive model creation. Data cleaning is a preprocessing step that transforms raw data into a format that is acceptable for the type of analysis being conducted. Sampling the raw data to identify subsets based on paper grade, noise treatment, handling missing data, normalization, and feature extraction are some common data cleaning methods [4].

The data cleaning techniques used are unique to the paper operation being analyzed. While most paper machines make multiple grades, each grade will exhibit other exclusive characteristics that require a unique set of data cleaning processes. For example, response variables may need to be transformed mathematically to correct non-normality. Predictors with non-constant variability may also need to be transformed. Predictors and response variables with too few non-zero data points may require the application of missing data techniques such as removal or substitution. Outliers in both the predictor and response data should be investigated and, when appropriate, removed.

Data Mining

Data mining is the process of discovering relationships and patterns in the data. While initial response-predictor relationships may have been discovered in the data surveying step, data mining takes a deeper look at the multidimensional relationships between the response and predictors. A data mining technique such as multiple regression and causal networks can be used to identify relationships and patterns in papermaking data. Figure 4a,b highlights the importance of considering high-dimensional relationships. Figure 4a depicts a simple and well-known relationship between paper strength and basis weight. However, as Figure 4b shows, the true relationship is more complex because it is influenced by many secondary, tertiary, and quaternary relationships. Choosing a prediction technique that incorporates these multidimensional relationships will increase predictive power and accuracy.

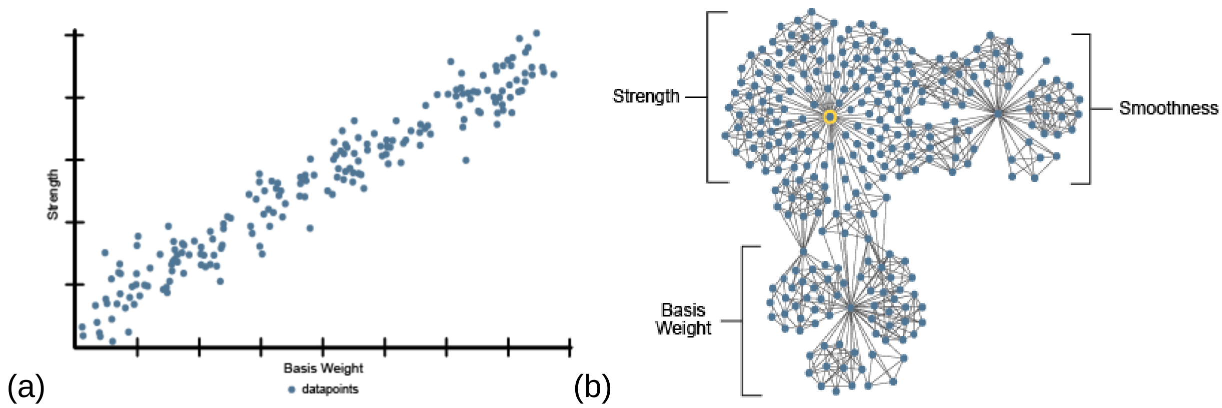

Figure 4. (a) Simple relationship between strength and basis weight; (b) Causal network showing multidimensional relationships between strength, basis weight, smoothness, and numerous other predictors

Developing the Predictive Model

After the historical data is collected and surveyed, the predictive model is built. Making paper quality predictions requires developing a supervised learning model, where historical data is used to "teach" the model based on past outcomes. However, the underlying process complexity requires a novel approach to developing a predictive model. Next, we explain the major challenges in predicting paper quality measurements in real time and highlight the novel approaches we have developed to overcome these challenges.

High-dimensionality. The number of sensors in a mill can range from hundreds to thousands. These sensors continuously collect data resulting in a high-dimensional and high-frequency data stream. Additionally, sensors are added and removed over time. This leads to temporally dynamic high-dimensional data.

Spatial and temporal dynamics. The variables in the data stream are spatially and temporally dynamic. Spatial dynamics are the correlations between variables measured at the same time, and temporal dynamics are the correlations within or between variables observed at different points in time. Additionally, these dynamics can be linear or nonlinear, which makes their estimation extremely difficult.

Measurement errors. The response variable, which is typically a quality measurement related to paper strength (e.g., Ring Crush, Mullen, etc.), is measured in a lab. These measurements commonly have a non-random error, often because of human data entry. Additionally, because they are measured in long intervals, it is difficult to identify and isolate the measurement error.

Observational data. Although experimental data is preferred to identify the causal relationships between variables, it is difficult to obtain because of the exceptionally high cost of experimentation in paper mills. Consequently, only observational data is available for modeling.

To overcome those challenges, a prediction system requires these salient features:

Regularization in spatial and temporal domain. Having too many variables in a prediction model results in poor accuracy. Therefore, regularization techniques, for example, L_0, L_1, or L_2, are often employed in high-dimensional processes [5]. Regularization yields a smaller set of the most important variables for prediction (variable selection). However, most conventional regularization techniques have two limitations, namely that they work in the spatial domain and on linear relationships between variables. We developed a nonlinear correlation estimator [6]. Subsequently, we developed a novel regularization approach that uses this estimator in the temporal space to overcome these limitations. Note that the correlation estimator can also find a linear correlation.

Additionally, to aid the variable selection, a causal relationship graph is required. Estimating this is rather straightforward with experimental data. However, because of the availability of only observational data we developed an approach using a constrained graphical lasso method.

Adaptive and evolving model. Paper manufacturing is a continuous process that changes over time. These changes are slow and sometimes abrupt. For example, a change in a raw material or the degradation of a machine part leads to a slow process shift. On the other hand, at or near the time of a paper break or grade change, the process rapidly experiences a shift. These scenarios lead to two considerations, namely the predictive model needs to be updated frequently and historical data quickly becomes obsolete. Thus, a continuously-evolving and adaptive model is required. The predictive model evolves over time as more data is collected. With new data streaming the old data slowly loses its relevance. The prediction system is built such that it automatically uses the knowledge gathered from the old data and combines it with the new data to construct an evolved (a relearned and more accurate) model.

Additionally, the prediction system self-adapts to a process change. For the model adaptation, we perform best prediction model selection, and its retuning, periodically and on-demand. The periodic runs are to ensure nothing is missed by the triggers. The on-demand is triggered automatically whenever a process shift is detected. Data with measurement errors are identified in real-time using F-tests and are isolated from the training data. Both of the above features are computationally intensive. To deploy them in real-time, we parallelized the processes through multiple processing units. The parallel processing enables the delivery of accurate predictions in real-time every 30 seconds.

Figure 5 shows the outcome of a predictive model that uses an off-the-shelf approach without considering the unique characteristics of the papermaking process compared with a model using a tailor-made, advanced approach such as the one discussed previously. The oversimplified model depicted in Figure 5a does not mirror the true process, while that of Figure 5b considers the complexity of the papermaking process and clearly delivers predictions that are more accurate.

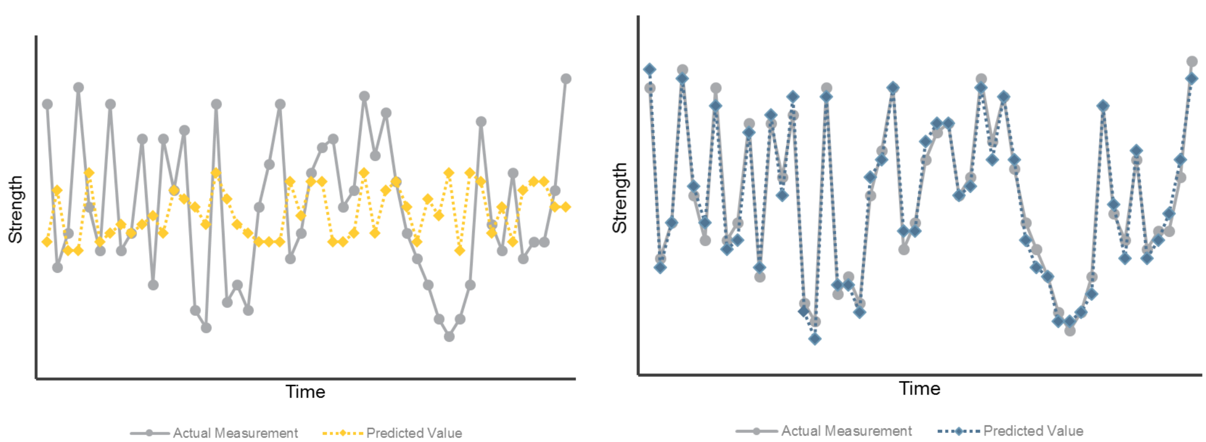

Figure 5. (a) Predictive model using poor prediction techniques that oversimplify the process; (b) Predictive model with the same data utilizing advanced techniques capable of handling paper machine process data

Predictive Model Accuracy

An accurate model provides predictions that are close to the actual value most of the time. That is, the absolute value of the prediction, minus the test value, should be close to zero. Prediction accuracy must also be measured in real-time to instill confidence in users acting on the predictions. One method to evaluate prediction accuracy in real time is to use a control chart to monitor the difference between actual and predicted values.

The prediction at the time of a reel turn up is most comparable to the lab-tested sample; thus, those values can be compared. Variability in the prediction and the process will guide the width of the control limits, which will change over time. Figure 6 depicts an accurate, real-time prediction. Figure 6a shows that the prediction, indicated by the green line and surrounded by a grey confidence interval, remains close to the actual value, indicated by the black line. Figure 6b depicts a prediction accuracy control chart that uses a metric related to the difference between the prediction and the actual value. This value remains within its control limits, signifying an acceptably accurate prediction.

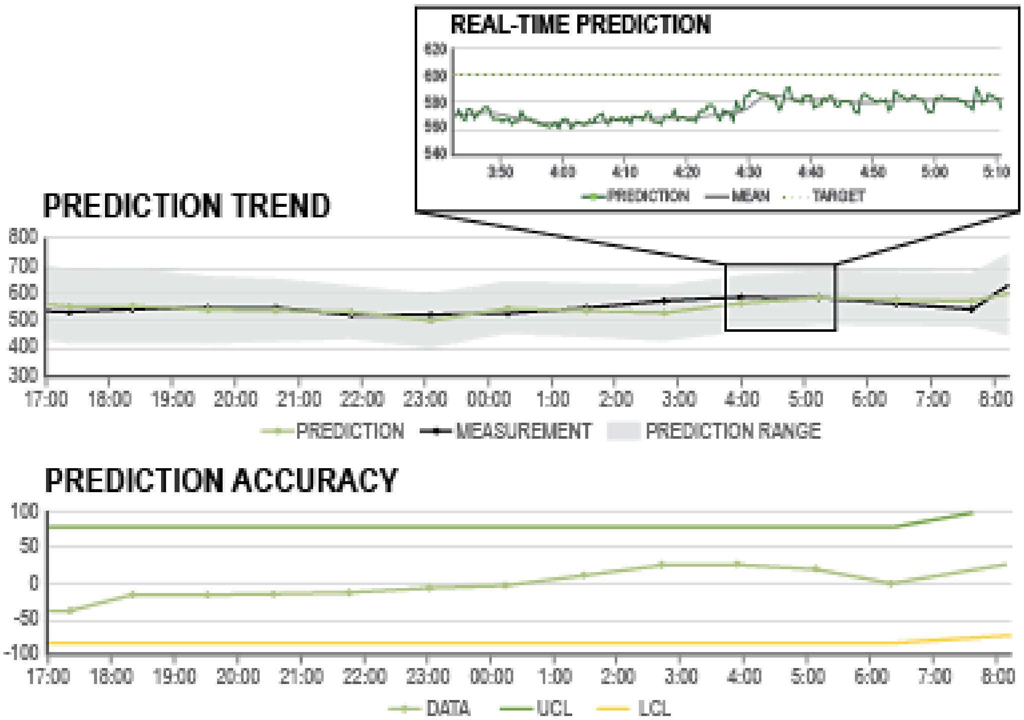

Figure 6. (a) Real-time prediction trend: Green points indicate predictions, grey band is the prediction confidence interval, black points indicate actual lab values; (b) Real-time prediction accuracy control chart where the values are a metric related to the difference between the prediction and actual value and the control limits are an acceptable range of variability for an accurate prediction

Application of Real-time Machine Learning

Although a predictive model that utilizes the correct techniques based on the type of data and the true relationships between the response and the predictors is required to make accurate paper machine quality predictions, the variables influencing the predictive model are always changing. Machine runnability, grade mix, fiber availability, chemistry usage, water characteristics, seasonal changes, operating crew, and upstream operations are among the many sources of variability in paper manufacturing. While some of these sources of variability may be directly captured in the predictive model, others are not. Figure 7 shows that without regular tuning, the predictive model quickly becomes irrelevant in a highly-variable application such as papermaking. However, manually tuning predictive models is time consuming and requires a subject matter expert. Instead, machine learning algorithms can be applied to constantly tune predictive models and preserve prediction accuracy in real time.

Predictive models utilizing machine learning adapt to changes in the process by learning from observations and interactions. This information drive shifts in the predictive model aimed at maintaining prediction accuracy. Adaptive predictive models do not need to be tuned manually due to the application of machine learning.

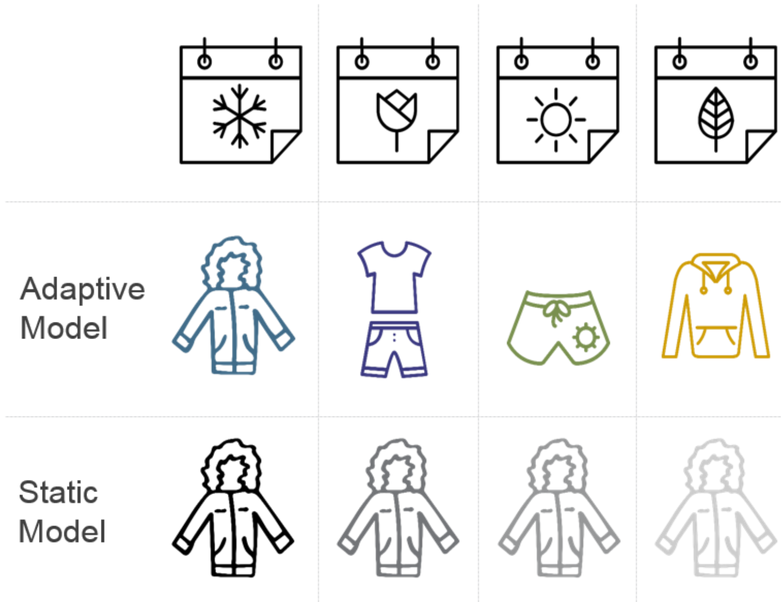

Figure 7. Depiction of prediction accuracy for static vs. adaptive models

Live Connection

After the predictive model is created and the machine-learning algorithms are in place for continuous tuning, the next step is to connect live streaming data. Figure 8 depicts the flow of information from the machine, through the live predictive model experiencing machine learning, and to the screen of an operator who can use the real-time predictions to make on-the-fly machine adjustments.

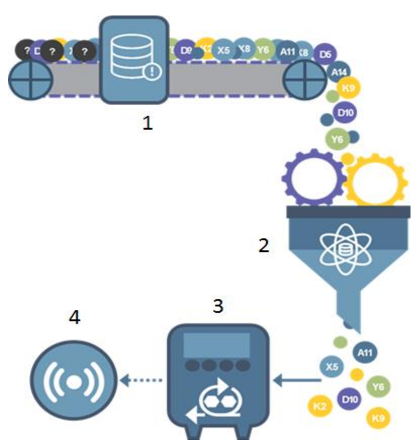

Figure 8. Flow of live machine data to real-time, adaptive predictive model outputs: (1) Real-time process data streams to model; (2) Model uses current conditions to make accurate predictions; (3) Machine learning algorithms tune model; and (4) Predictions are created and available for use

Summary

Figure 9 shows the steps in the process of establishing real-time, adaptive quality predictions on the paper machine,, including data collection, surveying, cleaning, and mining, followed by building the initial predictive model, ensuring modeling accuracy, applying machine learning to continuously update the model based on the current machine conditions, and connecting to live streaming data for real-time predictions. Next, we exemplify the value of utilizing real-time, adaptive predictions for process and quality optimization.

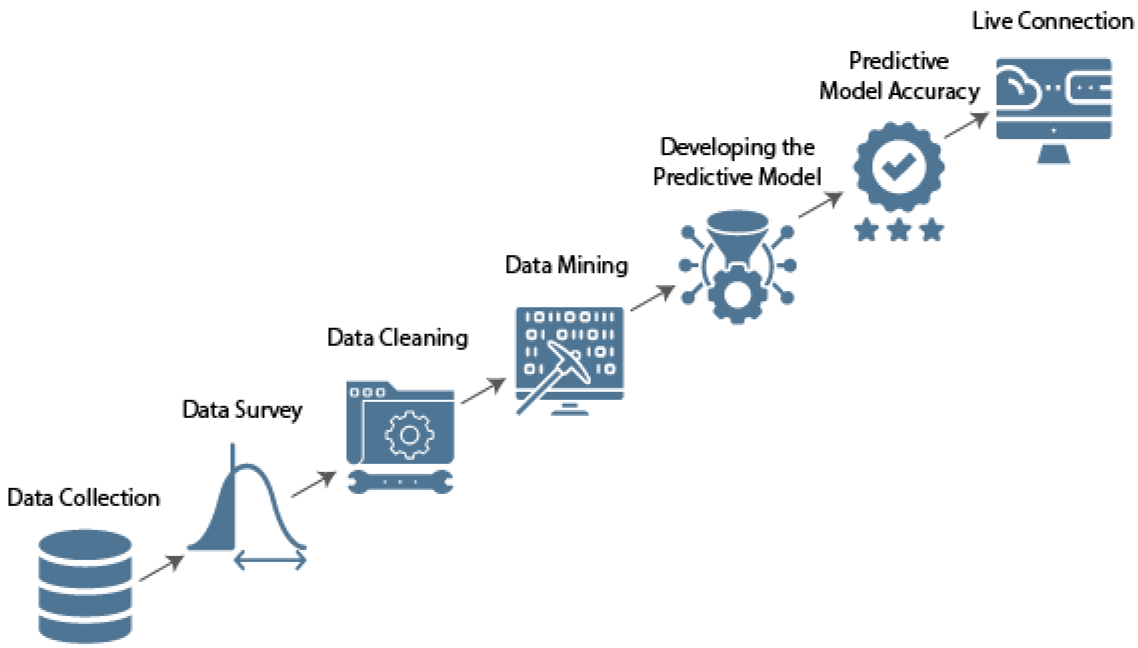

Figure 9. Summary of process for creating real-time and adaptive predictions

VALUE OF UTILIZING ADAPTIVE AND REAL-TIME QUALITY PREDICTIONS

Informed Decision-making

Packaging manufacturers use the results from periodic lab tests to make machine quality control decisions. These lab tests are infrequent and highly variable because of their manual nature and a machine direction strength profile is not generated. The availability of real-time predictions that tune themselves using machine learning provides information that paper machines currently do not have.

These predictions are generated using live process data; therefore, they do not suffer from the same sources of variation as do manual lab tests. The evaluation of a strength parameter is more accurate using a prediction when the prediction error is less than the error in the lab test. Error reduction due to the predictive model indicates that using the prediction to drive decision-making is superior to using a periodic lab test alone. Using an accurate, real-time prediction to inform on-the-fly process control decisions can translate to a reduction in process variability.

Figure 10 demonstrates a reduction in strength variability associated with using an accurate, real-time prediction to make control decisions. This variability reduction occurs because the prediction is inherently less erroneous than the corresponding lab test. Thus, when the prediction is used in place of the lab measurements to make quality control decisions, the system variability decreases and some of the inherent lab testing error is no longer transferred into the process.

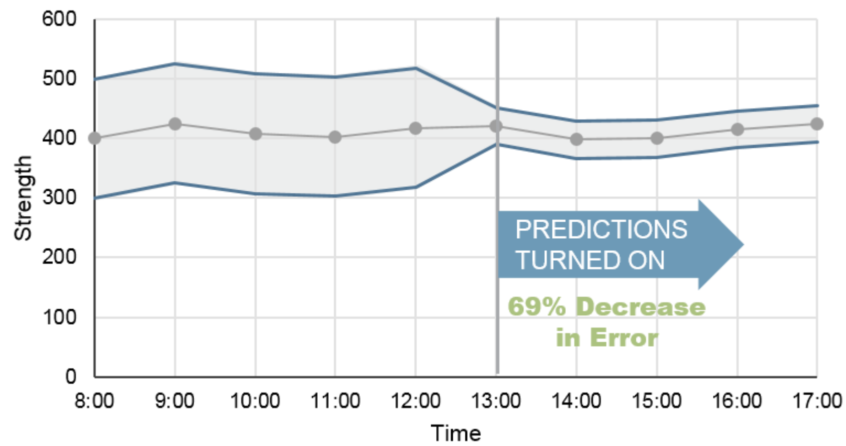

Figure 10. Error reduction achieved by utilizing real-time, adaptive predictions

Real-time prediction frequency is much greater than lab testing, in the range of 15–20 seconds compared with 45–90 minutes, respectively. This leads to a machine direction strength profile that offers operators the opportunity to confidently control their machine and make more frequent adjustments, thereby eliminating the need to overfeed expensive materials, such as fiber and chemicals, and reducing unnecessary costs.

Better fiber management can lead to significant savings: a one percent decrease in basis weight on a packaging machine producing 350,000 tons per year can save approximately $500,000. A one pound per ton decrease in dry strength chemical spend can save the same machine about $500,000. These savings are often attainable on a packaging machine that focuses on throughput at the cost of running slightly overweight or one that sets their chemical dosages higher than what is required to insure the product meets quality specification during infrequent process anomalies. Real-time predictions provide the information necessary to confidently eliminate excess. Figure 11 exemplifies the value created by eliminating excess raw material usage and increasing the production rate on a machine experiencing reduced process variability because of the use of real-time predictions.

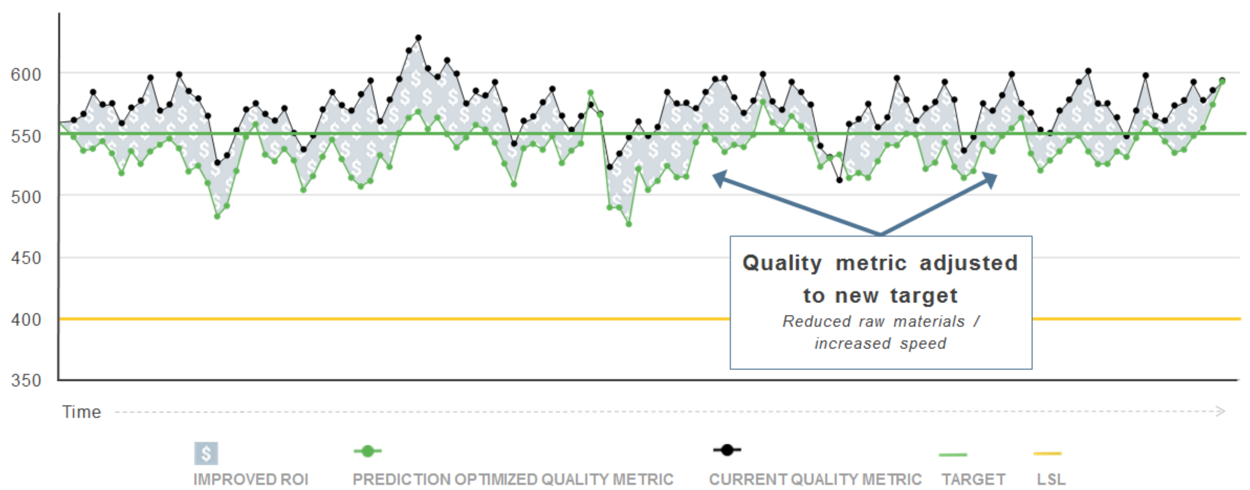

Figure 11. Value captured when utilizing a real-time prediction

Change Detection

Process shifts are common in paper manufacturing and sometimes go unnoticed until the reel of paper is tested in the lab or the downstream customer is converting the material. Real-time predictions help to identify process upsets as they are happening rather than after the fact.

One way to detect process change is to monitor the process variables, called predictors, which are most influential to the predictive model. An adaptive predictive model will have an ever-changing set of predictors depending on the current state of the process. This leads to two types of insights, namely monitoring changes in the set of predictors and monitoring changes in the individual predictors. When the set of predictors changes, this indicates a change in what is informing the prediction. Changes or patterns in individual predictors can allude to consequential changes and patterns in the predicted quality variable and can indicate an area of interest when investigating a process upset.

For example, a strength prediction may not use the process variable corresponding to dry strength dosage until that dosage increases above some threshold. If the set of predictors changes to include dry strength dosage, this indicates that it is now more influential to the strength prediction than it was before. Figure 12 depicts a scenario where a mechanical failure associated with a single predictor led to an uncommon trend in the predicted variable. A mill using real-time strength predictions noticed a disturbance in the form of a cyclical swing. They investigated the predictors at the time of the upset and discovered that the acid pump was also swinging irregularly. Additional inspection of the acid system revealed underlying mechanical issues and once corrected, the prediction and the acid pump stabilized.

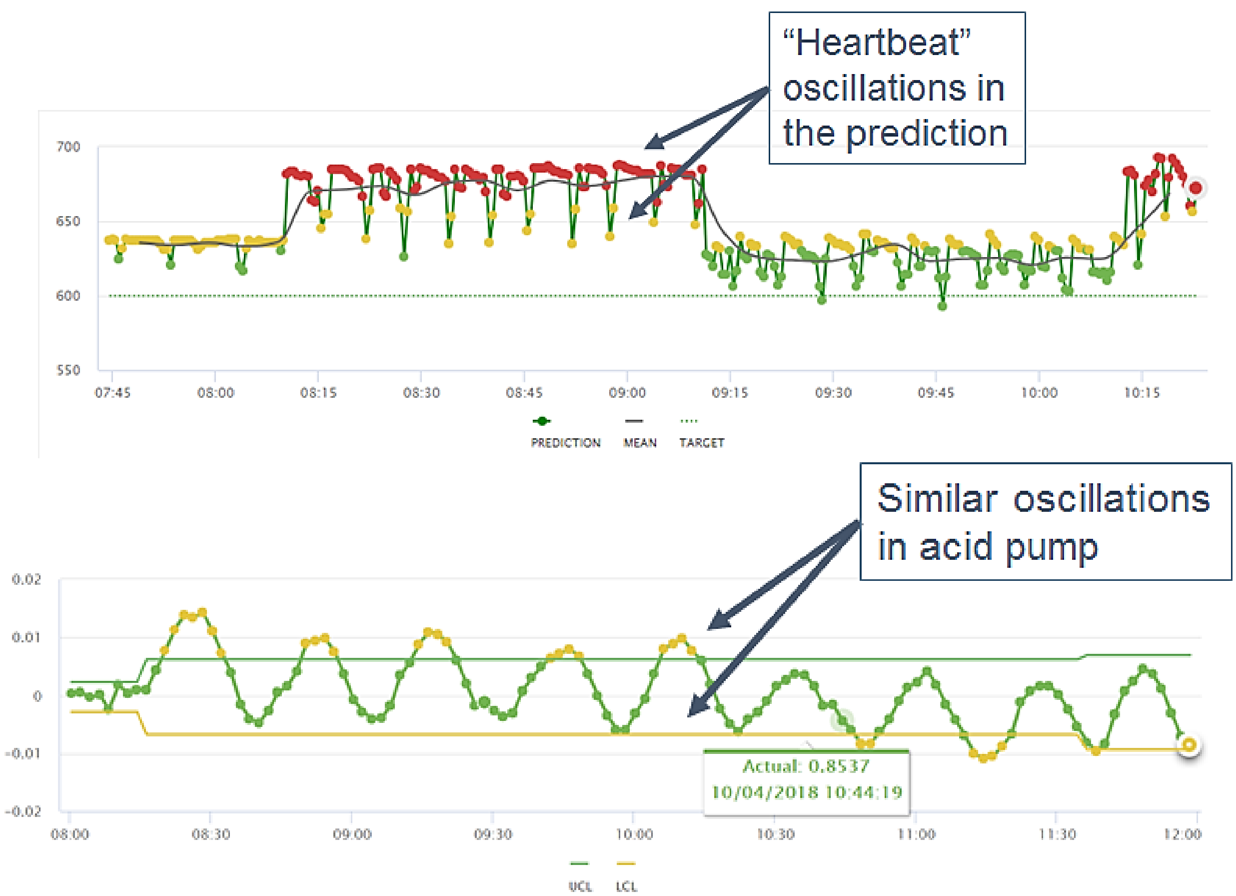

Figure 12. Similar oscillations in the strength prediction and an individual predictor (acid pump) leading to maintenance in the acid system

Process and prediction changes also can be detected by evaluating the stability of a prediction. Figure 13 depicts a scenario where a mill using a strength prediction ran a new chemical trial aimed at improving strength. This was a new variable in the process, and thus the predictive model was not utilizing it to create predictions. The strength chemical positively influenced strength, identified by increased lab strength values; however, the prediction continued making predictions as though the new chemical were not in the process. Over a long enough period, an adaptive predictive model would learn to compensate for this difference through machine learning. However, at the time of the chemical trial, the strength prediction helped verify the impact of the new additive. If a once-stable prediction suddenly departs from the actual values, it could be an indication of an outside disturbance (such as the addition of a new chemistry).

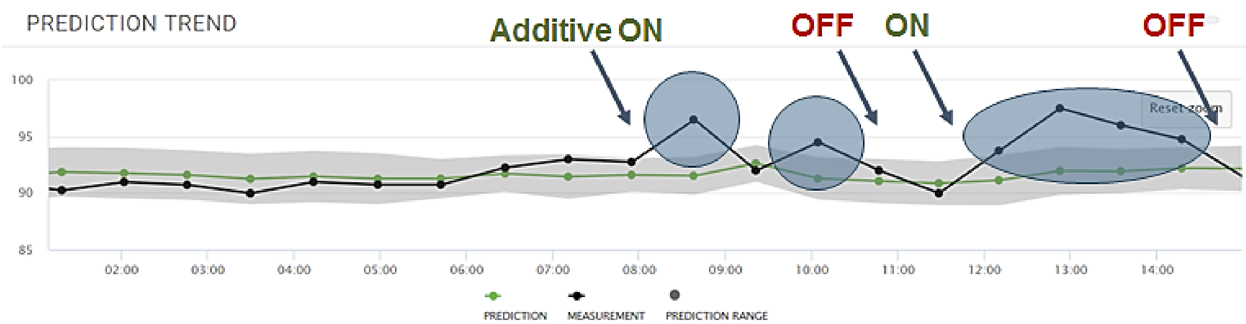

Figure 13. Prediction (green line) departing from actual values (black line) when a new chemical additive aimed at strength improvement was turned on and off

Tuning by Learning

In addition to improved quality control, variability reduction, and change detection, real-time and adaptive predictions benefit from the use of machine learning, which evaluates process changes and constantly updates the predictive model to ensure accuracy. This means human tuning is not required, alleviating the need for an employee dedicated to predictive model tuning.

Figure 14a demonstrates the effectiveness of machine learning on real-time quality predictions. Approximately five months of process data was used to create a strength predictive model. An additional one month of new data from the same process was used to simulate real-time strength predictions. This real-time data was looped in a simulator and, over time, the machine-learning algorithms continuously updated the model until the prediction perfectly matched the actual value. In Figure 14b, the prediction confidence interval (grey band) no longer exists because of the near-perfect prediction accuracy induced by automatic model tuning using machine learning over a long period. Although no real papermaking process will repeat itself exactly, this simulated example using real data indicates the ability of machine learning to maintain prediction accuracy in the face of process changes.

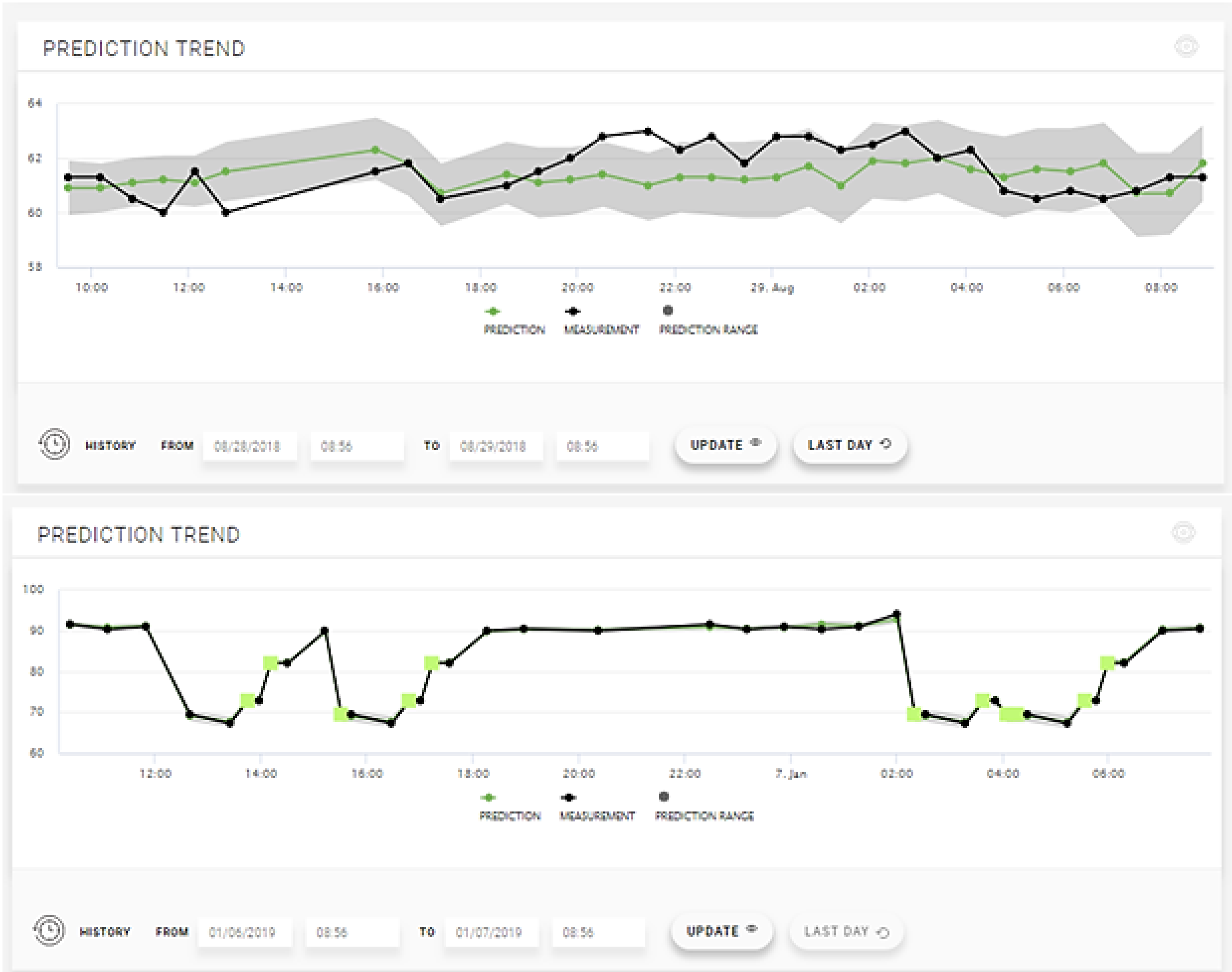

Figure 14. Effect of machine learning on prediction accuracy using papermaking process data: (a) Early stages of simulation; (b) Late stages of simulation resulting in near-perfect predictions

CONCLUSION

In this paper, we demonstrated the process of establishing real-time, adaptive quality predictions in an industry focused on reel-to-reel quality control. The predictive model building process includes data collection, surveying, cleaning, and mining followed by building the initial predictive model, ensuring modeling accuracy, applying machine learning to continuously update the model based on the current machine conditions, and connecting to live streaming data for real-time predictions. We also discussed the value of utilizing real-time, adaptive predictions for process and quality optimization in terms of informed decision-making, change detection, and tuning by machine learning.

REFERENCES

[1] Chatterjee, S., & Simonoff, J. S. (2013). Handbook of regression analysis. Hoboken, NJ: John Wiley & Sons.
[2] Hogg, R. V., McKean, J. W., & Craig, A. T. (2019). Introduction to mathematical statistics. Boston: Pearson.
[3] Biermann, C. J. (1996). Handbook of pulping and papermaking. San Diego: Academic Press.
[4] Osborne, J. W. (2013). Best practices in data cleaning: A complete guide to everything you need to do before and after collecting your data. Los Angeles: SAGE.
[5] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. Cambridge, MA: The MIT Press.
[6] ProcessMiner. (n.d.). ProcessMiner/nlcor. Retrieved from https://github.com/ProcessMiner/nlcor