Impact of TMP refining line interruptions and reject refiner operations on pulp and paper variability

ROBERT P. HARRISON, ALAIN A. ROCHE, AND PAUL R. STUART

ABSTRACT: This article develops a method for correlating TMP operations with pulp quality, and ultimately with paper quality, by focusing on process fundamentals such as specific energy and refining intensity. The case study is an Eastern Canadian newsprint mill that experiences variability in paper strength and porosity. Frequent interruptions in the four refining lines greatly affect the reject refining specific energy and other key parameters, many of which are not measured directly and must be calculated from other variables. Using multivariate analysis and other statistical tools, it was possible to link pulp quality back to the TMP and reject refining operations, taking into account the number of lines in operation, plate age, and process lags. Furthermore, it was possible using multivariate analysis (MVA) models to correlate roughly half the variability in final paper quality with the refining operations.

Application: This study presents a straightforward method for combining individual TMP operations, including reject refining, into a single statistical model for explaining quality fluctuations in pulp and final newsprint.

Canadian TMP newsprint mills currently face a double challenge: varying chip quality, and an increasingly demanding clientele who expect a better and more consistent product. Companies are now less vertically integrated and have less direct control over the wood supply. Many chips are now coming from outside sawmills, where they are seen as a secondary by-product, with little or no attention paid to quality assurance. Unfortunately, many existing TMP mills have built-in characteristics, such as limited chip storage, that make final newsprint quality highly susceptible to disturbances. This situation only exacerbates other economic factors that are making it increasingly difficult for North American newsprint mills to compete internationally.

To achieve more consistent paper quality, we must reduce the variability in the intermediate pulp. However, mills have certain productivity expectations and seek to maximize overall equipment efficiency. Many mills therefore operate at or above nominal capacity virtually all the time, leaving little or no "wiggle" room to compensate for outside disturbances.

At the same time, mills must deal with a glut of process data captured by millwide data historians. In theory, this is a gold mine of information; in practice, it is often difficult to show trends and relationships across the mill. Poor-quality measurements are partly to blame, but an equally important reason is the multivariate nature of the TMP mill itself.

Studying one or two variables at a time is a fruitless task, because the operating parameters, pulp measurements, and final paper quality are connected in such a way that the interactions between them are as important as the variables themselves. To be properly understood, therefore, the different variables in a TMP mill must be studied in combination. One way to do this is the statistical technique known as multivariate analysis, or MVA. Like all so-called "black-box" methods, which rely only on inputs and outputs, it has the pitfall that it can blindly attribute correlations between variables without any regard whatsoever for the actual underlying process.

Our goal was not to create a universal predictive model of the TMP process, which is virtually impossible without a series of designed experiments. Rather, the objective of our research is to develop a methodology for identifying the probable main sources of variation in important paper parameters using routinely collected data. This article presents part of the solution by proposing a way to correlate TMP operations with pulp quality, and ultimately paper quality, with reference to process fundamentals such as specific energy and refining intensity. The case study mill is an Eastern Canadian newsprint mill that experiences short- and long-term variations on the order of 10%-15% in final paper strength. Discrete events like primary and secondary refining line interruptions and changes in reject refiner operation are an integral part of this case study.

Rich vs. poor data sources

The impacts of pulp characteristics on final paper quality are well established. Newsprint strength, for instance, is known to depend on pulp fiber length, while other parameters critical to paper quality include freeness and fines content [1,2,3]. When using real mill data, however, it is often a challenge to show these links statistically because of the infrequency of measurements. Longer-term variations can be shown, such as weekly or monthly averages [4], but at shorter timescales, correlations are much harder to demonstrate.

This is true of the case study mill, whose characteristics were described in an earlier paper [5]. Pulp samples at the mill are automatically analyzed roughly every 60–120 minutes, as shown in **Table I**. The sampling process itself lasts only about 30 seconds, so these are grab samples, not composite samples. Along

Data Source	Frequency of Measurement	Number of Variables	Quality of Data
Wood chip quality	Grab sample every 8 h	A few physical characteristics	Very poor
TMP mainline operation	Once per second	Large number of energy, pressure, and flow rate measurements	Rich
Reject refiner operation	Once per second	Some key variables missing; must be calculated	Rich
Pulp quality – latency chest	Grab sample every 60–120 min	Limited to freeness, fiber length, and fines	Poor
Pulp quality – stock preparation	Grab sample every 60–120 min	Limited to freeness, fiber length, and fines	Poor
Final paper quality	Grab sample every 45 min	Many variables with much redundancy, especially for paper strength	Intermediate

I. Rich vs. poor data sources at case study mill.

with the inevitable measurement error and calibration issues, this approach tends to create much variability from one reading to the next. Moreover, pulp is a highly complex mix of organic compounds, both cellulosic and non-cellulosic, and biological in origin. Therefore, it cannot possibly be fully characterized by a handful of infrequent, macroscopic measurements.

One goal would be to automate the control of the TMP refiners to achieve the required pulp quality. Data from the papermaking section (longer loop) could be used to update the setpoints for pulp quality (shorter loop) as has been proposed in several commercial applications [6]. Before doing this, however, it is necessary to understand the impact of TMP operations on pulp and paper variability. The variables used in this study are shown in **Table II**.

In contrast to pulp, for paper the number of variables tested automatically is much higher, and there is a certain degree of redundancy among the numerous strength tests. However, measurements are still relatively infrequent. Tests are performed on a 30-cm-wide strip at the end of each reel (again, a small grab sample), corresponding to a period of roughly 45 minutes. For the purposes of this study, the following paper parameters were used: tear strength, burst strength, tensile stiffness index (TSI), and permeability to air (an indicator of porosity). These were selected based on preoccupations at the case study mill and known relationships with the parameters tested on the pulp.

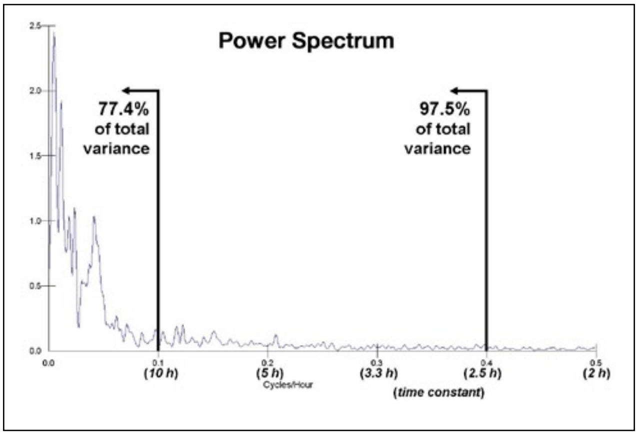
Sparser still are measurements on the incoming chips, tiny grab samples taken from the main conveyor belt only every eight hours. Furthermore, the tests performed on the chips are limited to density and moisture measurements, a rough size distribution, and rot content. Apart from very long-term trends, such as season-to-season, these chip data are far too limited to be useful for predicting final paper quality. Better and more frequent measurement of the chips would likely be extremely helpful, given the importance of wood species, growing conditions, harvesting techniques, and other such factors on final paper quality [7,8].

In stark contrast, the mainline refining and reject refining sections of the mill are extremely rich in frequent, plentiful data. The refiners and ancillary equipment are highly instrumented, with many operating parameters measured continuously. Fast variables, such as motor load, are logged in the data historian once per second, providing a very large amount of process information. However, these data present some important challenges:

• Certain critical operating parameters go unmeasured, e.g., refining consistency.

VARIABLE	UNIT
X-Variables	
Production rate (proportional to feed screw rotational speed)	t/d
Number of TMP lines in operation	—
Primary and secondary refining specific energy	kWh/t
Primary and secondary blowline consistency (calculated with simple mass/energy balance)	% solids
Primary and secondary refiner plate age	h
Standard deviation of motor loads (primary and secondary refiners)	MW
Reject refining specific energy	kWh/t
Reject refiner plate age	h
Y-Variables – Pulp Properties (PulpExpert)	
Canadian Standard Freeness	mL
Average fiber length (length-weighted)	mm
Fines content, defined as particles small enough to pass through a 200-mesh screen (76 μm)	% (mass)
Y-Variables – Paper Properties (Autoline)	
Tear, MD (index)	mN/(g/m^2)
Tear, CD (index)	mN/(g/m^2)
Bursting strength (index)	kPa/(g/m^2)
Tensile stiffness index (TSI), MD	kNm/g
Tensile stiffness index (TSI), CD	kNm/g
Permeability to air (indicative of porosity)	mL/min

II. Variables used to generate PLS model of TMP mill.



1. Power spectrum for tear strength, generated from hourly measurements over one month (August 2004).

• Variables are not in a form corresponding to known process fundamentals.
• Measurements are subject to instrument drift and calibration problems.
• Frequent starting and stopping of refiners.
• Gradual effects of equipment wear such as plate age.
• Presence of process lags.

Some of these issues, such as instrument drift and equipment wear, are present in any TMP mill. Others, such as lack of a refining consistency measurement, are specific to the case study mill.

Another particularity of the case study mill is overcapacity in the pulp section relative to the paper section. This results in frequent interruptions in each of the four main refining lines, which greatly affect the reject refining specific energy and other key parameters. A method for dealing with this problem is proposed below.

Despite these issues, it should be possible to exploit the richness of the refining data to model pulp and newsprint quality. First, though, we must determine whether the sampling frequency of the final paper properties is sufficient to be usable for MVA.

Frequency of paper sampling

As mentioned above, final paper quality is tested using small strips taken off the end of each reel. This approach is unavoidable because newsprint is sold in long continuous sheets which would be destroyed by intermittent cutting. This means, by definition, that any MVA models constructed using these data will at best be linked to long-term fluctuations in paper quality, as opposed to the shorter-term fluctuations that would occur within a single reel. The question remains whether these measurements properly represent long-term variations in paper quality or have been corrupted by the presence of short-term variations—the problem of aliasing. In other words, the question is whether the long-term paper quality variations (e.g., of period greater than two hours) are sufficiently large compared to the short-term variations (e.g., less than two hours) to allow one-hour sampling to track the majority of quality variations properly.

Following the technique recommended by Croteau et al. [9], we studied the power spectra of all the paper characteristics of interest, using one-hour averages. We found that three-quarters of the variability occurred at a frequency below 0.1 cycles per hour, i.e., at a time constant greater than 10 hours. Over 95% of the variability was below 0.4 cycles per hour, i.e., longer than 2.5 hours. This would suggest that the slower trends in final paper quality are indeed adequately represented and that this sampling frequency is sufficient for our purposes. **Figure 1** shows the power spectrum for tear strength only, but similar plots were found for all the paper parameters under study. There are two reasons for the long-term nature of these variations: the process, with its numerous tanks, screens, and recirculation loops, acts as a filter for pulp quality variations; and control loops at the paper machines tend to stabilize the short-term fluctuations in basis weight, moisture, and caliper that would otherwise affect paper properties.

It may seem paradoxical to use one-hour data to justify a one-hour time increment. Had we been doing designed experiments, rather than using existing data, no doubt we would have selected a shorter sampling frequency interval to be more conservative. However, such are the limitations when using real operating data.

Integrating TMP fundamentals into a PLS model

The main variant of MVA used to relate independent (X) variables to dependent (Y) variables is known as partial least squares, or PLS. As with any purely statistical method, without the use of a designed experiment, no causality can be inferred, only correlation.

MVA is a purely statistical tool and is automatically drawn to the most prominent trends in any dataset. Starts and stops of equipment, setpoint adjustments, routine calibrations, and other abrupt shifts will tend to dominate the latent variables, or "principal components," that are generated by the MVA algorithm. It is therefore necessary to pre-treat raw process data before using MVA, if only to remove periods of low production and other outliers and filter out noise. A detailed methodology for pre-treating TMP operating data was presented in an earlier TAPPI JOURNAL article [5].

Because MVA is a linear technique, it is also necessary to modify the variables to correspond better to the underlying process. TMP refiner operation is dominated by two nonlinear terms: specific energy, defined as mechanical energy per metric tonne of pulp produced, and refining intensity, or the specific energy per bar impact. Both must be properly controlled to ensure good pulp quality.

Specific energy is the central parameter of refiner operation. It must be high enough to ensure fiber separation and defibrillation, but low enough to avoid excessive fiber cutting, which can adversely impact the strength of the final paper [10]. Specific energy is not controlled directly, but rather indirectly, by means of the manipulated variables of throughput, dilution flow, and plate gap. It is relatively easy to calculate, since motor load is very accurately measured. Production rate is known from the volumetric feed rate, although the latter is susceptible to fluctuations in chip density.

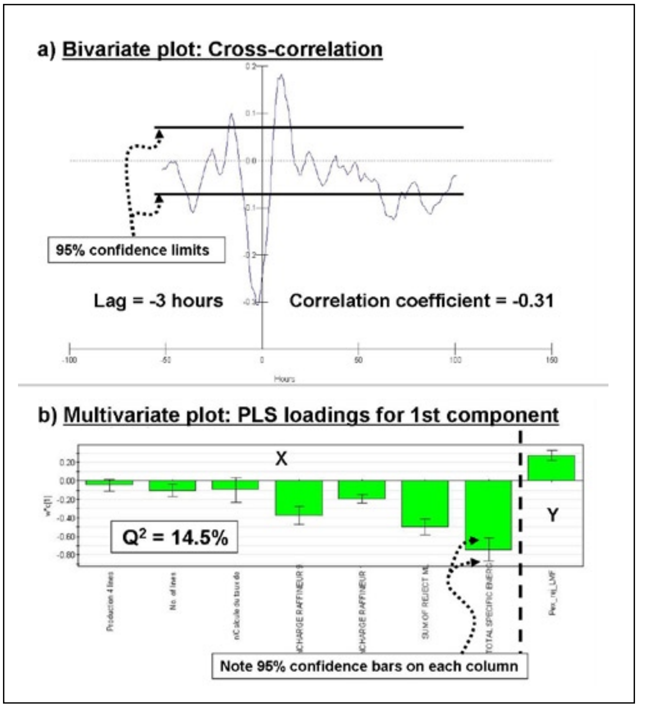
Refining intensity is much more difficult to measure in an industrial refiner. It indicates how the mechanical energy is applied, whether gradually over a large number of bar impacts, or suddenly by just one or two jarring impacts. The latter situation is highly damaging to fibers [11,12,13]. For a given plate configuration and disk rotation speed, refining intensity is largely a function of refining consistency [10]. This is an inverse relationship: higher consistency means a longer fiber residence time between the refiner plates, and hence more bar impacts of lower refining intensity for a given specific energy. At the case study mill, refining consistency is not measured directly in real time, but for Line 1 an online calculation based on a simple mass/energy balance is made available to operators. By comparing the equation in the distributed control system (DCS) logbook to real production data, it was possible to deduce the coefficients and extend this calculation to the other three lines.

The refiner plates have a normal lifespan of about 2000 hours, during which time the bars on the plate surface are gradually worn down, affecting the process characteristics. Less energy is required with an older plate to achieve a given freeness, but the refining intensity is higher because the plate gap is smaller, resulting in more damaged fibers [9]. Lama et al. [14] reported the effect of plate age on a mathematical model of TMP motor load for a twin refiner. Plate age displays both slow and very fast effects on the TMP refining process, because the plates wear down gradually over several thousand hours of operation and then are abruptly replaced with new plates. At the case study mill, the individual plate ages are continuously logged in the data historian and so are readily available.

Importance of reject refining

A very important, and sometimes overlooked, part of any TMP mill is reject refining. Because the case study mill experiences frequent starts and stops on the four main refining lines, the throughput at the presses, and ultimately in the reject refiners, is continually changing. Partly to compensate for this, the operators tend to increase the reject rate from roughly 30% to 40% after a line stoppage, meaning that the fiber length distribution of the pulp entering the reject refiners varies over time. These adjustments are made manually and not automatically, usually with a lag of several hours. Combined with occasional stoppages of the reject refiners themselves, this situation results in a highly variable and poorly controlled reject refining operation.

Unfortunately, the measurements at the reject refiners are insufficient to calculate the refining consistency, which could serve as an indicator of reject refining intensity. However, it is possible to calculate the specific energy at the reject refiners by adding the motor loads of the two reject refiners (they operate in parallel) and dividing by the throughput. The latter was estimated by simply multiplying the reject rate by the total production of the four main lines. The specific reject refining energy thus calculated showed a very large variability from hour to hour, ranging from 800 kWh/t to over 1400 kWh/t. By looking at the original variables, we concluded that this variability was due to fluctuations in both the numerator (motor load) and the denominator (throughput). Sometimes one of the two reject refiners was stopped, but the number of reject refiners in operation was not always a function of the number of working TMP lines.



2. Statistical relationship between reject refining specific energy (calculated from other variables) and pulp quality at the refined rejects outlet. 2a) Bivariate plot showing the cross-correlation between reject refining specific energy and fiber length (August 2003). 2b) Corresponding multivariate "loadings" plot showing all the reject refining variables.

Though based on several assumptions, this estimate of specific reject refining energy showed significant correlation with pulp quality at the refined rejects outlet, much more so than the original variables did. This was true for both bivariate (cross-correlation) and multivariate (MVA) statistical tests, as shown in **Fig. 2** for fiber length. Note that the calculated specific reject refining energy dominates the MVA plot in Fig. 2b (second column from the right) despite the presence of the original variables from which it had been calculated.

In each case, the amount of variability explained is not particularly high: a correlation coefficient of 31% in one case and a Q^2 of 14.5% in the other (Q^2 represents the proportion of the Y-variance captured by the PLS model; it is similar to R^2, but decreases with overfitting). This is not surprising, since reject refining is only part of the overall picture. These correlations, while only explaining part of the picture, are nonetheless statistically significant, as indicated by the 95% confidence limits on both plots.

Process lags

MVA is not a time-series technique and treats all observations exactly the same, whether they be sequential or three weeks apart. Process lags must be synchronized beforehand; otherwise, the algorithm may blindly combine data for different moments in time into a series of muddled, meaningless data points. At the case study mill, the process lags for the mainline pulp are not constant because of changes in tank levels. For the purpos-

LOCATION	LAG
Based on cross-correlation curves:	
Headbox feeder tank	0 h
Disk filter feed	-2 h
Refined rejects outlet	-3 h
Primary screen accepts	-5 h
Outlet of latency chests	-5 h
Assumptions, by simple extrapolation:	
Final paper	0 h
Reject refiners	-5 h
Primary and secondary refiners	-6 h

III. Average process lags used in PLS model, based on cross-correlation curves for average fiber length at various locations (August 2003).

es of this study, the monthly average lags between unit operations were used. These can be estimated by comparing cross-correlation curves [15], in this case for fiber length, measured at different points in the process, as listed in **Table III**.

We tried to link process lags to levels in the high-density tank, which tend to fluctuate significantly. However, we found no statistically significant relationships, probably because of the small number of data points available for any given tank level.

Linking paper quality back to TMP operations

To model final paper quality, it was necessary to combine all the different sections of the TMP mill into a single model. Essentially, we did to the data what occurs to the pulp within the real process, i.e., combine all the main lines as well as the reject lines, add process lags, and dampen high-frequency fluctuations.

Initially, we created models in which the four TMP lines were represented separately. It quickly became apparent that this approach led to serious problems. The models were exceptionally hard to interpret because they tended to reflect the idiosyncrasies of the four individual lines, rather than the entire effect of upstream operations on the pulp and newsprint. Thus, one week Line 1 might dominate, whereas for a different week, it might be Line 3. Not surprisingly, these models tended to be weaker overall.

Before the four main TMP lines were combined, each was first treated individually by removing all shutdown periods. Any hour during which the minimum production rate (i.e., smallest second-by-second value over the entire hour) was below 200 t/d was systematically erased. The hour itself was not removed, but the values for production, specific energy, and consistency were replaced with blanks. Combining the four main lines, therefore, consisted in taking the average of just those lines that were in operation for any given hour. This technique completely eliminated the effect of starts and stops, while giving the correct weighting to the lines in operation.

In the case of plate age, this approach was unnecessary, because the plates do not age when the line is shut down. The ages for all four lines were simply added together. To differentiate between the different kinds of refiners, the four primary refiners were considered as one group, the four secondary refiners as another group, and the two reject refiners as a third group. These total plate ages were then used as new X's in the MVA model to give the final choice of X's listed in Table II.

PLS Model	Number of Components	Q^2
Pulp Quality (freeness, fiber length, and fines content)	5	38%
Paper Strength (tear, burst, and TSI)	3	47%
Porosity (permeability to air)	3	69%

IV. Summary of PLS models obtained using hourly averages (August 2003).

Because of frequent grade changes, the original plots of tear and bursting strength showed abrupt shifts not related to the upstream TMP process. These would have destroyed any chance of linking these two parameters to the refining operation. Dividing these two parameters by the basis weight to create an index greatly reduced this effect, which underlines the importance of studying the original data before doing any MVA. Permeability was also clearly linked to basis weight, but in this case indexing was not helpful, so the models were limited to periods of 45-g/m^2 production (roughly two-thirds of the hours in the study period) for this parameter only.

All variables, both X and Y, were further treated to remove outliers and noise. Using another variant of MVA known as principal component analysis (PCA), we identified data points that did not fit the overall correlation structure. Often these outliers yield interesting clues about the process, such as shifts in operating regime. In this case, however, our goal was to represent the most typical TMP operation and relate it to slower trends in the paper, so the outliers were systematically removed. To counteract noise, all variables were also subjected to filtering. We used an exponentially weighted moving average (EWMA) with an alpha value of 0.8, corresponding to a first-order filter with a time constant of four hours—a rough approximation to the filtering provided by the process.

Using this approach, it was possible to link pulp quality back to TMP operations, including reject refining, as shown in **Table IV**. The Q^2 value listed is a combined metric for that group of dependent variables; the results obtained would have been different if each dependent variable had been modeled separately. **Table V** shows the TMP and reject refining operating variables associated with each of the dependent variables, in other words, which X-variables were grouped by the PLS model with the Y-variables. For the most part, the trends that we found correspond to what would be expected for a typical TMP operation.

With this multivariate approach, we were also able to find correlations between the upstream operations and the final newsprint properties. **Table VI** shows which TMP and reject refining operating variables were statistically linked to each of the paper quality parameters. Again, the results are consistent with expectations. One interesting result not reflected in

Sampling location	Dependent (Y) variable	Independent (X) variables showing strongest correlation within PLS model
Outlet of refined rejects tank	Higher freeness	Lower specific energy in reject refiners.
	Higher average fiber length	Lower overall specific energy; lower refining consistency; older plates.
	Higher fines content	Higher overall specific energy; higher refining consistency; newer plates; fewer TMP lines in operation; higher motor load standard deviation.
Primary screen accepts	Higher freeness	Lower overall specific energy; lower refining consistency; older plates.
	Higher average fiber length	Lower overall specific energy; lower refining consistency; older plates.
	Higher fines content	Higher overall specific energy; higher refining consistency; newer plates; fewer TMP lines in operation; higher motor load standard deviation.

V. Pulp Quality Model – TMP and reject refiner operating variables (X-variables) most associated with each dependent (Y) variable (August 2003).

Dependent (Y) variable	Independent (X) variables showing strongest correlation within PLS model
Higher tear strength	Higher overall specific energy; higher refining consistency; newer plates; more TMP lines in operation.
Higher bursting strength	Higher overall specific energy; higher refining consistency; newer plates; more TMP lines in operation.
Higher TSI	Higher overall specific energy; higher refining consistency; newer plates; fewer TMP lines in operation.
Higher porosity (permeability to air)	Lower overall specific energy, especially in reject refining; lower refining consistency; older plates; more TMP lines in operation.

VI. Paper Quality Model – TMP and reject refiner operating variables (X-variables) most associated with each dependent (Y) variable (August 2003).

the table is the presence of links between secondary plate age and strength (positive) and permeability (negative), exactly the opposite of those found for the other two plate ages. This could be attributable to coincidence or to some outside factor affecting the process. Again, it is important to highlight that MVA finds only correlation, not causality.

It is important to note that the Q^2 values presented above represent the proportion of the Y-variance captured by the model, and not the statistical significance of the model *per se*. The latter is indicated by the 95% confidence limits on the PLS histograms. When using real process data, it is common to have Q^2 values in the range of 40%, i.e., the model explains with statistical significance 40% of the Y-variance. The unexplained portion of the variance corresponds to unmeasured (or unmeasurable) process variables that impact the quality parameters or to measurement errors and noise, both of which are prominent in real operating data.

DISCUSSION

The use of historical TMP operating data, while a promising source of insight, can be difficult because of outliers, noise, and the multivariate nature of the process itself. MVA is a cheap, non-intrusive method for analyzing such data, although interpreting the results and recognizing and discarding findings that are merely coincidental can be quite difficult.

Using MVA and other statistical tools, it was possible to link one-hour values for pulp freeness, fiber length, and fines content back to the TMP refining operations, including reject refining. To get the best results, it was necessary to convert the existing variables into a form closer to the process fundamentals by creating nonlinear terms to represent specific energy and refining intensity. It was also necessary to take into account the number of lines in operation at any given time, the gradual wearing of the refining plates, and the presence of process lags. Of all these factors, plate age seemed to have the greatest effect on pulp parameters, even though actual plate wear is known to be nonlinear with time. Specific energy in the reject refiners also had a significant effect, notably on fiber length.

Refining consistency, calculated from other variables, was found to trend almost always with specific energy. As a result, the anticipated independence of specific energy and refining intensity was not observed. To verify this point, it would be necessary to conduct designed experiments in which specific energy would be held constant and the refining consistency would be manipulated in a controlled manner.

Using the same approach, it was also possible to explain roughly half the variability in final paper quality, represented in this case by strength, porosity, and linting propensity. This result is quite remarkable, considering that no paper machine operating variables were used and that little or no data were available on incoming chip quality. Of course, these results are correlations, not cause-and-effect assertions; it is possible that both the TMP operations and the paper properties are being affected by a third, unmeasured factor. This question will be the theme of our next article, along with a systematic comparison of different timescales and sampling locations.

Such an approach could be used for a two-tier control strategy in which the results of paper monitoring would serve to update setpoints for pulp quality in the TMP mill. Because the TMP process is a moving target, any such strategy would require an adaptive controller.

RECOMMENDATIONS

Using multivariate analysis and other statistical tools, it was possible to link pulp quality back to TMP and reject refining operations, taking into account the number of lines in operation, plate age, and process lags. Furthermore, it was possible to explain roughly half the variability in final paper quality by creating MVA models of the refining operations based on process fundamentals such as specific energy and refining intensity. Based on our results, it seems that there is room for improvement in the control of specific energy and the scheduling of plate changes. Another conclusion that can be drawn from this work is that even better results would likely be obtained with more online measurements, in particular for incoming chip quality and reject refining consistency. **TJ**

ACKNOWLEDGEMENTS

This work was completed with support from the Natural Sciences and Engineering Research Council of Canada (NSERC) Environmental Design Engineering Chair at École Polytechnique. We would also like to acknowledge Martin Fairbanks of Abitibi-Consolidated, Inc., for his invaluable advice and inspiration.

Received: September 23, 2006
Accepted: January 6, 2007

LITERATURE CITED

1. Saltin, J.F. and Strand, B.C., *Pulp and Paper Canada* 96(7): 48 (1995).
2. McDonald, D., Miles, K., and Amiri, R., *Pulp and Paper Canada* 105(8): 27 (2004).
3. Law, K., *Pulp and Paper Canada* 106(1), T5 (2005).
4. Browne, T.K., Miles, D., McDonald, J., et al., *Pulp and Paper Canada* 105(10): 35 (2004).
5. Harrison, R.P. and Stuart, P.R., *TAPPI J.* 5(8): 17 (2006).
6. Strand, W.C., Fralic, G., Moreira, A. et al., *Proceedings, 2001 International Mechanical Pulping Conference*, Helsinki, Finland, Vol. 2, p. 253.
7. Wood, J.R., *TAPPI J.* 84(6): 32 (2001).
8. Ding, F., Benaouida, M., Bédard, P., et al., *Pulp and Paper Canada* 106(2): T25 (2005).
9. Croteau, A.P., Nobleza, G.C. and Roche, A.A., *Pulp and Paper Canada* 94(1): T25 (1993).
10. Roche, A., Owen, J., Miles, K., et al., *Proceedings, Control Systems 96*, Halifax, Canada, p. 129.
11. Miles, K.B., "The Essence of High Consistency Refining." *The Marcus Wallenberg Foundation, Symposia Proceedings 12, Mechanical Pulping Scientific Achievements* (1998).
12. Miles, K.B. and Omholt, I., *Proceedings, 2003 International Mechanical Pulping Conference*, Quebec City, Canada, p. 179.
13. May, W.D., "*The Miles and May Model – a Presentation.*" *The Marcus Wallenberg Foundation, Symposia Proceedings 12, Mechanical Pulping Scientific Achievements* (1998).
14. Lama, I., Perrier, M., and Stuart, P.R., *Nordic Pulp & Paper Research Journal*, 21(4), p. 527-533 (2006).
15. Nobleza, G.C., Roche, A.A., and Croteau, A.P., *Pulp and Paper Canada* 91(7): T280 (1990).

INSIGHTS FROM THE AUTHORS

Even though modern-day mills have access to a lot of process data, there are serious challenges to extracting useful knowledge from this data. One critical problem is the sheer complexity of the dataset, with a very large number of variables and time periods from which to choose.

This paper follows on earlier work in which various data pre-treatment techniques were systematically applied to see which ones yielded the best models of TMP pulp quality. Here we have delved further into the process fundamentals in an effort to improve these models, focusing on refining energy and consistency, most notably in the reject refining section.

Within the mill, there are rich and poor data sources, which somehow must be combined into a coherent model. We addressed this by using as short a timescale as possible, then applying filtering to bridge the gap between the faster (process) measurements and slower (product quality) measurements.

The most surprising discovery from this work is how much "black box" models can be improved by incorporating knowledge about process fundamentals.

The next step in this work is to consider applying our experience in "black box" TMP modeling to process monitoring and process control applications.

Stuart

Roche

Harrison

Paul Stuart is a professor of chemical engineering at Ecole Polytechnique, the engineering school of the University of Montreal. He also holds the NSERC design engineering chair in process integration in the pulp and paper industry. Author contact: paul.stuart@polymtl.ca. Alain Roche is a research engineer at the Pulp and Paper Research Institute of Canada (Paprican) in Pointe Claire, Quebec, Canada. Robert Harrison recently obtained his doctoral degree in chemical engineering from Ecole Polytechnique.