Improving the reliability of newsprint quality data using integrated factor networks for paper machine control and analysis

Johan F. Saltin and William C. Strand

ABSTRACT: Factor analysis can be used to improve the reliability of newsprint quality by quantifying the interrelationships among strength and printing properties. Factor models are used not only for data reconciliation but also for interpreting the underlying factor networks. Using these techniques, changes in paper quality can be traced to the underlying causes. Applied to a newsprint paper machine, these techniques include the use of integrated factor networks for data reconciliation, disturbance detection, and identification.

KEYWORDS: Analysis of variance, data, handsheets, measurement, mechanical properties, networks, newsprint, paper machines, physical properties, printing, process control, quality, reliability, test methods.

One of the primary requirements of a robust process control system is that the signals used for control must be accurate. The need for accuracy is especially critical in controlling quality on a paper machine. As with all laboratory tests, measurements of paper properties contain measurement error. The error can be reduced through standardization and duplication of tests. However, the errors can never be totally eliminated.

Various tests in the pulp and paper industry are used to characterize the strength and printing properties of newsprint, as listed in **Table I**. However, all these measurements are not used for control. Often, one paper property is used for control, and all additional variables are simply monitored. If the sensor measuring the key paper property fails, the situation that occurs is that an erroneous measurement is being used for control, even though other noncontrol variables may be correct.

Factor analysis

A large portion of the variation in the quality tests listed in Table I can be attributed to a relatively few "common factors" that link seemingly unrelated properties. These common factors are the result of various intrinsic fiber properties such as fiber bonding, fiber length, and fiber orientation.

Factor analysis is a data reduction technique that can be used to analyze interrelated paper strength and optical properties such as those shown in Table I. Factor analysis has two primary goals:

1. To determine how many significant independent paper properties actually exist. In factor analysis, this determination is made by examining the degree of correlation among the paper quality variables.
2. To take common paper properties and combine them to form "composite" variables. These composite variables are a mathematical representation of the common factors that determine pulp quality. For the analysis of paper properties, the form of the equations used to calculate these common factors is as shown by Eq. 1.

$$F_1 = a_{1,1}\,tear + a_{1,2}\,tensile + \ldots + a_{1,n}\,density \qquad (1A)$$

$$F_2 = a_{2,1}\,tear + a_{2,2}\,tensile + \ldots + a_{2,n}\,density \qquad (1B)$$

$$\vdots$$

$$F_k = a_{k,1}\,tear + a_{k,2}\,tensile + \ldots + a_{k,n}\,density \qquad (1K)$$

Saltin is a research assistant at Dept. of Chemical Engineering, University of Idaho, Moscow, ID 83843. Strand is project manager at Pacific Simulation, 121 Sweet Ave., Moscow, ID 83843.

I. Strength and printing properties of newsprint

Strength properties

- Tear strength, MD and CD
- Tensile strength, MD and CD
- Stretch, MD and CD
- Tensile energy absorption, MD and CD
- Basis weight
- Caliper
- Density
- Formation

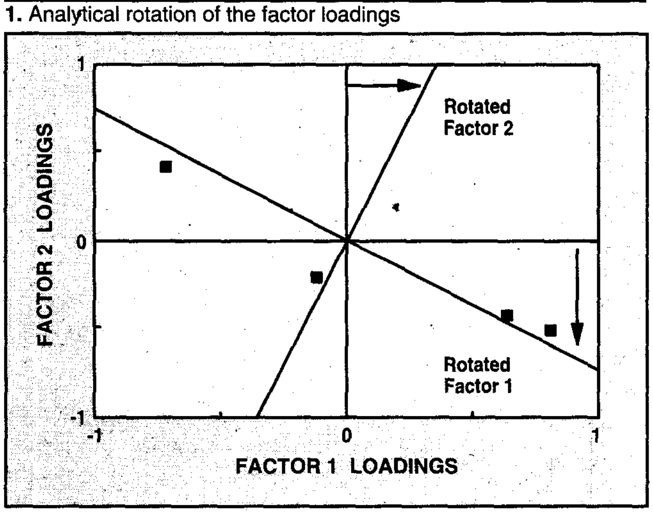
Printing properties

- Roughness, top side and bottom side
- Water drop
- Oil drop
- Scattering coefficient
- Air permeability
- Linting propensity

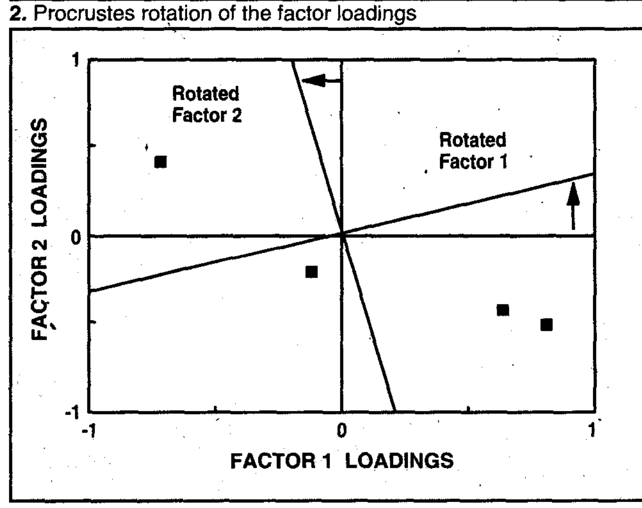
II. Effect of an analytical rotation on the factor loadings

	Variable 1	2	3	4
Factor 1				
Extraction	0.81	–0.71	0.63	–0.11
Rotation	0.75	–0.68	0.58	0.02
Factor 2				
Extraction	–0.52	0.41	–0.44	–0.22
Rotation	0.02	–0.05	0.00	–0.19

1. Analytical rotation of the factor loadings



2. Procrustes rotation of the factor loadings



where

F_i = the ith common factor

$a_{i,j}$ = model coefficients.

In factor analysis, the number of significant common factors is first determined. The number of common factors is always less than the number of paper properties analyzed. If there is a high degree of correlation among the paper properties, the number of common factors required will be small.

The next step is to determine the degree of correlation between each paper property and each common factor. These correlation coefficients are called the "factor loadings." The constants $a_{i,j}$ in Eq. 1 are then determined from the factor loadings.

There are many methods for determining the factor loadings, but the method chosen is one that gives the most meaningful interpretation from the standpoint of intrinsic fiber properties. (This will be discussed in detail in the following section.)

The common factor model can then be used to express each pulp variable in terms of these common factors (F_1) and a unique factor (e_1):

$$\text{Tensile} = b_{1,0} + b_{1,1}F_1 + b_{1,2}F_2 + \ldots + b_{1,k}F_k + e_1 \tag{2A}$$

$$\text{Tear} = b_{2,0} + b_{2,1}F_1 + b_{2,2}F_2 + \ldots + b_{2,k}F_k + e_2 \tag{2B}$$

$$\vdots$$

$$\text{Density} = b_{n,0} + b_{n,1}F_1 + b_{n,2}F_2 + \ldots + b_{n,k}F_k + e_n \tag{2N}$$

where

$b_{i,j}$ = constants

$e_{i,j}$ = unique factor (or error term).

III. Effect of a Procrustes rotation on the factor loadings

	Variable			
	1	2	3	4
Factor 1				
Extraction	0.81	−0.71	0.63	−0.11
Target	0.70	−0.50	0.50	−0.10
Rotation	0.73	−0.51	0.50	−0.12
Factor 2				
Extraction	−0.52	0.41	−0.44	−0.22
Target	−0.70	0.50	−0.50	−0.10
Rotation	−0.72	0.53	−0.51	−0.13

These two sets of equations, Eqs. 1 and 2, describe the interrelationship among the paper properties by a small number of completely independent common factors.

Rotation of the factor loadings

The common factors derived from the factor analysis can often be identified as intrinsic properties of the system. For example, factor analysis of the strength and optical properties of mechanical pulp handsheets yields two common factors that can be identified as fiber bonding and fiber length (*1–3*). However, when the system becomes more complex, the common factors derived from the factor analysis can be difficult to identify as intrinsic properties. One method that can be used to aid in the identification of common factors is rotation of the factor loadings.

The first step in factor analysis is to calculate the correlation matrix among all the variables. From this correlation matrix, the initial factor loadings are found using factor extraction. The values of the initial factor loadings (correlation between the variables and the common factors) are dependent on which method has been used for factor extraction. For the same correlation matrix, there exists an infinite number of different factor loadings, as a result of the indeterminate nature of the factor analysis model (Eq. 1, A–K). To change the values of the factor loadings, the loadings can be rotated. The rotation of factor loadings does not change the overall result of the factor analysis—it only affects the interpretation of the common factors.

Figure 1 is an example of an analytical rotation for two common factors, where the deviation around the factor loadings is to be minimized. In the original factor extraction, the factor loadings were distributed around each axis. The rotation moves the axis in such a way that three of the variables lie on one factor axis and the fourth variable lies on the second axis. In **Table II**, the original and the rotated factor loadings are given.

Figure 2 shows a second rotational method called a "Procrustes rotation." Procrustes rotation uses a hypothesized "target" for the factor loadings. The rotation is performed toward this target matrix. The target factor loadings are given in **Table III** along with the original and the rotated factor loadings. The advantage of this type of rotation is that the common factors can be rotated to an estimate of an intrinsic property. If the estimated values exist in the system, the factor loadings will be rotated to that estimate. However, if the estimate is not correct or the intrinsic property is not part of the system, the rotation will not be completed.

A Procrustes rotation is easily performed for a system with two common factors, such as in the example in Fig. 2. However, as the number of common factors increase, the rotation becomes much more complex since it is necessary to define a factor loading for each combination of variable and common factor. For a system with five common factors and twenty variables, it would be necessary to define 100 target values.

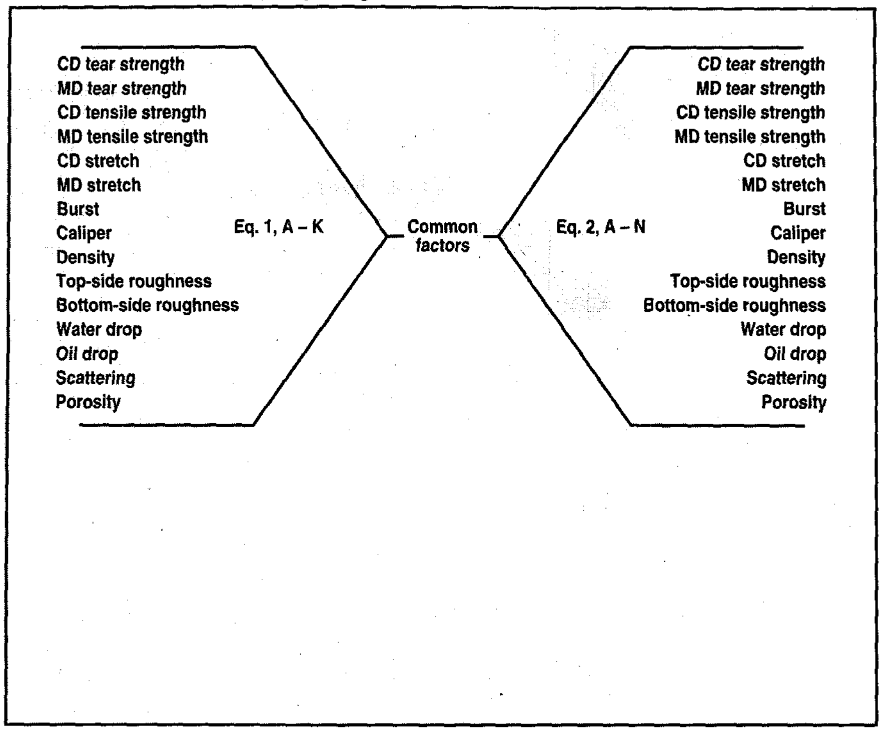
Therefore, a modified method for rotating the factor loadings has been developed for use in this analysis. Instead of supplying estimates for all target values, estimates of only a portion of the values are given. The remaining values are found through an iterative procedure.

The advantage of using methods such as this modified Procrustes rotation lies in the interpretation of the common factors. When each common factor has a physical explanation, it is possible to use the calculated values of the common factors as a tool for troubleshooting. Rather than a large number of variables being examined, a much smaller number of common factors can be used to monitor the state of a given process.

Data reconciliation using factor models

As discussed, two sets of Eqs. 1 and 2 describe the interrelationship among the paper properties by a small number of completely independent common factors. When the relationship is portrayed graphically, as in **Fig. 3**, these equations take the form of a "factor network." The factor network technique has the unique property of making it possible not only to calculate the common factors, but also to recalculate the paper quality data in the

3. Reconciliation of paper quality using the factor network



IV. Strength and printing properties of newsprint

Strength properties

Tear strength, MD and CD
Tensile strength, MD and CD
Stretch, MD and CD
Tensile energy absorption, MD and CD
Burst strength
Caliper
Density

Printing properties

Roughness, top side and bottom side
Water drop
Oil drop
Scattering coefficient
Porosity

form of "reconciled" data. The use of this property of the factor network for pulp quality reconciliation to remove random testing error in mechanical pulp handsheet tests has been discussed in an earlier publication (*2*).

The result of using the factor models in this fashion can be described as "single sample filtering." Filtering usually denotes some type of time averaging, such as a moving average or exponential filtering. However, in the case of single sample filtering, variables are filtered using other variables in the same observation—for example, filtering a tear index test using measurements of tensile strength, density, or light-scattering coefficient taken on the same sample. The result is that random test variations can be filtered without introducing time delays, which are inherent in most conventional filtering methods.

V. Target factor loadings for newsprint samples

Handsheet property	*Fiber bonding*	*Fiber length*	*Unbond surface*	*Fiber orientation*	*Pressing*
Caliper	–0.80	0.20	...	...	–0.40
Density	0.80	–0.20	...	...	0.40
Tear strength					
CD	...	0.80	...	–0.40	–0.40
MD	...	0.80	...	0.40	–0.40
Tensile strength					
CD	0.60	...	...	0.40	–0.40
MD	0.60	...	...	–0.40	–0.40
Stretch					
CD	...	...	...	0.40	–0.50
MD	...	...	...	–0.40	–0.50
Burst strength	0.60	0.20	...	...	–0.40
Roughness					
Top side	–0.60	0.20	–0.40	...	–0.40
Bottom side	–0.60	0.20	–0.40	...	–0.40
Porosity	–0.80	0.40	...	...	–0.40
Scattering	–0.40	–0.40	0.60	...	–0.20
Water drop	0.60	...	...	...	...
Oil drop	0.60	...	...	...	...

Ellipsis indicates that no target was supplied for the factor load.

Interactive analysis interpretation

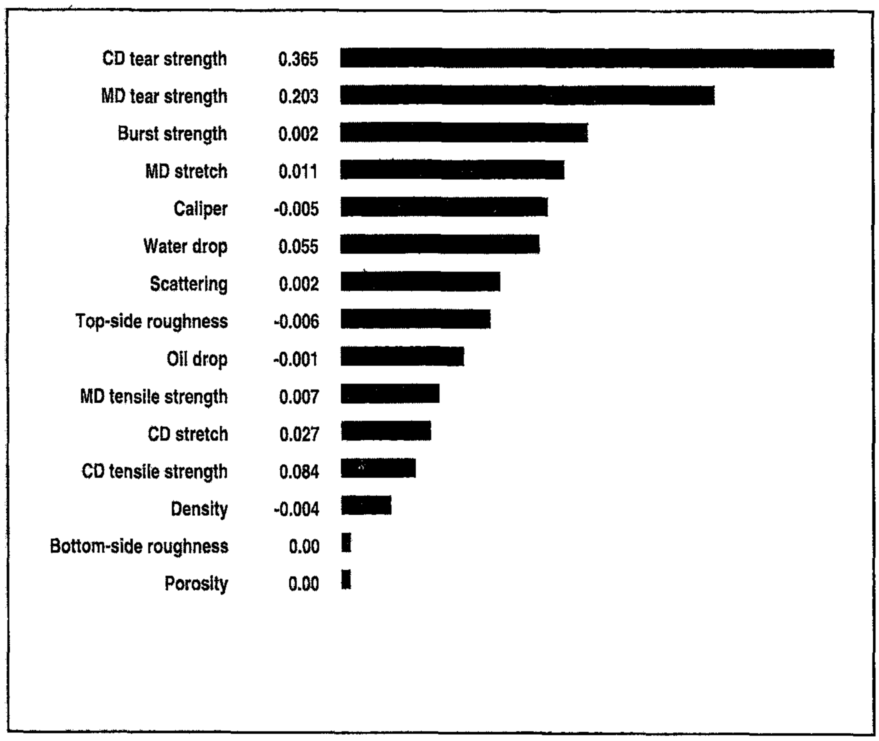
One of the difficulties in performing this type of analysis on a large number of variables is the interpretation of the results. An analysis with 20 variables and five common factors results in 100 factor loadings that must be reviewed and interpreted.

To aid in the interpretation and rotation of large factor models, the FactNet analysis system was developed (*4*). This analysis system is an interactive software package that uses a number of different types of graphi-

VI. Factor loadings for newsprint samples

Handsheet property	*Fiber bonding*	*Fiber length*	*Unbonded surface*	*Fiber orien-tation*	*Pressing*
Caliper	–0.73	0.25	–0.16	–0.14	–0.43
Density	0.76	–0.22	0.18	0.07	0.44
Tear strength					
CD	0.08	0.76	0.10	–0.27	–0.18
MD	–0.20	0.66	0.18	0.40	–0.26
Tensile strength					
CD	0.62	0.06	–0.22	0.24	–0.16
MD	0.67	–0.21	–0.30	–0.41	–0.21
Stretch					
CD	0.22	–0.08	–0.01	0.35	–0.61
MD	0.52	–0.03	–0.02	–0.17	–0.50
Burst strength	0.66	0.11	–0.33	–0.38	–0.22
Roughness					
Top side	–0.49	0.13	–0.61	0.19	–0.42
Bottom side	–0.56	0.16	–0.60	0.21	–0.37
Porosity	–0.66	0.36	–0.08	0.40	–0.06
Scattering	–0.46	–0.23	0.73	–0.04	–0.33
Water drop	0.49	–0.53	–0.11	0.03	–0.15
Oil drop	0.71	–0.30	–0.04	–0.05	0.16

4. Graphical representation of the CD tear strength filter



cal representations to aid the analyst in interpreting large factor models. Rather than attempting to review and interpret a large number of factor loadings, the analyst is presented with a network representation of the factor model with the loadings presented in color-coded form. Insignificant loadings can be removed graphically to present only the strongest and most significant effects. Rotation targets are interactively input and analyzed, allowing a large number of rotations to be analyzed and interpreted quickly. The result is that an analysis that would take over a week to perform, review, and interpret using conventional methods can be performed in several hours using this analysis system.

Analysis of newsprint data

Four databases were analyzed for this study:

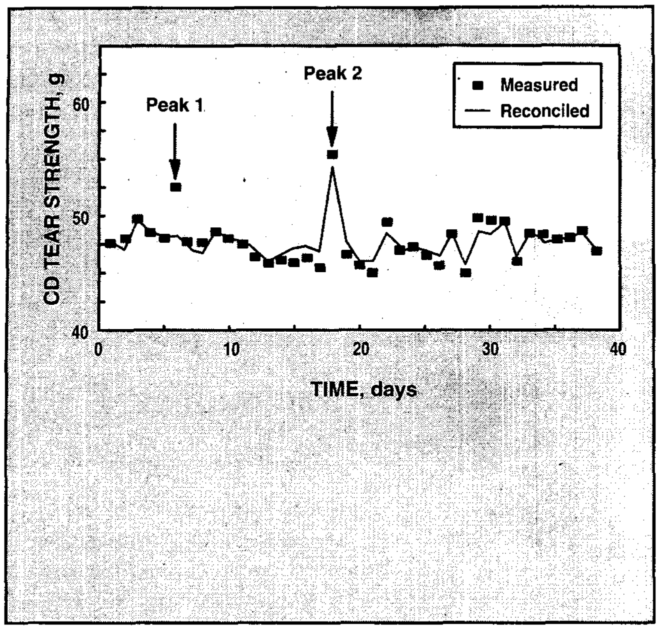
- Database 1: Daily averages of reel quality covering a one-year period for a paper machine producing newsprint quality from Scandinavian spruce thermomechanical pulp (TMP), stone groundwood pulp (SGW), and kraft pulp.
- Database 2: Daily averages of reel quality covering a nine-month period for a paper machine producing newsprint quality from a blend of kraft pulp and spruce, fir, and pine TMP.
- Database 3: Competitive testing of newsprint sheets from eleven installations in North America.
- Database 4: Competitive testing of newsprint sheets from eight installations in North America.

The analysis results for all four databases were similar. The analysis results presented here consisted of 250 samples including the test measurements shown in **Table IV**.

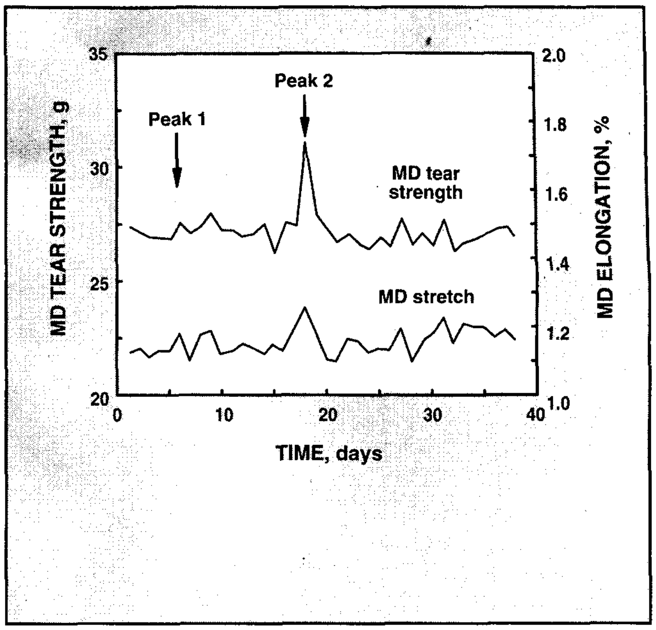
Interpretation of the mill data analysis

Factor analysis indicated the existence of five significant common factors that explained 80% of the variation in the 16 properties shown in Table IV. Unfortunately, the results cannot be presented in the color-coded, graphical form used by the analysis system. Therefore, the results are presented here in a more

5. Measured and reconciled CD tear strength



6. Measured MD tear strength and MD breaking length



conventional, tabular form.

The first step in using the analysis system is to make a preliminary identification of the five common factors. Experience with mechanical pulp handsheets indicates that two of the common factors are fiber bonding and fiber length (*1*). The methodology used in producing handsheets ensures random fiber orientation. However, paper machine operation results in varying degrees of fiber orientation that affect the interrelationships among quality variables. Handsheets are also produced using standardized pressing—a variable that may be varied on the paper machine. In addition, since newsprint furnishes may contain fillers as well as several different pulp types, unbonded surface area is a candidate as a common factor. Thus, we have five possible common factors—fiber bonding, fiber length, unbonded surface area, fiber orientation, and pressing (either wet pressing or calendering).

The second step in using the analysis system is to use process experience to determine approximate target factor loadings between the paper quality variables and the five preliminary common factors. The following general estimations can be made:

- Fiber bonding: high tensile strength, good surface properties
- Fiber length: high tear strength, poor surface properties
- Unbonded surface: good surface properties and scattering
- Fiber orientation: properties vary in the machine direction (MD) and cross-machine direction (CD)
- Pressing: poor strength, good surface properties, low caliper

These general estimations are then input in more detail in the form of target factor loadings. The target factor loadings that were used by the analysis system are shown in **Table V**.

The five common factors are then rotated using the sparse matrix Procrustes rotation discussed earlier in an attempt to match the target loadings. This method does not change the overall results in terms of analysis accuracy but simply changes the calculation of the intermediate common factors. The total amount of variation in the pulp quality variables explained by the five common factors remains constant, regardless of any rotational method. However, the amount of variation explained by any single factor can change dramatically, depending on the rotation. The results of the rotation are shown in **Table VI**. The high degree of correspondence between the target and actual factor loadings indicate that the preliminary identification of the five common factors was correct.

Paper quality reconciliation

Once the factor models have been created, the models can be used to "filter" the raw paper quality data using the methods described. The technique of paper quality reconciliation is simple in its approach. Because of the highly interrelated nature of paper quality variables, changes in any single variable should be mirrored by changes in related quality variables. Variations that are not mirrored are assumed to be random test variations or miscalibrations and are filtered using the factor models.

To allow the filter to be viewed in more detail, the analysis system provides a graphical representation of the filter, as shown in **Fig. 4** for CD tear strength. The length of the bars in the histogram represent the importance of this variable in the filter calculations, while the number represents the actual coefficient used in the filter calculations.

For example, the most important property in filtering the CD tear strength measurement is the CD tear strength measurement itself. The model constant (0.36) indicates that a change of one unit in the CD tear strength measurement results in a calculated change of 0.36 units in the reconciled tear strength. The reason the filter only accepts 36% of the change as valid is that other measurements such as MD tear strength and MD sstretch should change at the same time as the CD tear strength. If these changes occur simultaneously, the filter calculations result in a reconciled CD tear strength value that more closely matches the raw value.

An example of this filtering for CD tear strength for one of the test mills is shown in **Fig. 5**. Figure 5 contains both the raw and reconciled tear strength measurements over a 40-day period. Note the presence of two peaks in the CD tear strength measurement.

The filtering results in a significantly lower reconciled value for the first peak (52.7 compared to 48.5), while the raw and reconciled values are quite similar for the second peak (55.5 compared to 54.5). The reason for this can be seen when the plots for MD tear strength and MD stretch are examined (**Fig. 6**). Note that there is no change in these variables corresponding with the first peak in CD tear strength. Since the factor models indicate that CD tear strength has a strong correlation with these variables, the factor model filters out a portion of the peak. In contrast to the first peak, MD tear strength and MD breaking length both show changes that correspond with the second peak in CD tear strength. The filter thus calculates a reconciled CD tear strength, which more closely mirrors the raw measurement value.

As discussed earlier, this type of "single sample filtering" does not produce time delays which are normally associated with conventional filtering. The use of reconciled paper quality for paper machine control has several obvious advantages. A measurement with reduced random variations can be used in a tighter control loop, allowing rapid response to process and quality disturbances.

Factor networks as input–output models

Currently, factor models are being used in several facets of paper machine control. Work has now been completed in the development of factor models that integrate both dependent paper quality variables and independent process variables. These integrated factor networks can be used for data reconciliation in the same fashion as described earlier.

Additionally, these integrated factor models can be used to calculate paper quality from whatever combination of dependent and independent variables is available at the time (*5*). This type of integrated factor model combines the best attributes of factor modeling and analysis methods, such as partial least squares, that distinguish between dependent and independent variables. **Tj**

Literature cited

1. Strand, B. C., 1987 Pulping Conference, TAPPI PRESS, Atlanta, p. 61.
2. Strand, B. C., Mokvist, A., and Jackson, M., 1989 International Mechanical Pulping Conference, EUCEPA, Helsinki, p.362.
3. Strand, B. C., Mokvist, A., Ferritsius, O., Jämte, J., and Sköld, H., On-line prediction of mechanical strength and optical properties, 1992 CPPA Annual Conference, CPPA, Montreal, p. B53.
4. FactNet Users Manual, Version 1.0, Pacific Simulation 1992.
5. Saltin, J. F. and Strand, B. C., presented at CPPA Control '92, CPPA, Whistler, BC, 1992, p. 53.

Received for review June 29, 1992.

Accepted Aug. 5, 1992.

Presented at the TAPPI 1992 Engineering Conference.