Metrology—impact of sampling theory

Manual sampling methods provide information of questionable validity, thereby increasing process spread and causing targets to be set higher than optimum. Continuous measurement and control gives dual benefits of improving the process and eliminating the need for a safety factor in targets.

BY JOHN E. DEWITT, ASST. GEN. MGR., PAPER INDUSTRY DIVISION, INDUSTRIAL NUCLEONICS CORP.

The previous article in this six part series surveyed the economic opportunities in the paper industry and showed how on-line analytical instrumentation, properly used, can make a significant contribution to a company's financial results.

The ideal papermaking system should produce products which meet the customer's specifications at the lowest material and machine cost. Paper has more than its share of customer specified characteristics. These include optical properties, surface finish, caliper, color, flatness, strength properties, and a host of others. Most of these are dependent on basis weight and moisture to some degree.

Traditionally, the acceptable range for basis weight is the nominal specification ±5 per cent. Although this tolerance is quoted as both plus and minus, the lower boundary is most often the reject limit. Moisture is more likely to cause rejects when it is too high. Consequently, the basis weight target is set high enough to avoid violating the low limit; the moisture target is set low enough to keep streaks from exceeding the upper limit.

In both cases, more accurate setting of targets offers great economic potential through material savings and throughput increases. This potential cannot be fully achieved with conventional manual sampling for two reasons. First, the actual value of the basis weight and moisture is not easily determined. Second, the process variation is often so great that a large safety factor must be included in the target setting, so that random samples will not be out of tolerance. This can also be indicative of low product quality; for, in general, the more uniform the basis weight or moisture, the more uniform those other related properties.

The possible harmful effects of using inadequate information from manual sampling to control a process and set targets are not as widely known or fully understood as they should be. The degree to which samples represent a process depends primarily on the sample size and the nature of the process from which it is taken. Defining the complex relationship between the information from sampling and process spread is complicated by the two dimensional nature of the paper sheet.

Mr. DeWitt attended Ohio State, studying electrical engineering. He joined Industrial Nucelonics, Columbus, Ohio, in 1954 in the tobacco division, where he became manager. Mr. DeWitt transferred to the paper division about a year and a half ago.

Nature of sheet variations

Variations of weight or moisture occur in both the machine direction and across the machine. In Fig. 1 the cross machine (CD) and machine direction (MD) variations are depicted above the sheet section. Total variation at any point on the sheet is the sum of both components.

Machine direction variation is made up of several components, as illustrated in Fig. 2. Graph A depicts a steady trend or drift away from target. This could occur in either direction. Graph B shows long term recurring variations with periods of at least several minutes, corresponding to sheet lengths of thousands of feet or longer. Graph C shows high frequency variations corresponding to sheet lengths ranging from a few feet up to several hundred feet. The typical process is a composite of all frequencies, as illustrated in graph D.

The relative amount of variation at each frequency depends upon the sources of the variation and can differ widely from one machine to another. It also depends upon the type of control used as well as the machine condition. Likewise, the amount of cross machine variation depends

Fig. 1. Sheet variations are two directional.

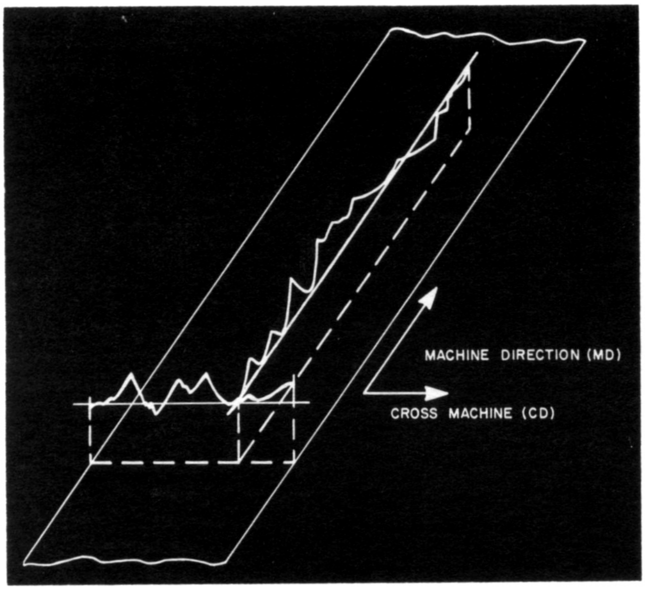
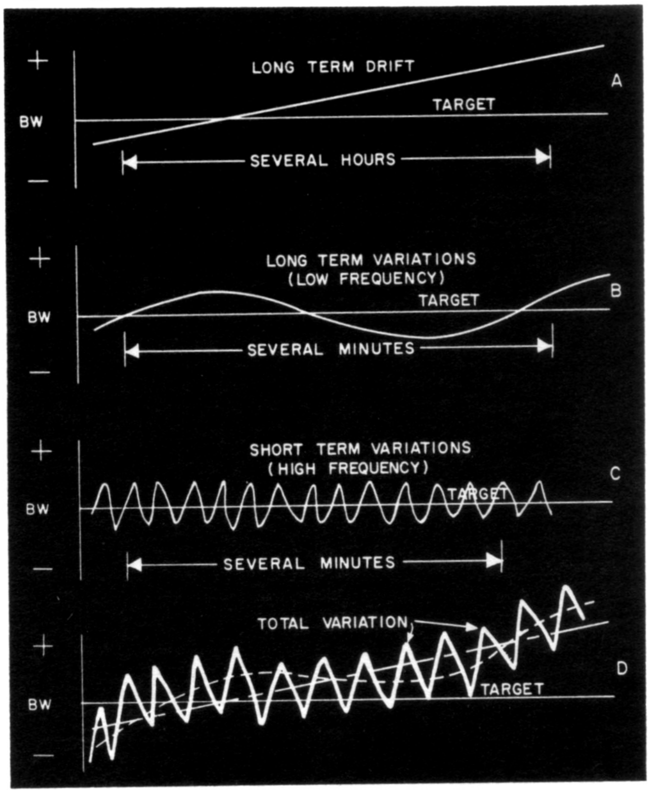


Fig. 2. Types of machine direction variation.



upon the ability of the crew to level the profile. Numerical examples later in the article are based on typical variations.

Inadequacy of conventional sampling

Consider the problem of using information from typical reel end samples to control a process with variations like those shown in Fig. 2D. The objective of such control can only be to remove or reduce the very long term variations and trends. Fig. 3 illustrates the possible effects of sampling and control on this process. To simplify the example, all variations except the trend have been removed.

Assume samples are taken at end of each reel. Sample 1 indicates heavy weight, and a corresponding correction is made. However, the heavy trend continues, and by the end of the next reel, sample 2 indicates heavy, another correction is made, etc. . .

The resulting effect on the actual variation and average is clear. These samples are only partially effective in correcting the trend and achieving the correct average because they do not provide enough information. In fact, each represents the weight *only* at the sample point and is incapable of indicating what happened within the previous reel or predicting what will happen in the next. It should also be clear that sampling and correcting more frequently could reduce the variations and move the average closer to target.

Sample uncertainty increased by short term variations

Suppose the long term and short term components were now present. Fig. 4 shows how they contribute "noise" which always increases the uncertainty of information from samples. Since end-of-reel samples are so short in the machine direction, there is high risk of sampling at peaks which are not at all representative of the average. The operator has no choice but to assume each sample is unbiased and therefore indicates the true average. No correction would be made on the basis of sample 1, since it indicates the weight is on target. No correction would be made even though the actual weight is heavy. Sample 2 would correctly indicate heavy weight.

This uncertainty of sampling information is expressed as "sampling error." The total sampling error includes:

1) Sampling distribution error (E_{SD}): this is the error associated with the choice of the particular point on the sheet from which the sample is taken. It is a function of the sample size and the total variability of the process from which it is taken.

2) Sample handling and sample preparation error (E_{SH}). Examples of sampling handling error are moisture equilibration of samples between the time they are cut from the sheet and the time they are weighed on the scale, and errors in drying out the sample. In a few seconds the moisure equilibration error can produce a one per cent change in apparent basis weight. Sample preparation errors are typically those of determining the area of the sample. A $^1/_{16}$ in. error in cutting each dimension of a one ft. square represents approximately a one per cent error in the basis weight—even if the weight of the sample is precisely determined.

3) Instrument error (E_I): this is the error of the measuring device itself. In the case of an off-line measurement, it is the error of the weigh scale or other laboratory equipment. In the case of an on-line measurement, it is the error of the basis weight or moisture gauge.

The total sampling error is the root mean sq. of the components:

$$E_T = \sqrt{E_{SD}^2 + E_{SH}^2 + E_I^2}$$

The sampling distribution error (E_{SD}) is the most difficult to calculate since it depends, among other things, on the process variation itself.

Fig. 3. Effect of infrequent control on a process with drift.

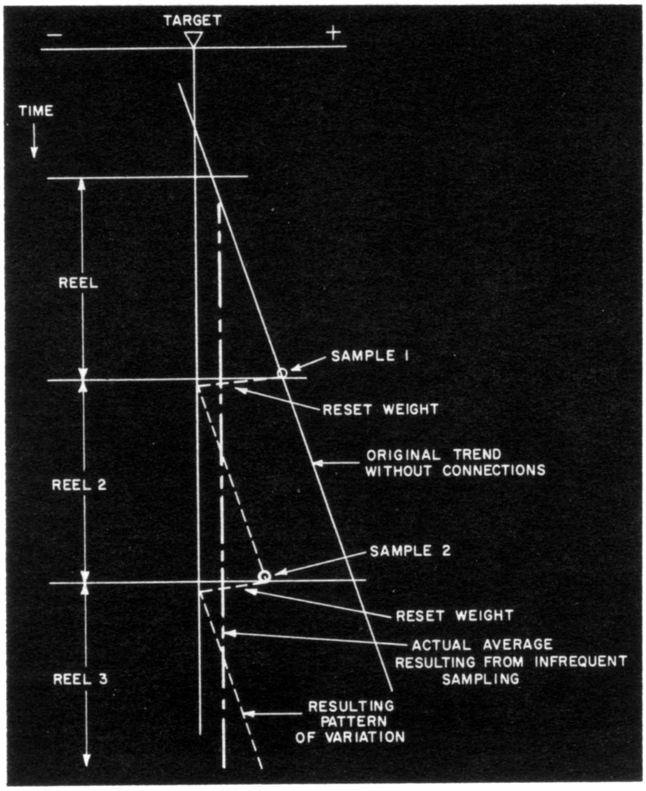
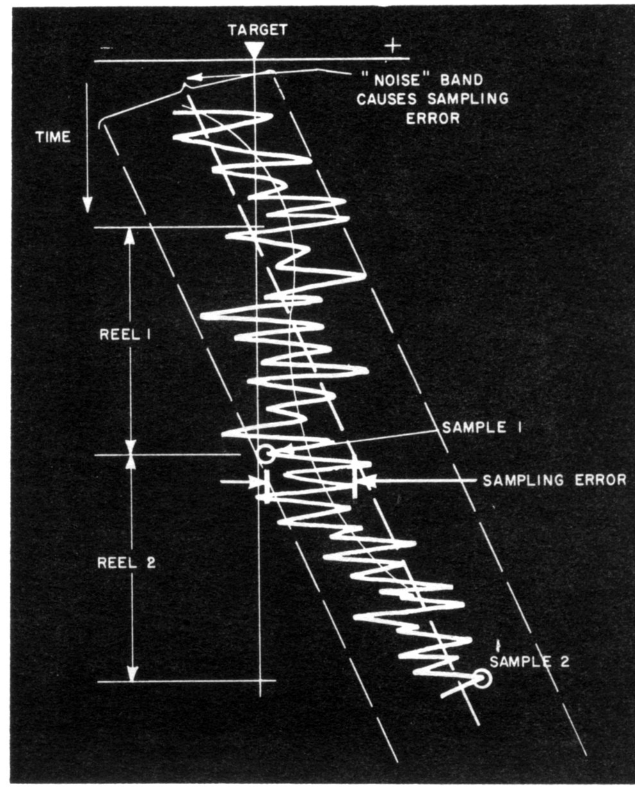


Fig. 4. Effect of high frequency process variations on samples.



The material in the preceding sections is qualitative and helps to clarify the basic concepts of process variation and sampling. Translating these concepts into numerical examples and practical methods requires further analysis and some definitions of terms. Variability of a process can be described in many different ways, such as range, fraction defective, average deviation, standard deviation, and the like. Fig. 5 illustrates some of these terms.

Probably the most efficient and commonly accepted statistic for defining dispersion of a process about its average is the standard deviation (usually represented by σ). The squared value of standard deviation is the variance (V). $V = \sigma^2$. Variance is a highly useful statistic for analyzing processes having several components of variation, for the total variance is simply the total of all the components: $V_T = V_1 + V_2 + V_3 \ldots$ *

Analysis of variance is useful for identifying and isolating some sources of process variation, provided sufficient data can be obtained. This must precede any effort to reduce product variations. Continuous scanning measurement can provide this information.

Scanning gauge gives more information

Another way of viewing the sampling problem and its effect on the determination of basis weight or moisture variations can be seen from Fig. 6. This shows a section of paper with the machine direction and cross direction variations depicted above the sheet. The ft. long sample strip taken at the end of the reel provides very little information about the overall variations in the machine direction. The cross machine (profile) information it provides is useful only if it truly represents the average profile, which is not likely.

A typical scanning gauge sampling path is shown for comparison. The gauge measures continuously and moves across the sheet while the paper is being made. Unlike the end-of-reel sample, the scanning sample is representative of the machine direction variations over the full width of the sheet. The information gathered from such a scan is therefore much more representative of the total sheet variations than is the single end-of-reel sample. For example, at a machine speed of 1,000 fpm and a scan time of 120 seconds, one scan of the gauge represents 2,000 ft. of paper. This contrasts the one ft. long manual sample at reel end! Furthermore, scanning gauges are usually programmed to scan repetitively, thereby effectively providing information representative of the entire reel. The relative sampling errors of the end-of-reel and scanning samples in estimating the average is roughly the ratio of the sq. roots of their respective sizes; $\sqrt{n_1}/\sqrt{n_2} = \sqrt{1}/\sqrt{2000} = 1/45$. That is, the sampling error of the end-of-reel sample is about 45 times as much as that of *one* scan of a continuous gauge. There may be twenty or more scans in one reel, thus reducing the effective "sampling error" of the scanning gauge still further. A more rigorous method of estimating sampling error will be covered later.

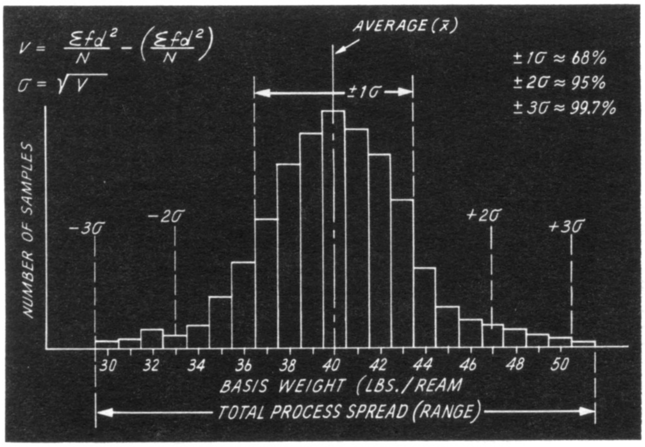
Fig. 6 also illustrates an important advantage of a scanning gauge over one operating at a single set point along the sheet. The scanning gauge signal is more representative of the total sheet area because it includes the cross machine variations. A single point measurement could approach the scanning gauge in accuracy only if machine direction variations were the same at all points across the sheet. In numerous cases it has been observed that this is not the case.

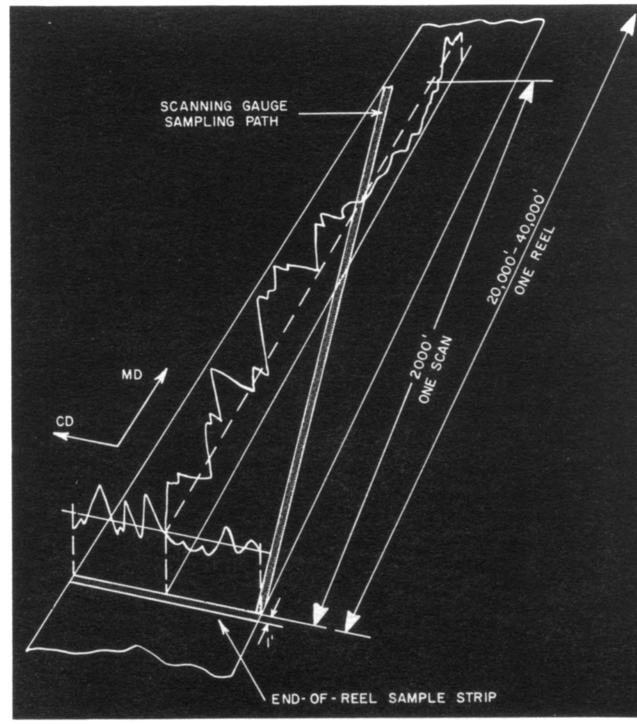
Variance partition—a powerful analysis tool

Another important advantage of the scanning gauge is its ability to provide information to isolate and identify the major components of total process variation. These in turn can be related to specific causes in the process and

* *"Quality Control and Industrial Statistics," by A. J. Duncan, Richard D. Irwin, Inc., Homewood, Ill. (1955).*

Fig. 5. (below) Basic statistical definitions. . . Fig. 6. (right) Comparison of end-of-reel sample and scanning gauge path.





effort applied to reduce them.

Variance partition analysis is a method whereby the total variability of a sheet process, as well as that of the major contributing components, can be determined from approximately 15, or more, diagonal scans. The data from these 15 scans is divided into typically 20 or more cross machine segments per scan and entered into a statistical matrix. From this the variances can be determined for the cross machine component and for the machine direction components covering different frequency bands in the machine direction.

Fig. 7 shows the decomposition of variance that results from this type of analysis. The machine direction variances are subdivided into two frequency bands. The first is the band of frequencies which corresponds to dimensions on the sheet from approximately one ft. up to the length of one scan. These are termed the short term, or high frequency, variations; this variance component is designated V_{STMD}. The second component represents frequencies corresponding to variations in the sheet longer than one scan. This is the long term component of variance (V_{LTMD}), which can be calculated from the distribution of scan averages. Total machine direction variance is $V_{TMD} = V_{STMD} + V_{LTMD}$.

The cross machine component of variance (V_{CD}) is computed from the average or composite profile of the total number of scans sampled. Total process variance is the sum of the individual components: $V_T = V_{STMD} + V_{LTMD} + V_{CD}$. The total standard deviation is $\sigma T = \sqrt{V_T}$.

The composite profile is the average of all the scans included in the analysis. Thus all machine direction components are averaged out, leaving a good clear representation of the true profile.

The example chosen has the following component values: $V_{LTMD} = 0.24$, $V_{STMD} = 0.16$, and $V_{CD} = 0.36$. In this case the units of variance are "pounds squared" so that the standard deviation is expressed in pounds. If there were a reason to do so, we could also compute the standard deviation of each of the components. In this example $\sigma_{LTMD} = \sqrt{.24} = .49$ lbs., $\sigma_{STMD} = .40$ lbs. and $\sigma_{CD} = .60$ lbs. These components could also be expressed as a percentage of the nominal target weight of 40 lbs.

Another output of the variance partition is the "variance profile" which shows the amount of machine direction variation at each of the sampled points across the machine. The not too surprising result of several analyses to date is that machine direction variation is usually not the same at all points across the sheet. This incidentally points up one of the serious problems of single point gauging—how does one select the most representative point? Considerable work yet remains to relate this information to specific causes in the machine. The method has been programmed for a computer to speed the calculations.

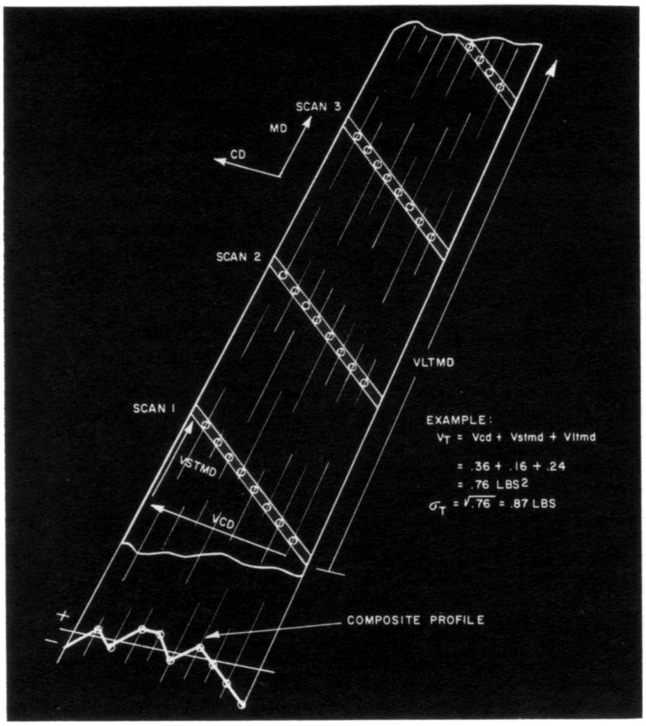
Another practical advantage of variance partition is that the various sampling and instrument errors can be treated just like any other components of variance. They can be added to the actual process variance components to obtain the total variance.

A scanning gauge is the only means of obtaining enough information, non-destructively, to make the variance partition analysis.

New look at sampling error

Information from the variance partition analysis provides new insight into the problem of estimating sampling distribution error and leads to a relatively simple procedure. Consider what really happens in taking a sample, for example a one ft. wide strip across the machine at the end of a reel. Weighing this sample and dividing by its area gives the average weight. Thus this sample has averaged out, or "filtered," all variations within the sample. A sample only one ft. long in the machine direction contains virtually no machine direction variance, but contains all the profile variance.

The sampling distribution error is directly related to the remaining unfiltered process variance, that is—all the variance components not included in the sample area. Since $V_T = V_{CD} + V_{STMD} + V_{LTMD}$, and V_{CD} is "filtered out" by a reel-end sample, the sampling distribution error is caused by the long term and short term machine direction variations remaining. The net effect is to add to the actual process variance a sampling error variance equal



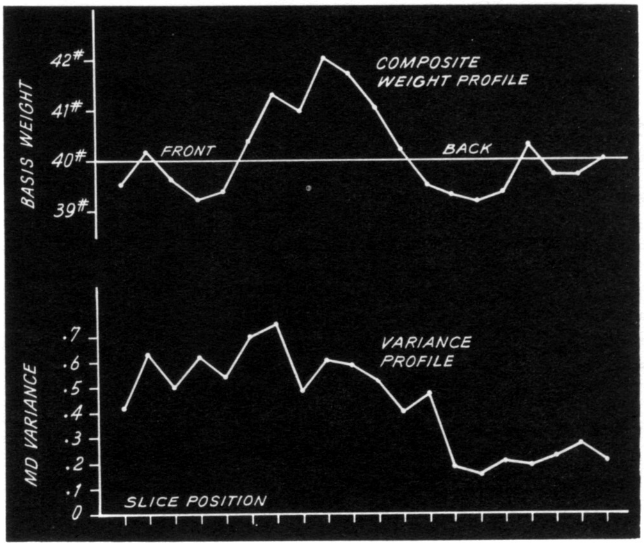


Fig. 7. (left) Variance partition analysis. . . Fig. 8. (below) Results from variance partition.

to $V_{STMD} + V_{LTMD}$.

Scanning gauge "samples" filter out *both* V_{CD} *and* V_{STMD}, leaving only V_{LTMD} to produce sampling error. This explains the smaller sampling distribution errors of a continuous, scanning measurement. An operator could in fact improve the process by using scanning gauge information even if he continued to make the corrections manually. However, a far more dramatic effect occurs when automatic control is used to reduce the long term process variation—sampling error virtually disappears! This is illustrated in this table:

Summary of components of variation

Variance component	Manual sampling, control	Continuous measurement Auto control	A/C & process optimization
V_{CD}	.36	.36	.18
V_{STMD}	.16	.16	.09
V_{LTMD}	.24	.08	.08
$V_{sampling}$	.40	.08	.08
$V_{handling}$	.16	—	—
$V_{sample\ preparation}$	.04	—	—
$V_{instrument}$	.04	.16	.16
Total variance	1.40	.84	.54
Total σ lbs./ream	1.18	.92	.73
2σ " "	2.36	1.84	1.46
3σ " "	3.54	2.76	2.19

Reducing process variation

It is possible to reduce the inherent process variations and thus obtain a narrower total process spread. For example, the average weight value of each scan can be computed automatically and used to control the longer term variations in the machine direction. The slice can be adjusted to reduce the cross machine component of variation. Wet end tuning and other process adjustments can reduce the short term machine direction variations.

The term "compatible automatic control" describes control which has an action time approximately equal to the sample averaging time used as the basis for each correction. The non-compatibility of the manual reel-end sample and the control action based on it is easily seen. The sample itself may represent approximately $^1/_{30}$ of a second, or less, of production. The information from that sample, however, is used to set the stock or steam valve for the next 20 minutes or so of production. In this case, the sample period used as the basis for correction is only 1/36,000 as long as the period over which the control action will be effective.

Corrective action at the end of each scan establishes a more compatible sampling and control system. In a system of this type, the average weight or moisture of the sheet is determined for each scan length (typically $1^1/_2$ to two min. of production). The correction based on deviation of this average from target is applied to the stock valve or the steam valve. The next correction is not made until the process transportation lag time has elapsed. This is typically three min. Thus the sample averaging time is approximately two min. and the effective duration of the control action based on it is approximately three min. This is a far more compatible situation than the 36,000:1 ratio of control period to sample period for the end-of-reel sample.

Compatible automatic control and a scanning gauge effectively reduce the long term variations of greater than scan length to as little as $^1/_5$ to $^1/_3$ of their former value.

Experience proves that the other variance components can be reduced by suitable adjustments of the process. Specifically, the cross machine component can be reduced by "leveling the profile." The variance partition analysis simplifies this by isolating this component and by providing the average profile to work with.

The short term machine direction component of variation can often be reduced by "wet end tuning"(*) and other process analysis and adjustment directed at isolating and removing process disturbances of high frequency nature. Advanced techniques, beyond the scope of this article, such as spectral density analysis, can assist in identifying and

* *"A Practical Approach to the Reduction of Machine Direction and Cross Direction Basis Weight Variations," by Antoni Rocheleau, in TAPPI, Vol. 48, No. 9 (September 1965).*

continued on page 22

continued from page 19

isolating the causes of some kinds of variations by associating their frequency of occurrence with the rotational periods of moving parts.

A thorough "Results Operations Program" has the objective of reducing *all* components of variation, thereby gaining a substantial improvement in the total process uniformity. These programs will be covered in forthcoming articles in this series.

Practical target setting—a numerical example

The concepts developed thus far can now be applied to an example of setting targets for an optimum balance of uniformity and economy. The operating target must be set far enough above the specification low limit to account for all components of process variation and all error components. All of these components are assumed to be random and are combined on a root mean square basis to be certain they are properly weighted.

Consider for example the data of Fig. 7, with an additional one per cent error in basis weight due to moisture equilibration; a one half per cent error due to sample cutting; and a one half per cent error in the weight scale. Based on a nominal 40 lb. sheet, the total process spread and all error components are summarized as follows:

$$\begin{aligned}\sigma_T^2 = & \text{ (sum of all process variance components)} \\ & + \text{(sampling distribution variance)} \\ & + \text{(1 standard error due to handling)}^2 \\ & + \text{(1 standard error due to sample preparation)}^2 \\ & + \text{(1 standard error of the scale)}^2 \\ & = 0.40 + 0.76 + (0.4)^2 + (0.2)^2 + (0.2)^2 \\ & = 1.40 \text{ lbs.}^2\end{aligned}$$

and $(\sigma_T) = \sqrt{1.40} = 1.18$ lbs./ream. This could also be expressed as $\frac{1.18}{40} \times 100\% = 2.95\%$. Corresponding calculations for continuous measurement with automatic control, and continuous measurement with automatic control plus process optimization, are summarized in table 1. The instrument error is assumed to be one per cent in both cases. Note that the data in the previous table approximates the example used for the economic model in the first article of this series.

Optimum targets can now be set with precision and confidence. The target must be far enough above the specification low limit to take into account the total observed spread of the process—including the true process variation, all sampling errors and instrument errors.

Total process spread (range) is approximately 6σ (or from -3σ to $+3\sigma$). If the target is set at a weight 3σ above the low limit, approximately 0.15 per cent of the product will fall below the limit. This may be an unrealistically low reject rate in some cases. If so, the target might be 2σ above the low limit so that approximately 2.5 per cent of the product will be lighter than the limit. Using 2σ limits, a simple table can be constructed showing the target set point in terms of lbs. above low limit for the three modes of operations.

Operation	Target in lbs. above low limit
Manual sampling, manual control	2.36 lbs.
Continuous measurement and automatic control	1.84 lbs.
Continuous measurement, automatic control and process optimization	1.46 lbs.

Continuous measurement, automatic control and process optimization permit the target to be set $2.36 - 1.46 = 0.9$ lbs. lighter than with manual sampling and control. Using our example of a nominal 40 lb. sheet, this amounts to a weight reduction of approximately 2.25 per cent. For a machine consuming $5 million in materials each year, the savings due to reduction of basis weight alone would be over $112,000.

Even greater improvements may also be possible in moisture uniformity and average because the sampling errors can be relatively much larger than for basis weight. Improved target setting offers potential savings in material by substitution of water. The greatly improved uniformity might also permit substantial throughput increases for additional economic benefits. And, of course, the possible competitive advantages of the better uniformity must not be overlooked. These other process and economic opportunities will be explored in future articles.

Summary

This article has described a systematic method by which process variations can be analyzed and targets optimized. Continuous on-line scanning measurement is far superior to infrequent manual sampling in providing information about the process. A number of key conclusions are:

A. Two important factors which limit process economics and product quality are the process spread and the uncertainty about critical product characteristics.

B. Both product quality and process economics can be improved by the application of compatible instrumentation and control systems.

C. The improvement obtained can only be achieved through the use of adequate instrumentation coupled with automatic control, both properly used.

D. Instrumentation provides the basic knowledge of the process which allows precise setting of targets.

E. A good estimate of the product spread can be obtained from a variance partition analysis.

F. The system that has been described for the determination of targets basically produces a higher quality product for the customer at a significantly reduced material cost to the producer. □