Statistical Process Control: Assignable Causes and Data Forecasting

By Thomas M. Tam and Paul R Scherer

A mathematical model Is used to evaluate chemical process data and isolate assignable causes. Any remaining non-assignable cause (process randomness) is then treated by the classical Shewhart method. This methodology enables design of process control strategy, improvement of process consistency, and control of the process at the "State of Statistical Process Control" by demonstrating that the process is stable, predictable and capable.

The objectives of this study were (1) to understand and forecast process behavior, (2) to improve control strategy, and (3) to demonstrate that the process meets criteria for a process at the "State of Statistical Process Control." The criteria were stability, predictability and capability.' The method employed was measurement of the sources of variation (assignable causes) relative to process randomness, and forecasting the value of process parameters based on historical data.

When the sources of variation are understood, the capability of a process can be estimated and a control strategy can be designed in order to meet cost (e.g., cost of frequent process maintenance) and performance (e.g., C_{pk} = 1, 2, 3 . . .) requirements. To accomplish this, (a) assignable causes must be separated from the process data and (b) process randomness must be evaluated, using the traditional Shewhart method.[2]

Data forecasting' is important because it ensures that corrective action, such as replenishment, is taken before bath parameters exceed control limits. The forecasting capabilities of the Shewhart and cumulative sum methods are limited because they allow only a constant process mean with random variation and do not take into account process drift. The method presented here takes into consideration all the assignable causes such as degradation of the chemicals, drag-in or drag-out, consumption, replenishment, evaporation and dilution. Variation is allowed when assignable causes have been shown by history or study to be indifferent to hardware quality. The method is a novel but simple way to correct assignable causes, to perform data forecasting, and to analyze process randomness.

Background

Table 1 shows the equations used to correct assignable causes and evaluate process randomness.

Data Forecasting—The forecasting system uses a set of historical data, M,., (Last Measured Values), in a discrete time series, and the assignable causes to estimate a future value, F, (Latest Forecast Value). The estimate is modified with each new M_t (Latest Measured Value). Equation (1) expresses this concept. The constant, K_1, is a dilution factor, if dilution is necessary during process adjustment. The term K_2 is the drift factor fora given process; it includes process consumption evaporation, drag-in or drag-out, chemical degradation, etc. It is determined by the slope of the process drift with respect to time, using historical data. The term t is the time factor expressed in days. The last term, A, of Equation (1) is the change in concentration from the amount of a chemical added during process adjustment.

Table 1
Equations Used for Data Evaluation

$$F_t = K_1 \times M_{t-1} + K_2 \times t + A \quad (1)$$
$$D_t \;\mu\; M_t - F_t \quad (2)$$
$$\sigma_{GS} = \alpha \times D_t + \beta \times | \sigma_{GS-1} | \quad (3)$$
$$S = D_t/\sigma_{GS} \quad (4)$$

Nomenclature
F_t = Latest forecast value
M_{t-1} = Last measured value
M_t = Latest measured value
K_1 = Dilution factor
K_2 = Drift factor
t = Time factor (days)
A = Change in concentration from the amount of chemicals added
D_t = Latest Delta value
D_{t-1} = Last Delta value
σ_{GS} = Latest geometrically smoothed standard deviation
σ_{GS-1} = Last geometrically smoothed standard deviation
σ_{avg} = Average standard deviation
α and β = Weighting factor
S = Warning signal
UCL = Upper control limit
LCL = Lower control limit

Latest Delta Value, D_t—Assuming that M, and F, are accurate, the difference between them will be zero. In reality, each of the values has its own uncertainties. Equation (2) combines the uncertainties into a single term, D_t. If all the assignable causes have been corrected, D, should be randomly distributed.

Geometrically Smoothed Standard Deviation, σ_{GS}—This concept arose from the Mean Absolute Deviation (MAD) proposed by Trigg.[4,5] The σ_{GS} is a geometrically smoothed value. It represents the recent randomness of the process. Equation (3) calculates the latest σ_{GS}. The two constants, α and β, are the weighting factors. The calculation emphasizes the most recent σ_{GS} and declines geometrically with past history.

Warning Signal (S)—Equation (4) calculates the ratio of D_t/σ_{GS}. If the value is greater than ±3, the difference between F, and M, is statistically significant:

$$-3 \times \sigma_{GS} < D_t > 3 \times \sigma_{GS}$$
$$-3 < D_t/\sigma_{GS} > 3$$

If the σ_{GS}s are relatively constant and the values of S are always less than ±3, then it can be said that the process is stable and predictable.

Upper and Lower Specified Limits (USL and LSL) vs. Upper and Lower Control Limits (UCL and LCL)—Historically the specified limits are the values specified in a contract or established in a standard. The control limits apply to the distribution of the mean, which represents the natural behavior of a process. In order to show that a process is capable of meeting the specified limits, the process spread (6 x o_{avg}) must be less than the tolerance (USL-LSL).[6] Assuming that a process decays continuously because of various assignable causes, the question then becomes: what latitude is allowable in the most practical and cost effective process? Expressed another way, what capability index (C_{pk}) is desirable? C_{pk} is defined as the smaller of either X - LSL/(3 x o_{avg}) or USL - X/(3x o_{avg}), where X is the process mean and o_{avg} is the average standard deviation.

Results and Discussion

A chemical seal process with more than three years of data was chosen as the object of this study. Table 2 shows the chemical parameters of the process being controlled.

Significant Figures—The first concern is the accuracy and precision of the analysis. The data must retain all the significant figures supported by the analytical method, the last digit of each figure representing the inaccuracy of the method. Figure 1 (a) shows a control chart in which process data were rounded to two significant figures. The trace of the graph resembles a staircase, with no apparent trend. Figure 1 (b) shows a control with the same set of data retaining all the available resolution. The data appear continuous and the process trend is apparent. If Equations 1-4 are used to treat the data of Fig. 1 (a), an inaccurate warning signal will be generated. This is because of improper rounding, e.g., the number between 1.84 and 1.75 oz/gal, if rounded to 1.8, causes the σ_{GS} to be close to zero. A small change in the process, or a random analytical inaccuracy, becomes Statistical y significant if, for example, 1.74 and 1.75 oz/gal are rounded to 1.7 and 1.8 respectively. A large and erratic warning signal will result. ' "

Table 2
The Chemistry of a Selected Chemical Seal Process

Solution	Tolerance, (USL-LSL)*	σ_{GS}(6xσ_{GS})	C_p**
Alkaline cleaner	8.0-5.0 oz/gal (3.0)	0.02 (0.12)	25.0
Aluminum etch	16.0-14.0 oz/gal (2.0)	0.2 (1.6)	1.3
Alodine coat	2.0-1.0 oz/gal (1.0)	0.007 (0.042)	23.8
Hot acid etch	16.0-14.0 oz/gal (2.0)	0.3 (1.8)	1.1
Hot alkaline etch	2.0-1.0 oz/gal (1.0)	0.04 (0.24)	4.2
Acid etch	16.0-14.0 oz/gal (2.0)	0.1 (0.6)	3.3
Acid anodizing	13.8-12.4 fl oz/gal (1.2)	0.1 (0.6)	2.0
Neutralizing	7.5-5.5 oz/gal (2.0)	0.03 (0.18)	11.1
Sealer I	2.3-1.7 oz/gal (0.6)	0.05 (0.30)	2.0
Sealer II	8.3-6.9 oz/gal (1.4)	0.25 (1.5)	0.9

*USL = Upper specified limit; LSL = Lower specified limit
**Capability ratio (C_p) = Tolerance/6xσ_{GS}

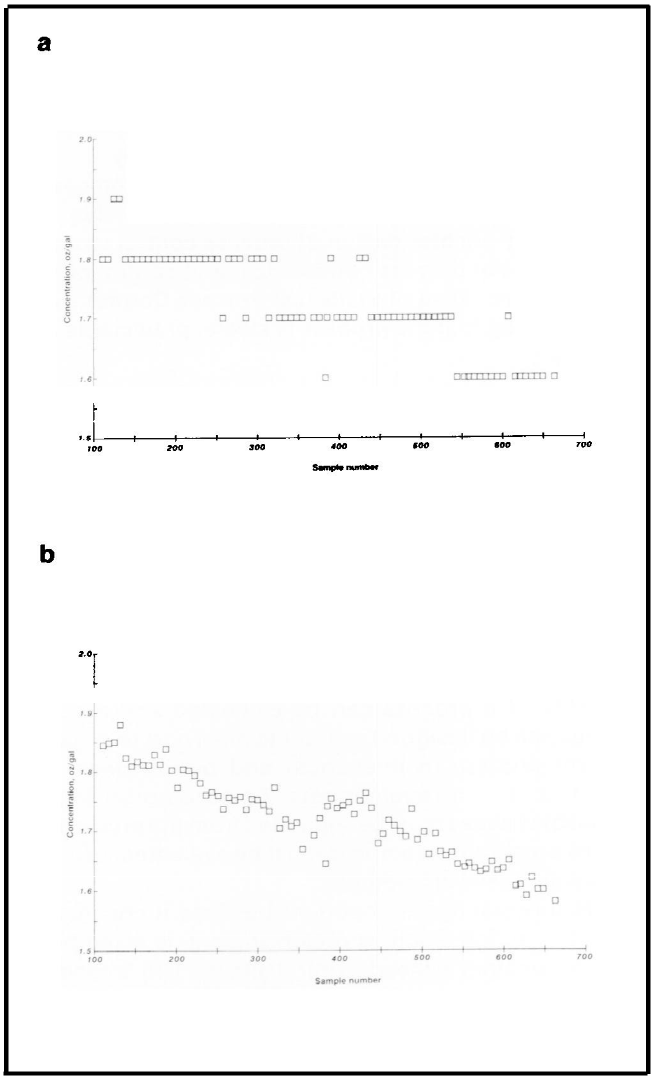


Fig. 1—Control charts of alodine solution. (a) Data rounded to two significant figures. (b) Data at four significant figures.

Assignable Causes—Figure 2 shows the control chart of a neutralizing solution that goes through a cycle of continuous decay and replenishment. The Latest Measured Value, M_t, is shown with the Latest Forecast Value, F_t, superimposed. It is apparent that the two sets of values agree quite well. Also shown in this figure is the Latest Delta chart (see Equation (2), Table 1). Its trace shows no apparent trend and is randomly distributed, evidence that all assignable causes have been corrected and that only process randomness remains. F_t is now being used to forecast the near future of the process and implement preventive maintenance to keep the processes within the defined control limits.

Stability—The geometrically smoothed standard deviation, σ_{GS}, is used to show that the process randomness is stable. The value is calculated using Equation (3), where α = 0.182 (α = 2/(n+1), n = 10 (emphasizing the latest 10 measured

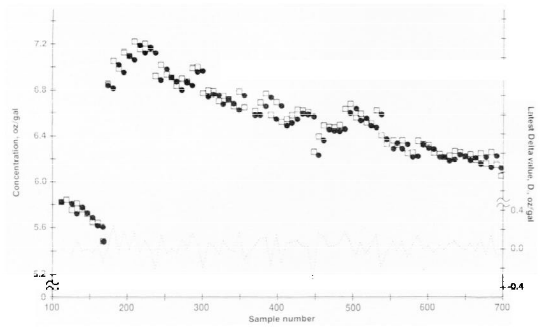


Fig. 2—Neutralizing solution. (a) Control chart with Latest Measured Value, (b) Latest Forecast Value, (c) Delta chart.

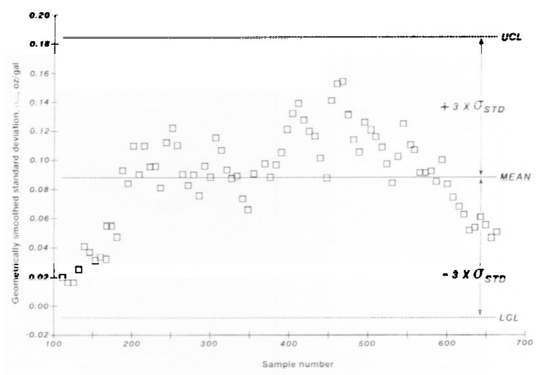


Fig. 3—Geometrically smoothed standard deviation, σ_{GS}, of the neutralizing solution.

values), and $\beta = 1 - \alpha$. Figure 3 is a plot of σ_{GS} vs. sample number. The center line is the average σ_{GS}, and the positions of the upper and lower control lines represent ±3 times the standard deviation of the average $\sigma_{GS,avg}$. A stable process should not have any point lying beyond the ±3 standard deviations limit.

Predictability—Figure 4 shows a plot of Warning Signal S (calculated by Equation 4) vs. sample number for the neutralizing solution. Five values, predating this study, were observed to exceed ±3. Assignable causes were not identifiable; however, the more recent history indicates that this process is predictable.

Capability—The objective was to control the process within a C_{pk} of 2. The capability ratio, C_p, is evaluated first. Column 3 of Table 2 shows the last two years' average standard deviation of the process solution randomness, o_{avg} and Column 4 shows the corresponding C,. Three solutions: aluminum deoxidizer, hot acid etch, and sealer II have C_p less than 2, therefore need improvements in terms of control technique (e.g., better level controller, sampling and analytical technique). As for the remaining solutions, it is thought that the process can be controlled to meet C_{pk} immediately by controlling the assignable causes. The approach taken was to calculate and maintain the new upper and lower control limits. Figures 5 and 6 give before-and-after examples of the new methods for controlling these solutions to meet the desired C_{pk}.

Automatic Data Reviewing System—In our company it is necessary to control over 100 process solutions consistently. The method just described would be cumbersome if applied manually. Computer procedures have been written to handle the routine calculations. Figure 7 shows the block diagram of the procedure. The input data are laboratory analytical results and assignable causes. The output data are the Latest Forecast Value, F_t, the Latest Measured Value, M_t, and the Warning Signal, S.

The calculated results are evaluated according to the logic diagram of Fig. 8. Four different warning signals,

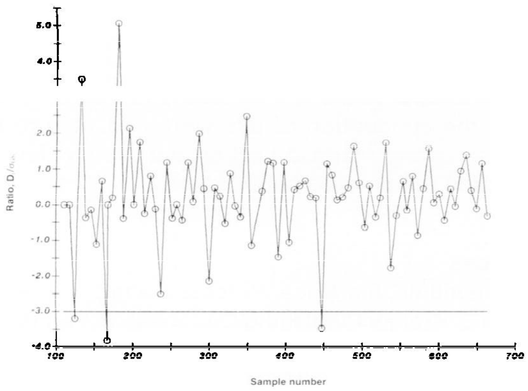


Fig. 4—Control chart on the ratio of D_t/σ_{GS} of the neutralizing solution.

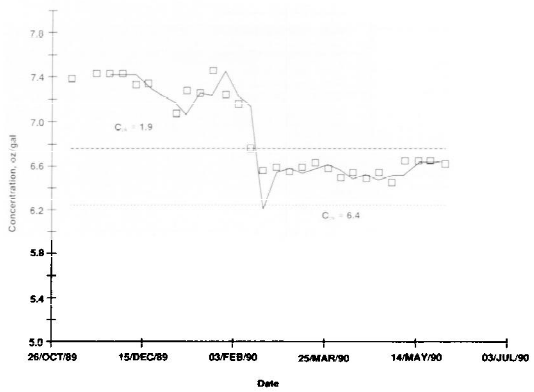


Fig. 5—Example 1 showing improvement on C_{pk}.

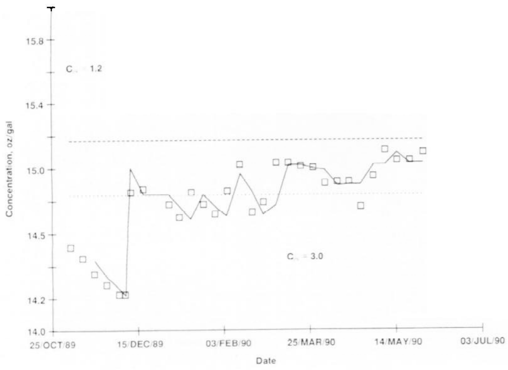


Fig. 6—Example 2 showing improvement on C_{pk}.

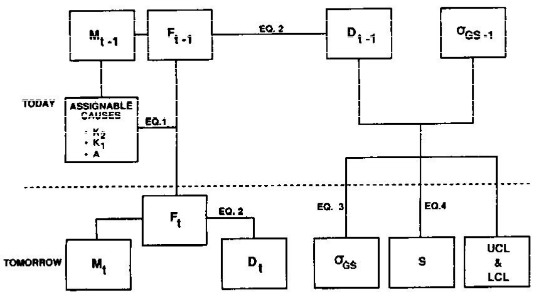


Fig.7—Block diagram of the Automatic Data Evaluation System, using Equations 1-4.

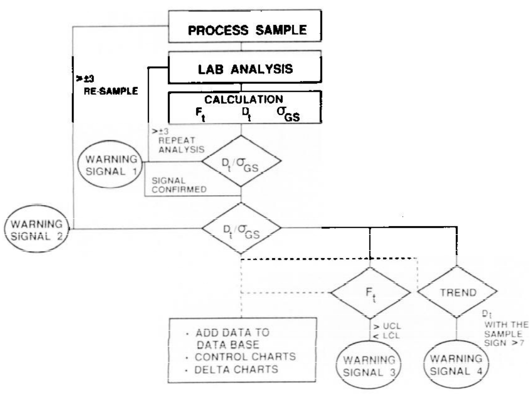


Fig. 8—Logic diagram for data review and for multiple warning signals.

Table 3
Statistical Rules for Posting Warning Signals

Warning signal	Rule	Action Item
1	$D_t/\sigma_{GS} > \pm 3$	Confirm laboratory results
2	$D_t/\sigma_{GS} > \pm 3$ (confirmed laboratory result)	Investigate process solution/equipment Re-sample
3	LSL>F_t<UCL	Process adjustment
4	Number of D_t with the same sign >7	Re-evaluate assignable causes

based on statistical rules and possibility of error, are posted to laboratories and engineers to call for review or corrective action. The statistical rules are given in Table 3. Additional rules may be added when more experience is gained in use of this system, or when less variation is desired. This automatic data reviewing system has improved the integrity of the process control data and decreased the response time for corrective action.

Conclusion

In the application of a data forecasting procedure to control chemical process solutions, simple mathematical equations have been developed to evaluate process data. The essence of these equations is quantitative consideration of assignable causes. By comparing the difference between forecast and measured values, it becomes possible to assess the randomness of the process solution. The randomness is then treated by the classical Shewhart method. In this manner the behavior of the process randomness can be characterized.

In several instances, small but abnormal changes in the process data were observed. They were the result of equipment malfunction, or sampling or analytical error. The sources of assignable variation and process randomness are now understood, permitting determination of optimum process control strategy. Finally, using the data forecasting equation, effective preventive maintenance of process solutions can be implemented.

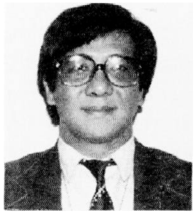
Acknowledgments

The authors appreciate the collaboration of Dr. Ross A. Quinn at the conception of this work and wish to thank Perry C. Cansanay, Marta O. LeBron and Leonor Manansala for providing the process control data.

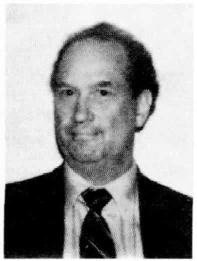
References

1. R.R. Clements, *Statistical Process Control and* Beyond, Robert E. Kreiger Publishing Co., Malabar, Florida, 1988; pp. 1, 23.
2. S.C. Puri, *Statistical Process Quality Control,* Standards-Quality Management Group, Ottawa and Washington, DC, 1984; p. 77.
3. R.C. Brown, *Smoothing, Forecasting and Prediction of Discrete Time Series,* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1963.

4. D.W. Trigg, *Operations Research Quarterly, 15, 271 (1964).*
5. G.S. Cembrowski, J.O. Westgard, A.E. Eggert and E.G. Toren, Jr., *Clinical Chemistry, 21,* 1396 (1975).
6. S.C. Puri, *op. cit.,* p. 127.
7. W.J. Genck, *Chemical Processing, 52,54* (Dec., 1989).



Tam



Scherer

About the Authors

Dr. Thomas M. Tam is a material process engineer in the Materials and Process Control Center of Lockheed Missiles & Space Company, Inc., 1111 Lockheed Way, Sunnyvale, CA 94086. He has 10 years' industrial experience in electrochemical plating, chemical sensors, metal surface finishing, and statistical process control. Dr. Tam earned his PhD. from the University of California at Davis. Before joining Lockheed, he was a post-doctoral fellow for two years at the Lawrence Berkeley Laboratory.

Dr. Paul R. Scherer la a senior staff engineer In the Materials and Process Control Center of Lockheed Missiles & Space Company, Inc. He has 22 years' Industrial experience in analytical chemistry, instrumentation, computers and process control. He holds a PhD in analytical chemistry from the University of Arizona at Tucson.