

Applied Chemical Process Statistics - Bringing Industrial Data to the Classroom

Milo D. Koretsky
Department of Chemical Engineering
Oregon State University

There are three kinds of lies: lies, damn lies and statistics.

-Mark Twain

ABSTRACT

Alumni, co-op interns and the department's Industrial Advisory Board have all identified Analysis of Process Variation, Statistical Process Control, and Design of Experiments as important elements missing in OSU's undergraduate chemical engineering curriculum. Even though many students had, in fact, taken introductory statistics, they were unable to apply it in an industrial setting. In response to this input, a 1 credit elective course, *Chemical Process Statistics*, was introduced in Fall 1996. It was offered again in Fall 1997. The initial course offerings were well received (3.77/4.00 and 3.91/4.00 on student evaluations). The intent is to integrate this course into the ChE core as a 3 credit course.

This course provides a brief introduction to the use of applied statistics in the chemical process industry. Whenever possible, concepts are introduced through examples. Industrial data has been provided by Merix and Hewlett-Packard, on copper and tantalum etch processes, respectively. Thus, concepts discussed in class are immediately applied to real industrial data. Statistical analysis of the data is discussed in terms of the physical process. In this way, the statistics and the science are coupled. In the case of the copper etch for Merix, student email accounts are linked to the process control software so that students get "Process Out of Control Messages" in real time. Course concepts are cemented with a plant trip at the end of the term to see statistics in action.

INTRODUCTION

Undergraduate chemical engineering education emphasizes *analysis* and, then, *design*. In the typical curriculum, the majority of the technical credit hours are devoted to fundamental science (e.g., general chemistry, physical chemistry, and organic chemistry) and engineering sciences (e.g., mass and energy balances, thermodynamics, transport processes, reaction engineering, process dynamics and control). The student is then asked to *synthesize* this material in unit operations and then the capstone design course. However, the majority of graduates are hired as Process Engineers whose main focus is on *production*. Topics such as statistical process control, process capability and capability indices are essential to manufacture quality products at reduced costs.¹ In fact, upon accepting their first job offer, most entry level engineers, enroll in in-house statistics related courses such as *Practical Data Analysis*, *Statistical Process Control*, and *Design of Experiments*.²

When alumni, co-op interns and the department's Industrial Advisory Board were surveyed about strengths and deficiencies in content in the OSU chemical engineering program, the overwhelming short-coming was in the area of these production-based statistics topics. However, an examination of the transcripts of graduating seniors over several years revealed the majority had, in fact, taken introductory statistics. Apparently a formal statistics course plus the *ad hoc* inclusion in senior laboratory did not provide students with sufficient exposure to integrate into the production-based statistics they need in industry. Consequently, a 1 credit elective course, *Chemical Process Statistics*, was introduced. The intent of this course is to form a bridge between the abstract mathematical formalism presented in statistics classes and the process technology of chemical engineering. This approach is especially useful for the more hands on learners, many of whom go on to be process engineers.

Chemical Process Statistics provides a brief introduction to the use of applied statistics in the chemical process industry. Whenever possible, concepts are introduced through examples. Industrial data has been provided by Merix and Hewlett-Packard, on copper and tantalum etch processes, respectively. Thus, concepts discussed in class are immediately applied to real industrial data. Statistical analysis of the data is discussed in terms of the physical process. In this way, the statistics and the science are coupled. It is intended to be a "high level" overview and to compliment, not replace, the more in depth but theoretical classes offered at Oregon State University.

The objectives of *Chemical Process Statistics* are:

- 1) Develop an awareness of the utility of statistics in assessing experimental data and operating industrial processes.
- 2) Describe the basic concepts and nomenclature associated with applied statistics, Statistical Process Control, and Design of Experiments.
- 3) Work through real industrial examples to gain experience with these tools.
- 4) Utilize computer software to aid in statistical analysis.

SUMMARY OF COURSE CONTENT

To put the specific case studies in context, a brief description of the content is provided below. These topics can be found in many introductory statistics textbooks.^{3,4}

- Discussion of Distributions: central tendency, dispersion shape.
- *All processes show variation*
- The sources of the variation can be grouped into two major classes: *common causes* and *special causes*.
- When special causes are present the distribution varies and is not predictable. Special causes must be eliminated before process improvement can begin. A process with special causes present is said to be *out of control*.
- When common causes are present the distribution is constant and predictable. Process improvement can reduce the *variability* associated with common causes.
- A process which exhibits only common causes is said to be *in control*, no matter how large the dispersion is. However, it might not be *capable*.

- Capability is a measure of a process's ability to consistently produce products which meet the prescribed specification limits. Capability indices, such as C_p and C_{pk} , may be calculated and determine the capability of a process. Typically C_{pk} of 1.33 or higher is considered capable. C_p and C_{pk} are defined as follows:

$$C_p = \frac{\text{Specification Tolerance}}{\text{Process Spread}} = \frac{USL - LSL}{6\sigma},$$

$$C_{pk} = \text{minimum} \left[\frac{\bar{x} - LSL}{3\sigma}, \frac{USL - \bar{x}}{3\sigma} \right]$$

where USL and LSL are the upper and lower specification limits, respectively, σ is the standard deviation of the sample and \bar{x} is the mean.

- Use of *control charts*, to monitor processes, assess variation and determine process capability
- Use of *Design of Experiments (DOE)* for planning experiments which efficiently use resources to obtain the maximum information about the effect of input variables (factors) on the output (response) of a process and the interaction of the input variables.

CLASS EXERCISES

Students were asked to pick a process in their everyday life which could be sampled at least 3-4 times/ week, (weight, time to drive to school, wait in line at the grocery store, time of run...). They regularly collected data and set up a run chart on excel. As concepts were introduced in class, they were asked to apply it to their own data. These included, identifying any special causes of variation, calculating upper and lower control limits from the data which only show common cause variations, and calculating measures of process capability, C_p and C_{pk} , for their process. Students were asked to report the latter values in class. Most of their processes had capability indices well below 1. One outcome from this exercise is the realization of Deming's statement:

A state of statistical control is not a natural state for a manufacturing process. It is instead an achievement, arrived at by elimination, one by one, by determined effort, of *special causes* of variation.⁵

As concepts are introduced in class, students are asked to apply them to real sets of data that might interest them, rather than random sets of numbers. They are also asked to draw conclusions, i.e., *apply* the statistics. For example, they were asked to construct Box plots from different sets of data in the *AICHE Salary Survey*.⁶ Among the conclusions from this analysis were: starting salaries don't really matter (the spread for starting salaries is a lot tighter than later in one's career); those with more management responsibility received more pay; and women were not paid as much as men.

INDUSTRIAL CASE STUDIES

The heart of *Chemical Process Statistics* is applying the concepts listed above to real manufacturing data from a chemical process. Statistical analysis of the data is discussed in terms of the physical process. In this way, the statistics and the science are coupled. The two major cases that have been used are a tantalum etch process from Hewlett-Packard and a copper etch from Merix.

Example I: Tantalum Etch process (batch vs. semi-continuous)

Data from two process alternatives in the manufacture of ink jet printer pen heads at Hewlett-Packard corporation are compared. A schematic of these process alternatives is shown in Figure 1. The data represent the measurement of the width (μm) of a portion of thin-film tantalum that is defined through a wet chemical etch. The original process (Fig 1a) is a batch process in which 24 wafers (1 lot) are all etched at once in a bath of etchant. The etchant acids include $\text{HAc}:\text{HNO}_3:\text{HF}$ at a specified ratio (not privy to us). The bath is used until it is no longer suitable for etching, then it is repoured. This is a timed Ta etch lasting 2:40 hours. Data is collected by a IVS measuring which has a reported accuracy of 0.0096 mm. This new process (Fig 1b) is a continuous process in which each wafer is processed individually in a tool that is constantly bleeding old etchant solution while gaining fresh solution. There are six processing chambers in the tool, so the first wafer goes into chamber 1, the second to chamber 2 etc. It is not unreasonable to assume that all six chambers do not perform exactly alike, so the tool can be thought of as really six separate tools.

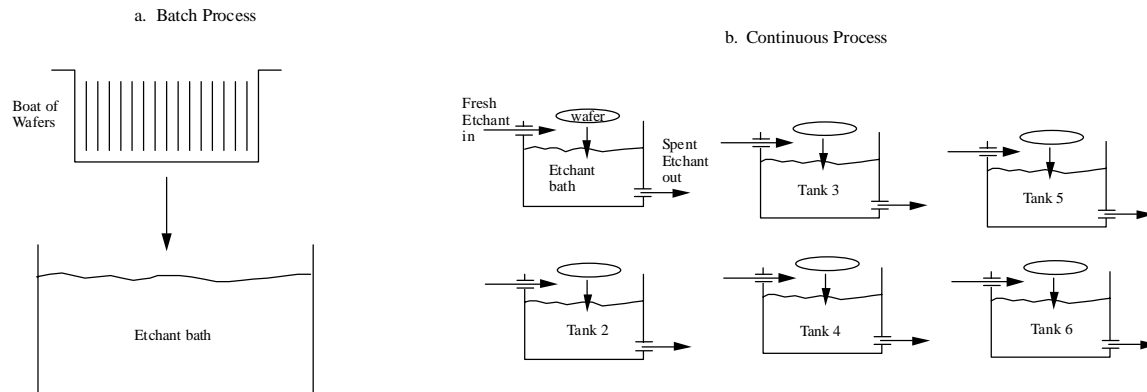


Figure 1. Comparison of the a) batch and b) continuous Ta etch processes from Hewlett-Packard The data was taken as follows: 11 lots of wafers were measured, 10 wafers measured from each lot, 5 "sites" were measured on each wafer. Students were asked to perform summary data such as lot to lot, wafer to wafer and site to site means and standard deviation. They also constructed control charts from these data. Conclusions reached included:

- The lot to lot variation is lower on the continuous process, because the etchant solution is kept at steady state, while in the batch process, the etchant solution is constantly changing over time.

- The wafer to wafer variation (within a lot) is lower on the batch process because all the wafers in the lot see identical conditions while the wafers in the continuous process are being processed individually at unique conditions. (Even though the continuous process is at steady state, there are still slight variations in etch rate amongst the 6 chambers)
- The continuous process represented an improvement over the batch process.

Example II: Real Time Control of Copper Etching in Printed Circuit Board Manufacture

A principle unit process in the manufacture of printed circuit boards is the patterned copper etch. This process defines the “wiring” of the electronic circuit which connects the components which are mounted on the board. A schematic of the process flow is shown in Figure 2. Copper is etched by the following reaction in which cupric ion (Cu^{2+}) reacts with metallic copper to form cuprous ion (Cu^+):

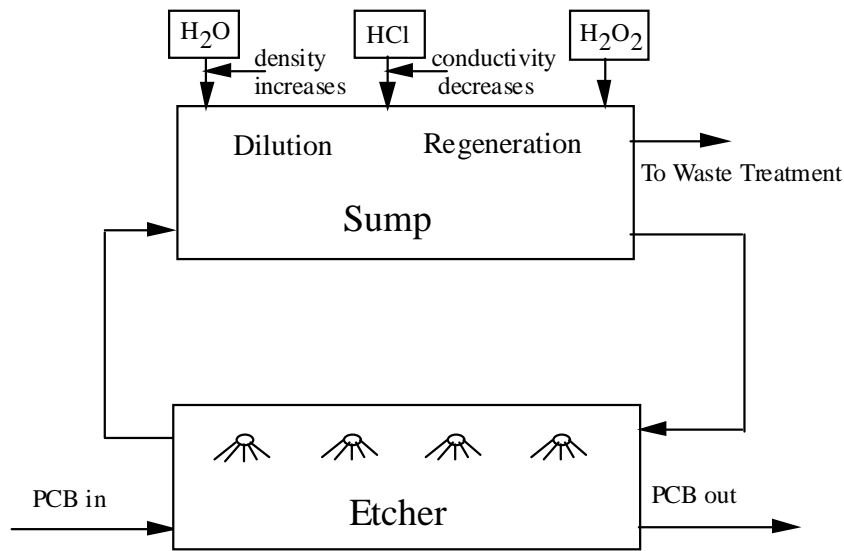
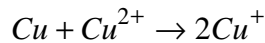
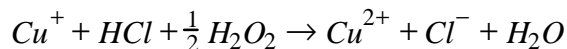


Figure 2. Copper etching of printed circuit boards (PCB) and etchant regeneration.

Etchant is then regenerated as follows:



As time goes on:

1. Cu^{2+} accumulates in the bath as the copper clad is etched.
2. HCl is consumed by reaction 2

3. H_2O_2 is consumed by reaction 2

To control this process we must

1. Determine how much H_2O to add to dilute the cupric ion accumulation.
2. Determine how much HCl to add to account for how much is diluted by control 1 above and by consumption in reaction 2.
3. Determine how much H_2O_2 to add to account for how much is diluted by control 1 above and by consumption in reaction 2.

Attention was focused on items 1 and 2.

Each student was asked to subscribe to a class email list. During the quarter, they got data *in real time* from laboratory analysis of Cu and HCl concentrations. They were asked to incorporate these data into a control chart, and include upper and lower specification limits as well as the target value. An example of such a chart is shown in Figure 3. They were also asked to identify special causes. For the sake of this exercise, they looked at four possible indicators: (i) a single point outside the specification limit; (ii) seven points in a row on one side of the center line, i.e., a run; (iii) five points all going in the same direction, a trend; and (iv) two points in a row just inside the specification limit.

The class list was also incorporated into the Merix process control server so that the students received “out of control” messages that went to the process engineers at Merix, as shown in Figure 4. These messages not only stimulated student interest but also allowed students to see that “out of controls,” occurred in bunches, i.e., they were not spread out in regular intervals. Whenever possible, the physical cause of an out of control message was presented. These stemmed from technical reasons, such as an improper value for the conductivity controller thermal compensation coefficient, operational reasons, such as start up of a new line (shown in Figure 3), and incidental reasons, such as failure to turn on a density controller after a new hydrometer was installed.

One particular benefit of this system is the intricacy of the control. It is desired to measure and respond to the concentrations of Cu^{2+} and HCl . However, they cannot be measured on the production line. Therefore, specific gravity and conductivity are tracked and correlated to

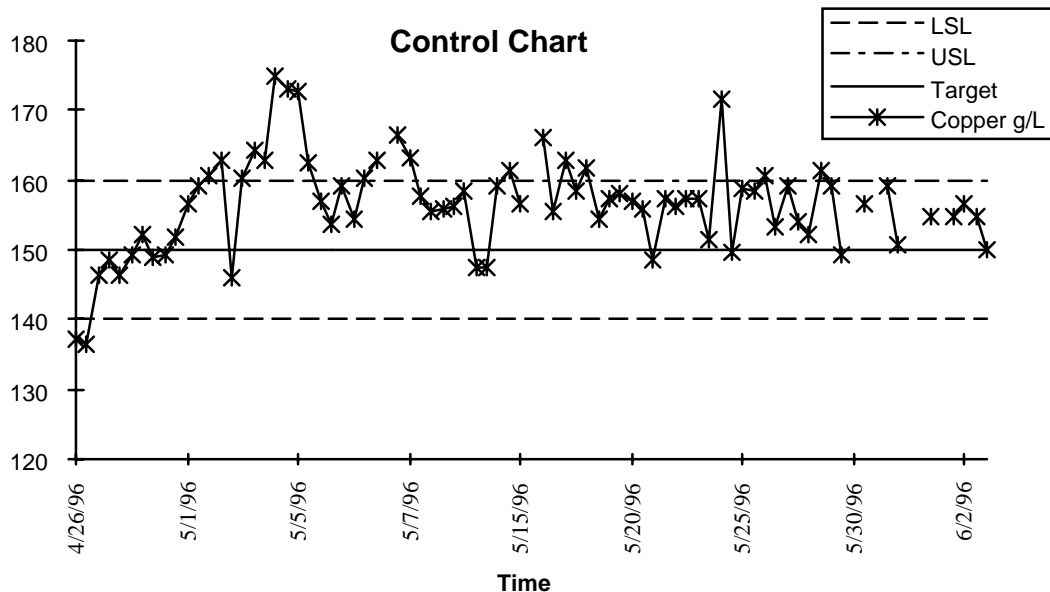


Figure 3. Control chart constructed by students from data obtained on Merix Cu etch line.

concentrations via laboratory quantitative analysis. When the conductivity falls below a “low” setpoint, concentrated hydrochloric acid is added to the chemical sumps until the conductivity reaches an upper setpoint. Specific gravity is monitored by the position of hydrometer float in relation to an inductive sensor. When the density of the solution increases, the hydrometer rises. When the float rises above the sensor position, the inductive contact opens. Water is added until the density decreases and the float re-makes the sensor contact. The conductivity controller setpoints are numerically entered in the PLC program and can be changed. The density controller sensor position can be adjusted. This system also includes an adjustable “stop” which prevents the float from falling below the sensor. If the stop were not there, a decrease in density would cause the float to drop, opening the contact and adding water, thus further decreasing density.

```

Lab Production

At 01:52 on 10/16/96  Copper of 6850 DES 3 Etcher

PROCESS OUT OF CONTROL MESSAGE
Concentration/data = 148.5265
Tank Spec Limits   :  USL =    160.0  LSL =    140.0
Tank Control Limits:  UCL =    159.3  LCL =    149.2
History Comments:

ERROR CONDITION:  OOC 2 SIGMA (LOW)
PLEASE REPLY WITH AN ASSIGNABLE CAUSE OR CORRECTIVE ACTION

```

Figure 4. “Out of Control” message students received as email

One of the problems with this control system is interactions between the measured variables, i.e., their relation to the chemical constituents. Adding concentrated *HCl* not only increases

conductivity, but also specific gravity. Cu^{2+} also does the same. It is not hard to imagine that a control system can be configured that adds *HCl* when the density is too low and adds water when the conductivity is too high will also work in this application (and is, in fact, used in other systems). What this means is that the system can take a long time to stabilize. An adjustment for Cu^{2+} will impact *HCl* and vice versa. The system can also stabilize at three different points. For instance, at a particular setting, one can have “good” Cu^{2+} and *HCl* readings or “high” Cu^{2+} and “low” *HCl* or “low” Cu^{2+} and “high” *HCl*. Lab analysis is used to avoid these situations.

Plant Trip

The course culminated in a plant trip to Merix to see statistics in action. This tour included inspection of the etcher from which data had been obtained, a tour of the analytical laboratory to see how Cu and HCl measurements were made, a presentation of the process control software Merix has developed and how it works, and a general walk through of the entire PCB manufacturing process. They were also presented with historical documentation of the improvement of the processes by tracking of C_{pk} in time. Finally, many students were excited to see the same control charts they constructed in homework assignments, in use and presented on the walls of the factory.

ASSESSMENT

There has been no formal attempt to assess the effectiveness of Chemical Process Statistics on OSU ChE graduates in industry. In fact, the first students to take the class (F 1996) have just graduated (S 1998). However, the initial course offerings have been well received. Student evaluations were 3.77/4.00 and 3.91/4.00 for Fall 1996 and Fall 1997, respectively. This compares to the College of Engineering average of 3.15/4.00.

Anonymous student comments from the course evaluation follow :

“This was an excellent introductory course for ‘laying down the basics’ of statistics in ChE language. Your examples of process control from industry and specific situations at Merix really brought home the material and why it’s so useful.”

“I thought the class was informative and insightful. This provides a good link between ChE topics and statistics.”

“The 2 examples, Merix & HP, were great. I learned more about how to handle real life systems in this course than ST 314 (Introduction to Statistics for Engineers) in much less time.”

“This class was very useful at putting things from my internships in perspective.”

“I really learned from this class. I think it should be required for ChE majors since statistics is a tool that engineers should have. Knowing how to get and take data is what we learn, but to know what it means is essential.”

An additional measure of the impact of this course was evident during the plant trip. Students were interested, engaged the engineers at Merix in detailed question about control strategies, and generally appeared comfortable with concepts presented in class.

SUMMARY

A 1-credit elective course *Chemical Process Statistics* has been developed to form a conceptual bridge between formal statistics courses and the process technology of chemical engineering. This course emphasizes statistical application to high volume production. Examples from processes at Merix and Hewlett-Packard were used to provide a “hands-on” context. The course has been well received by students. Plans are to increase it to a 3 credit course, with additional time spent on design of experiments.

ACKNOWLEDGMENTS

Without the help, support and commitment of industry this course could never have been developed. The author is especially grateful to Michael Matthews who integrated the class into the Merix statistical process control software. Information from, support of, and technical discussions with Adrian Kriz and Lori Thompson of Hewlett-Packard and Mike McMaster and Eric Warninghoff of Merix have also been of great value.

REFERENCES

1. Shunta, J.P., “Use of Statistics to Identify Process Control Opportunities,” *Chemical Engineering Progress* **92** (11), 47 (1996).
2. Eckert, R.E., “Applied Statistics Are ChE Educators Meeting the Challenge?” *Chemical Engineering Education*, Spring 1996, 122.
3. Montgomery, D.C., *Introduction to Statistical Quality Control*, 2nd Ed., Wiley, New York (1991).
4. Box, G.E.P., W.G. Hunter and J.S. Hunter, *Statistics for Experimenters*, Wiley, New York (1978).
5. Deming, W.E., “On Some Statistical Aids towards Economic Production,” *Interfaces*, **5**, 5 (1975).
6. *1996 AIChE Salary Survey*, American Institute of Chemical Engineers, NY (1996).

BIOGRAPHICAL INFORMATION

MILO D. KORETSKY Assistant Professor of Chemical Engineering
Oregon State University, Corvallis, OR 97331-2702 email: koretskm@cmail.orst.edu
B.S., M.S. University of California, San Diego; Ph.D. University of California, Berkeley
Research interests in materials processing for the microelectronics and high technology industries.