



STATISTICAL METHOD MAKES A COMEBACK

Chemical and drug firms warm to experiments that open a statistical **WINDOW INTO REACTIONS**

RICK MULLIN, C&EN NORTHEAST NEWS BUREAU

ANYONE WHO HAS washed dishes can tell you that a lot of stuff clings to dirty plates, glasses, and silverware after a meal. And anyone who has developed an automatic dishwasher detergent to clean the mess can tell you that a lot of chemicals need to come together to make it work. With all the variables involved, finding the right molecule to perform any one function in a detergent could involve hundreds of experimental test runs in the development lab.

Or it could take 20.

BASF recently used a statistical modeling method called design of experiments (DOE) to test its Trilon M chelating agent in phosphate-free detergent formulations. By establishing a means of manipulating multiple variables at one time—the effects of temperature and washing time on the interaction between chelating agents and surfactants, for example—the company was able to significantly reduce the number of tests needed to create effective detergents with Trilon M.

Novomer, a green polymers company, had similar results in developing a second-generation catalyst that converts waste gas streams into plastics and chemicals. Scott D. Allen, the firm's vice president of catalyst development, received the American Statistical Association's annual Statistics in Chemistry Award last year on the basis

of the catalyst work. The corecipient of the award, Bradley Jones, is a principal research fellow at JMP, the supplier of the software both BASF and Novomer used to execute DOE modeling in the laboratory.

BASF and Novomer are among the chemical and pharmaceutical companies to have achieved breakthroughs in recent years using DOE. The technique, created nearly 100 years ago, nowadays relies on computers to generate graphic images of a complex system's response to changes in variables. It is experiencing a renaissance as scientists find that they can use DOE to reduce the number of experiments they conduct and get a better handle on their chemical processes.

Behind the technique's rebirth is the rising importance of statistical analysis in research and manufacturing. Also key is a 2011 guidance document from the Food & Drug Administration promoting its use in submitting New Drug Applications. And advances in DOE software are delivering new insights to chemistry researchers.

The adoption of DOE, however, has been hampered by a cultural friction between chemical engineers, who crave tools that

TEST CYCLE BASF used design-of-experiments methods to develop dishwasher detergent formulations using its new chelating agent.

promote multi-variable experimentation, and laboratory scientists steeped in the "first principles" approach of the scientific method.

The chemical industry was directly involved in the early industrial applications of DOE. "The first big uses of designed experiments in the chemical industry probably took place in England during World War II," says Jones at JMP, which is a division of SAS, the industrial software giant.

George E. P. Box, a statistician and engineer at the storied British chemical firm Imperial Chemical Industries, put a statistical technique initially used in agriculture to work after the war to boost the efficiency of industrial manufacturing. Box went on to coauthor a seminal textbook on DOE with John Stuart Hunter, a former statistician at American Cyanamid.

THE CHEMICAL SECTOR continued to take the lead into the 1970s. "DuPont actually taught a course on DOE that they gave to their own internal staff and offered to others as a service," Jones says. DuPont's Quality Management & Technology Center trained researchers from other firms into the 1990s. Over time, however, the chemical industry has fallen behind other sectors—the semiconductor industry, in particular—in applying the statistical model, according to Jones.

DOE evolved differently in the pharmaceutical industry. "Clinical trials are designed experiments," Jones says. "Statistical modeling is necessary because you want to be very sure that whatever you are testing is better than a placebo and not toxic." It's important, he adds, that patients taking the drug and the placebo match demographically.

"But I think it would be great if the drug companies were a little less conservative," Jones says. "I think that they and the FDA have created a bit of an unhealthy intellectual attitude over the years. Drug companies are so afraid submissions will be judged harshly that they are not taking advantage of a lot of opportunities."

FDA addressed this timidity, he says, with the issuance of its Process Analytical Tech-

"When time is important and material costs are high, you need a robust process."

nology guideline in 2004, which allowed companies to make fundamental changes in manufacturing plants for quality and efficiency purposes. It followed up in 2011 with Quality by Design (QbD) guidance that allows parallel evaluation of development and manufacturing data in New Drug Applications. The latter guidance explicitly greenlights DOE-generated data in FDA filings.

AMID THE RENEWED activity, JMP—pronounced “jump”—is about to release a new version of its software, including a feature called definitive screening. Developed by Jones, definitive screening allows for factors to be measured at three levels rather than the two levels used in classical screening designs. JMP’s software includes other statistical modeling tools and can be integrated with its parent company’s database products.

Stat-Ease, a small company specializing in DOE software and consulting services, has also seen a rise in interest among chemical and pharmaceutical companies. At the vanguard, according to Mark J. Anderson, the firm’s general manager, are chemical engineers, who generally have more statistical training than chemists. Engineers, he adds, tend to share practices across industries, helping to spread DOE techniques.

“FDA’s QbD guidance has been a big thing for the process industries,” he says. “It’s directed at the pharma industry, but it migrates. There has been a cyclical move of DOE back to chemistry from pharma.”

Stat-Ease’s DOE software, called Design-Expert, combines the basic two-factor approach to designing experiments with a technique called response surface methods, a regime of using a sequence of designed experiments to obtain an optimal outcome. The approach particularly supports chemical mixture design. As with all software for DOE, Design-Expert produces a graphic representation of how a system responds to changing variables. “The product of the experiment is a visualization,” Anderson says.

Design-Expert generates three-dimensional topographical or contour maps that can be overlaid on each other to shade out system responses that don’t meet requirements. “It leaves a window or a sweet spot” illustrating the mix of variables and conditions that produce a desired performance or effect in a product or process, Anderson explains.

James S. Dailey, global technical key account manager in BASF’s care chemicals division, worked on the Trilon M detergent

MILESTONES IN DESIGNED EXPERIMENTS

■ In the 1920s, British statistician **Sir Ronald A. Fisher**, considered the father of modern statistics,

develops the first designed experiments using two planting fields for multivariable testing in an effort to boost crop yield. He coins the phrase “design of experiments” (DOE) to describe the method.



■ During World War II, **George E. P. Box**, a statistician working for Imperial Chemical Industries in the U.K., applies



Fisher’s principles to analyze nerve gas used by the Germans. He goes on to apply DOE to industrial applications. Box coauthors “Statistics

for Experimenters,” a seminal text on DOE, in 1978, with John Stuart Hunter, a Princeton University statistician who had worked at American Cyanamid, and William Gordon Hunter, a statistician at the University of Wisconsin, Madison.

■ In the 1970s, DuPont’s Quality Management & Technology Center trains DuPont employees on DOE and offers the training to other companies. The company continues the service into the 1990s.

■ In 2011, the Food & Drug Administration issues guidance on a DOE-based quality management approach called Quality by Design. FDA explicitly encourages the inclusion of multivariable data in New Drug Applications. The agency’s goal is to improve the quality of drugs and ensure supply-chain safety.

tests. A 20-year veteran at the company, Dailey had conducted his first designed experiment only a few years earlier. He was won over by the technique’s ability to account for multiple variables in a limited number of experiments and produce a graphic representation of how changes in any of the variables affect a system.

“We can model in a place where we

haven’t even done an experiment, and we can go back and confirm it,” Dailey says. “We can take new data and put it back into the model and improve it.”

DOE was not a popular technique in the industry until about 2008, Dailey acknowledges. “Even now, I would say, it’s not being used as much as it should be.” The availability of software necessary to do the math, he says, has been a limiting factor. “When I started here, all the software was on a central computer. If you tried doing DOE on one of these, it would be very difficult.”

Dailey’s lack of statistical training was also a barrier. “I’ll be honest. I am not a statistician,” he says. “I’m a chemist. I wonder if what we teach in college and grad school isn’t part of the issue here. We teach the scientific method—you change one variable at a time, and you do a lot of experiments.”

Novomer’s Allen says he started to learn about DOE when the company hired a vice president of R&D who had used the technique at Eastman Kodak and whose brother, also a former Kodak statistician, works at JMP. Novomer agreed to be a beta user of the version of the software incorporating definitive screening.

Allen, a chemist by training, says he initially understood DOE to be a tool for process scale-up. “We started using it for that purpose,” he says, “but JMP’s new features allowed us to apply it to discovery.”

The process paid off by giving researchers new insights into how catalysts perform. “We didn’t want to carry over assumptions from the previous generation of catalysts, so that meant taking a look at quite a few factors,” he says. “One of the factors we identified in that system had a much different influence on catalyst activity than we had seen previously. It jumped out at us. That got us thinking more about what is going on in the system.”

ANALYSIS OF DATA from the experiments led to something more radical than basic system optimization, Allen says. “It was really a matter of going all the way back to the basic fundamentals of that reaction and looking at them without any bias.”

Consolidating experiments saves time and money, Allen says, but the real value of DOE is the window it gives researchers into a chemical reaction. “If you have an experiment or system that is out of control or has sources of random error, you

AMERICAN STATISTICAL ASSOCIATION (BOTH)

are not necessarily going to see that doing traditional experiments.”

Drug firms and their suppliers say they have stepped up the use of designed experiments in development work. “It is definitely increasing, particularly in late-stage development,” says Robert Duguid, associate research fellow in the chemical development department at Albany Molecular Research Inc., a contract research and manufacturing firm. “We are proposing to all our clients that we do DOE analysis for robustness, testing processes before they go into commercial production.” The company uses Stat-Ease software.

FDA’s QbD guidance has advanced the adoption of DOE in the pharmaceutical industry, Duguid acknowledges. “And it’s certainly affected what we are doing.”

Walter Kittl, executive vice president for technical operations at Siegfried, a Swiss pharmaceutical chemistry firm, says DOE has become important in its drug formulation business. “We use it heavily in micronization,” Kittl says, noting that the milling process involves a broad range of variables. “It is impossible to do linearly, so you need to do it with design-of-experiment software, taking a statistical approach.”

Kittl, who formerly worked at Roche, says spray drying of pharmaceutical chemicals is another area where DOE has caught on at Siegfried. “When time is important and material costs are high, you need a ro-

bust process,” he says. Siegfried uses Stat-Ease’s Design-Expert and software from the Swiss firm Aicos Technologies.

Drug companies also have identified the need to gain better control of their development and manufacturing process. The FDA guidance not only encourages the application of statistical tools such as DOE, according to Julia O’Neill, director of engineering at Merck & Co.’s West Point, Pa., facility, it is forcing companies to use them.

“There is a new expectation that manufacturers will be able to demonstrate they are using statistical methods for monitoring results and monitoring processes to ensure they remain in a validated state,” O’Neill says. “That drives an expectation back upstream that statistical methods will also be used during the design and qualification stages.”

O’Neill, a chemical engineer who came to Merck from the chemical maker Rohm and Haas eight years ago, acknowledges that, prior to the recent push from FDA, the chemical industry had made more headway than the pharmaceutical industry on DOE. In doing so it also paved the way for the drug industry. “There is a whole set of tools just waiting to be taken off the shelf and used to achieve similar goals in pharma,” O’Neill says.

BUT THE CULTURAL divide between chemical engineers and chemists may be harder to cross. “I have heard people say

that statistics are for people who don’t know the science,” she says. “I totally disagree.”

Some chief executive officers say they are anxious to close the divide. “Concepts like statistical design of experiment have gained a lot of ground in recent years, but for mainstream chemists they are somewhat anti-intuitive,” says Siegfried CEO Rudolf Hanco. “They are 180 degrees against what you learn in grad school, where you learn to look for a cause-consequence relationship and the fundamental rule is you only change one parameter at a time.”

In an industrial setting, where the focus is on efficiency, changing one parameter at a time takes too long, Hanco says. “With DOE, you create a hyperdimensional space from which you determine the highest optimum of whatever parameter you are looking at. The fact that this is against the nature or at least the education of most technical people has led to a situation where the uptake was very slow in the industry.”

But Hanco, like others, points to DOE’s power to win converts who see the results of using multivariable experimentation. Dailey at BASF would count as a convert. “DOE saves time and money and allows you to predict performance,” he says. “It’s allowed us to increase the speed of development, and I think it is increasing our innovation. I get really excited about it.” ■

Reprinted from *Chemical & Engineering News*, Vol. 91 No. 13 pp. 25–27

Copyright © 2013 by the American Chemical Society and reprinted by permission of the copyright owner.