

Chemometrics and PAT: An Opinion on Current Status and Recommendations for the Future

Jerry Workman, Jr.

*Kimberly-Clark Corp., Analytical and Measurement Technology Group,
West Research and Engineering, 2100 Winchester Road, Neenah,
Wisconsin 54956 U.S.A. [Mailto:jworkman@kcc.com](mailto:jworkman@kcc.com)*

“In chemistry, one’s ideas, however beautiful, logical, elegant, imaginative... are simply without value unless they are actually applicable to the one physical environment we have; in short, they are only good if they work...”

– R. B. Woodward

Chemometrics Defined...

- "Chemometrics is what chemometricians do." Anon.
- "Chemometrics has been defined as the application of mathematical and statistical methods to chemical measurements." B. Kowalski, Anal. Chem. 1980, 52, 112R-122R.
- "Chemometrics is the chemical discipline that uses mathematical and statistical methods for the obtention in the optimal way of relevant information on material systems." I. Frank and B. Kowalski, Anal. Chem., 1982, 54, 232R-243R.

Chemometrics Defined... Page 2

- "Chemometrics developments and the accompanying realization of these developments as computer software provide the means to convert raw data into information, information into knowledge and finally knowledge into intelligence." M. Delaney, Anal. Chem. 1984, 261R-277R.
- "...research in chemometrics will contribute to the design of new types of instruments, generate optimal experiments that yield maximum information, and catalog and solve calibration and signal resolution problems. All this while quantitatively specifying the limitations of each instrument as well as the quality of the data it generates." L. S. Ramos et al., Anal. Chem. 1986, 58, 294R-315R.

Chemometrics Defined... Page 3

- "Chemometrics, the application of statistical and mathematical methods to chemistry..." S. Brown, Anal. Chem., 1986, 60, 252R-273R.
- "Chemometrics is the discipline concerned with the application of statistics and mathematical methods, as well as those methods based on mathematical logic, to chemistry." S. Brown, Anal. Chem. 1990, 62, 84R-101R.
- "Chemometrics is the discipline concerned with the application of statistical and mathematical methods, as well as those methods based on mathematical logic, to chemistry." S. Brown et al., Anal. Chem. 1992, 64,22R-49R.

Chemometrics Defined... Page 4

- "Chemometrics can generally be described as the application of mathematical and statistical methods to 1) improve chemical measurement processes, and 2) extract more useful information from chemical and physical measurement data." J. Workman, P. Mobley, B. Kowalski, R. Bro, Appl. Spectrosc. Revs. 1996, 31, 73-124.
- "Chemometrics is an approach to analytical and measurement science based on the idea of indirect observation. Measurements related to the chemical composition of a substance are taken, and the value of a property of interest is inferred from them through some mathematical relation." B.Lavine, Anal. Chem. 1998, 70, 209R-228R.

Chemometrics Defined... Page 5

- **“Chemometrics (this is an international definition) is the chemical discipline that uses mathematical and statistical methods, (a) to design or select optimal measurement procedures and experiments; and (b) to provide maximum chemical information by analyzing chemical data.” Bruce Kowalski, in a formal CPAC presentation, December 1997**

Chemometrics Defined... Page 6

- **Chemometrics is so often linked with Process Analytical Chemistry, again defined by Kowalski “as the discovery and development of new and sophisticated analytical methods for use in-line as an integral part of automated chemical processes.*” Some have said that process analytical chemistry is 90% hardware and 10% chemometrics. To an engineer that quantitative statement means one may be able to do without it.**

*B. Kowalski (personal communication) 12/97

Chemometrics Defined...Page 7

- What we have then is a process – we make measurements – we collect data – we use chemometrics to obtain information – we review the information and attain real knowledge. If chemometrics is difficult to clearly define and communicate, what are its advantages and disadvantages?

Advantages of Chemometrics . . . Page 1

- Chemometrics provides speed in obtaining real-time information from data.
- It allows high quality information to be extracted from less resolved data.
- It provides clear information resolution and discrimination power when applied to second, third, and possibly higher-order data.
- It provides methodology for cloning sensors – for making one sensor take data “precisely“ as another sensor.

Advantages of Chemometrics . . . Page 2

- It provides diagnostics for the integrity and probability that the information it derives is accurate.
- It promises to improve measurement quality.
- It improves knowledge of existing processes.
- It has very low capital requirements – it's inexpensive.

Case studies: When chemometrics is combined with PAT it CAN provides multiple benefits...Page 1

(1) safer plant and process operations through real-time monitoring and prevention of potentially dangerous process upsets; (2) assurance that processes and plant environments are in compliance to environmental regulations; (3) an increase in process plant operability through timely adjustments in processes possible using real-time data;

Case studies: When chemometrics is combined with PAT it CAN provides multiple benefits...Page 2

(4) improved product quality through maintenance of tighter control limits; (5) minimization of waste products through process optimization; (6) product production cost minimization through tighter target limits and more accurate production scheduling; (7) optimization of production capacity resulting from increased process operability and continuous product quality verification

Case studies: When chemometrics is combined with PAT it CAN provides multiple benefits...Page 3

(8) elimination of *possibly* the greatest challenge to 100% compliance and analytical accuracy for a process - namely, sampling error; (9) provides information and technology for real-time feedback learning and control.

Disadvantages of Chemometrics...Page 1

(1) Anyone with a computer can generate complex numerical solutions that can be impossible to interpret; (2) Complex mathematical solutions can leave plenty of room for misinterpretation unless strict guidelines (viz. standards) are followed; (3) Chemometrics requires a change in one's approach to problem solving from univariate to multivariate thinking - this

Disadvantages of Chemometrics...Page 2

requires a paradigm shift and a process model understanding closer to reality; (4) most “best practices” still need to be collected and codified into useful standards - requiring *much* cooperative effort between practitioners, theoreticians, and regulatory groups.

Old Vs. New Scientific Method

A new method requiring not a thought ritual, but rather a method involving many inexpensive measurements, possibly a few simulations, and chemometric analysis. The new method looks at all the data from a multivariate approach, whereas the old method requires the scientist's assumed powers of observation, from a univariate standpoint, to be the key data processor.

The Old Scientific Method (used for hundreds of years)

- Stating the problem
- Forming the hypothesis
- Observing and Experimenting
- Interpreting Data (traditionally *univariate* – ‘ponder and grimace’ stage)
- Drawing overly simplistic conclusions related to complex processes and phenomena
- Assume the process is in control

The New Scientific Method (for problem solving using PAT)

- Measure a process (any chemical phenomenon or process)
- Analyze the Data (multivariate analysis)
- Iterate when necessary
- Create and test (verify) model
- Develop a more fundamental, multivariate understanding of complex processes
- Make sufficient measurements to *verify* the process is in control

**One designed experiment is
worth a thousand educated
opinions...**

And so what is the information content of a thousand well measured results vs. that from a presumed process model with a few selected measurements?

Reluctance to Change...Page 1

APPROVED CHEMOMETRIC-BASED METHODS?

- **The American Society for Testing and Materials (ASTM) Practice E1655-00. ASTM Annual Book of Standards, West Conshohocken, PA 19428-2959 USA, Vol. 03.06 (2001) 573-600.**
- **AOAC Official Methods of Analysis, 15th Ed., Arlington, VA 22201-3301 USA, Vol. 15 (1990), 74-76.**
- **P. Williams and K. Norris (eds.), Near-Infrared Technology in the Agricultural and Food Industry, American Association of Cereal Chemists, St. Paul, Minnesota, Vol. 1 (1987) 291.**

Reluctance to Change...Page 2

CHEMOMETRIC-BASED METHODS IN COMMITTEE

- **Note for Guidance on the use of NIRS by the Pharmaceutical Industry and the Data to be Forwarded in Part II of the Dossier for a Marketing Authorisation, EMEA/CVMP/CPMP, 09/01 draft.**
- **Guidelines for the Development and Validation of Near Infrared (NIR) Spectroscopic Methods, Pharma. Analytical Sciences Group (PASG), 01/01 draft.**

The typical “process chemometrics” project₁

- **The process decisions are in the domain of the chemical engineer, plant manager, and quality group. Their process decisions are based upon process modeling, temperature, pressure, flow, and mass balance, which are considered adequate and always have been adequate. Decisions are made in the plant and through various engineering groups. These decisions are made based upon past experience and current academic training. The reason then that changes are so slow, and most resist change to include chemometric-based sensors is precisely due to resource deficiencies - in time, talent, attention, motivation, and economic incentive.**

The typical “process chemometrics” project₂

- The process engineer and manufacturing personnel require motivators to change behavior patterns, such as: recognition for accomplishment, demonstrated process improvement, no risk, convenience, economical choices; thus the risk/reward ratio must be near zero.
- The company has a separate list of requirements, such as: improved process performance, increased profits, maintenance or improvements in quality, convenience, economical to implement, low risk; thus the [Rewards/(Risk + Cost)] ratio must be a very large number.

The typical “process chemometrics” project₃

- Chemometrics supplies a perfect fit to these requirements by providing the expertise (time and talent) into the resource equation, by minimizing cost by data analysis techniques requiring some sensor and computer time, and demonstrating a potential benefit in understanding a process providing priceless information for improving it. *Risk is minimized due to the flow of real-time information.*

Check-list for making chemometric sensors work ₁

- Test underlying assumptions continually and thoroughly at lowest possible cost.
- Prepare multiple alternatives solutions
- Commit to implementing technology but not to one use or application of technology – look for multiple technologies and multiple uses
- Control time and cost commitment issues.
- Avoid overload of staff - 2 substantial projects each is optimum.

Check-list for making chemometric sensors work ₂

- Is there an internal customer market for this technology or approach?
- Can we deliver the technology reliably and cost-effectively?
- Can we take small exploratory forays into less challenging opportunities, learning and adapting as we proceed?
- Continually codify and diffuse the information relative to project, including risks, potential benefits, and proposed solutions

TABLE 1

Industrial chemometrics attribute map

	BASIC	Discriminator	Energizers
Attribute No.	Non-negotiable	Differentiator	Excitor

TABLE 2

Chemometrics Project Portfolio (Potential Cost vs. Risk Score = 1-9)

Cost of Project

High (3)	3	6	9
	2	4	6
	1	2	3
Low (1)			
Technical Risk (uncertainty of Sensor(s) and chemometrics used in approach)	Low (1)	Medium (2)	High (3)

TABLE 3

Chemometrics Project Portfolio (Project Selection Score = 1-27)

Cost vs. Risk Score

High (6+)	18-27	12-18	6-9
Medium (3+)	9-15	6-10	3-5
Low (1+)	3-6	2-4	1-2
Value to Corporation (\$)	Low (3)	Medium (2)	High (1)

Select lowest value projects first and don't often venture above **6**

The new value rules in technology₁

- Neolithic Age – Using renewable energy sources to produce food (spear requires high human energy input)
- Industrial Age – Using non-renewable energy sources to produce goods (Mechanical clock requires less energy input and has higher information content than spear)
- Information Age – Substituting information for energy to produce knowledge intensive goods (Pentium chip requires less energy than clock but has much greater information content)

Leading the changes? ₁

- **New technologies are usually inferior to present state of the art**
- **Today's technology leaders dismiss the new technology**
- **New technology moves forward very rapidly after take-off**
- **Success creates the seeds of complacency due to arrogance**
- **Competency traps itself in the status quo**
- **To survive, the competent must seek to replace themselves with new competencies**

Leading the changes? ₂

- Old technology insists on improved execution of the ‘wrong’ thing, not an emphasis on doing the ‘right’ thing...

Leading the changes? ₃

Stages of change:

- **Denial**
- **Resistance**
- **Negotiation**
- **Acceptance**

Leading the changes? ⁴

Needs of People going through change:

- Empathy
- Information
- Involvement

Leading the changes? ⁵

Managing adaptation and change

- GATHER FAST, CHEAP INFORMATION, AND CORRECT ANY PROBLEMS
- LOTS OF INFORMATION (NOT DATA) GIVES POTENTIAL LEARNING FOR SUCCESS
- THE SIZE OF YOUR INFORMATION PILE INDICATES THE LEARNING POTENTIAL FOR FUTURE SUCCESSES

Leading the changes? ⁶

Requirements for Chemometric-based PAT

- STANDARDIZED ALGORITHM CODES FOR CALIBRATION
- STANDARD SAMPLES FOR INSTRUMENT MONITORING AND CALIBRATION TRANSFER
- STANDARD 'OUTLIER' DETECTION METHODS
- STANDARD ANALYZER FUNCTIONALITY TESTS
- STANDARD CALIBRATION AND VALIDATION PROTOCOLS (BASED ON SOUND PRINCIPLES OF EXPERIMENTAL DESIGN)

**Gather the expertise to write
useful consensus standards
with periodic revision**

Note that standard methods will lag somewhat behind the newest technologies until an experience base is collected for deriving a consensus opinion.

Example: E1655-00

ASTM Stnd. Vol. 3.06, 2001, pp. 573-600.

- Peer reviewed by approx. 100 ‘skilled in the art’
- Includes scope and use descriptions, instrumentation requirements, calibration mathematics, statistics, pre- and post-processing, Outlier statistics, calibration and validation protocols, troubleshooting guidelines, quality statistics, protocols for updating models, terminology, and a questionnaire to check compliance with the Standard.