MS2 Data Mining & Visualisation Project Update

AJM Consulting

Issue 3

January 2007

On-Line Analysis

The rate of adoption of the innovative data mining technology in the MS2 Process Analysis System has been astounding, as you can see from the editorial on page 2. One question has been dominant in virtually all the projects we have undertaken so far - "when can MS2 be used on line?"

So far, the data mining tools developed for MS2 have enabled analysis of historical data, with the aim of increasing understanding of the complexities of process manufacturing. The business benefit of this is obvious.

But the capability for early detection of an encroaching process abnormality is at least as valuable and in most cases probably more so.

We are therefore responding to this need by commencing development of an on-line version of MS2, which will analyse multivariate process data as it becomes available, and provide valuable information when the process starts to become sub-optimal.

Our plan is to build on the success we have had so far, by extending the system into the on line world, whilst still providing the ease of use, flexibility and powerful visualisation for which MS2 has become so appreciated.

As with development of MS2 so far, we propose to invite a select number of companies to become development partners. This approach has worked exceedingly well since we gained valuable knowledge of actual process problems and the partners gained solutions to their specific problems at much reduced costs, helped in part by the funding assistance awarded to us by Yorkshire Forward.

If you would like to find out more about the substantial opportunities for manufacturing improvement which the new developments will enable, and would like to discuss how your company could benefit by becoming a partner in this exciting project, just contact us. We will be delighted to discuss our plans in more depth with you.

Inside this issue:

Development of the MS2 Process Analysis System received funding assistance from DTI's Research & Development Fund, administered by Yorkshire Forward.

MS2 Training

In response to demand from users we are providing the first training course in data mining and process analytics on January 31. Whilst this obviously uses the MS2 system it is also of more general interest to people who are thinking about using multivariate technologies. The course content includes a general overview of process analytics and data mining, data import and pre-screening, univariate statistics, parallel coordinate visualisation, multivariate statistics and principal component analysis, together with practical issues of using these techniques.

There will also be "hands on" experience in using MS2 effectively, with particular emphasis on effective analysis of differing types of real problems.

The course will be held at our Europarc office in Grimsby.

We are particularly pleased to be working closely with the Centre for Process Analytics and Control Technology at Newcastle University and are grateful for the assistance we receive from this renowned centre of multivariate research and development.

Onward and Upward!

After a very successful first year, we look forward to new challenges in 2007

Welcome to the third issue of our data mining project newsletter. As we look back on the first year of the MS2 Process Analysis System development it is startling to realise how far we have come. As of the start of January 2007 we have 20 sites using or implementing the MS2 system - far more than we originally forecast.

Virtually all have ordered maintenance contracts, enabling them to take up enhancements as these are released. Several are now using MS2 for multiple projects. The first training course is at the end of January and is heavily subscribed.

In previous newsletters we have described the various key technologies on which the system is based, and as major developments occur these have been described as well. The latest, Multiway Principal component Analysis (MPCA), which integrates batch trend data analysis into the system, is described on page 3.

Whilst 2006 saw the launch of MS2, 2007 brings us some exciting challenges. There are more algorithms in the pipeline, new and enhanced data screening functions and lots of powerful new visualisations. We will continue to work closely with the Centre for Process Analysis and Control Technology at Newcastle University and, of course, are delighted that Professor Julian Morris has joined us as Associate Director.

The big news of the new year is on-line functionality development (see page 1). For companies which want early benefits from this technology, please contact us now.

Alan Mason

Managing Director

Case Study - Fibres Worldwide

Fibres Worldwide Limited is a manufacturer of Acrylic fibres for industrial and speciality end uses based in Grimsby, NE Lincolnshire. The business focus is in two areas: Carbon Fibre Precursor and Speciality Textile Fibres. Carbon Fibre is processed into flame retardant and insulating materials, carbon-carbon composite aircraft brakes as well as other Carbon Fibre applications. It is used in continuous and chopped form to produce low weight, high strength composites for a wide variety of end uses ranging from portable electronic equipment to high performance bicycles and sports equipment to wind turbine blades and oil rig risers. The Speciality Textile product range includes the Amicor family of Anti-bacterial and Anti-fungal fibres, the Outlast range which utilizes Phase Change Material to regulate body temperature and the Super White range of high colour purity textile fibres.

Manufacturing is a continuous process involving several interconnected stages leading to a high degree of complexity and interaction. As with all manufacturing processes there is a need to understand which are the key process parameters in order to focus on these for both controlling the process to meet the stringent requirements of Aerospace and Medical end users and also for developing the process to break in to new, more demanding markets. To assist in this process, Fibres Worldwide has invested in the MS2 Process Analysis System, which has already provided insights into process complexity and will continue to be used by the Company in the future.

"Following our introduction to the concept of Data Mining and Multivariate Analysis we were aware of the potential benefits to a process as complex as ours but we were concerned about the suitability for a continuous process. By working closely with AJM, they were able to manipulate data from a large number of databases and produce a Parallel Co-ordinate Visualization with a timescale. Subsequent analysis including Principal Component Analysis has yielded valuable information about Key Process Parameters and has also shown where data that is difficult and expensive to collect is adding little or no value. We are looking forward to continuing to work with AJM to further develop the package and are now considering including Customer Data also."

Neil Barker,

Fibres Worldwide Ltd

Multi-Way Principal Component Analysis

Multi-way PCA extends the ability to provide multivariate analysis to data which represents changing parameters within a batch. An obvious example of this is the common situation of a temperature curve which varies as the batch progresses. Usually, several trends are recorded (such as pressure, agitator motor current demand and so on) rather than just a single trend. It is therefore used when there are multiple samples taken for one or more variables in a batch.

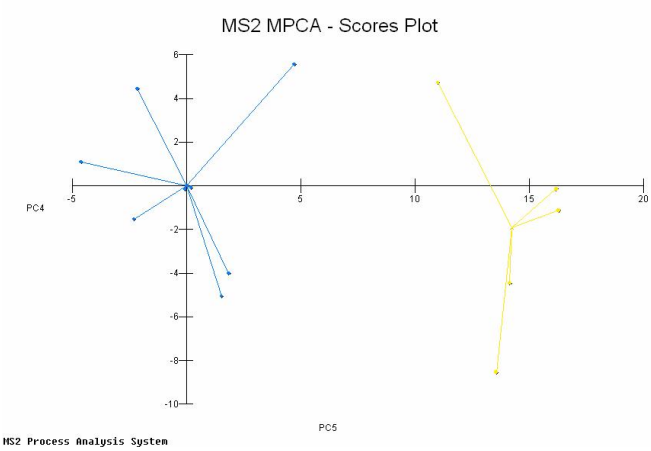
In order to do this, the multi-way PCA is performed on a 3 dimensional grid of data that has batches as one dimension, variables as a second and samples as the third. This must be unfolded into a 2 dimensional grid, in effect "slicing" the cube into layers which contain the data for all the variables for all the batches at a specific time sample.

The results from the multi-way PCA are similar to any other PCA, in that there is a small number of values for each batch. Typically no more than around 6 multi-way principal components are necessary to contain virtually all the information. However, unlike ordinary PCA, this information relates to particular differences in trend curves for a batch, compared to a set of known good batches. These principal components can therefore be plotted in a parallel coordinate visualization together with primary process variables and quality results, thereby giving a highly visual interpretation of the relationship between trend abnormalities and other process values.

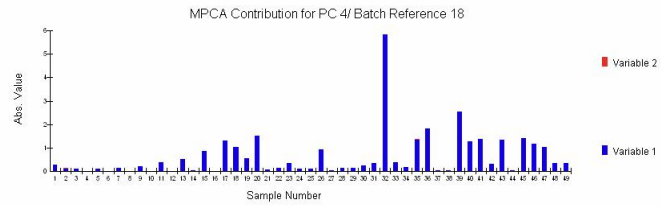
MS2 allows batches to be filtered, by any number of parameters, and analysed relative to good batches. It also enables drill down to the individual batch trend data to highlight those batches which have been filtered.

The "unfolding" process is handled automatically within MS2, which also validates the data to ensure that batches can be compared "like for like" and that no data sets which are impossible to analyse are included.

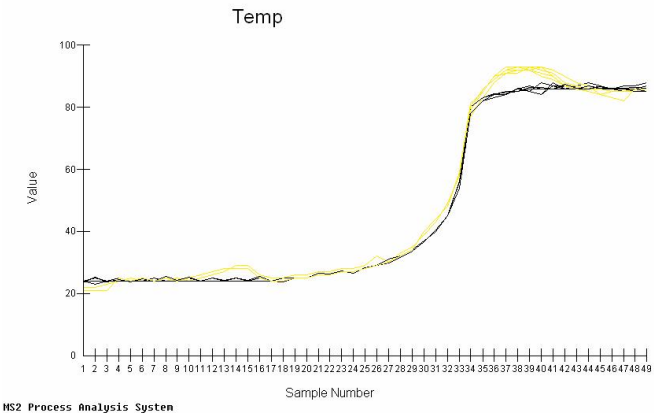
Other functions calculate other multivariable results, such as Q (the square prediction error) which provide an even more powerful visualization of deviating from a normal batch. A comprehensive range of integrated tools enable drill down to the various levels of the calculation, giving a simple and highly visual display.



In this example, multiway principal components for a selection of bad batches are calculated (yellow), against known good batches (blue). Orthogonal plots of the PCs show clear differences.



Drilling down to the contribution plots for individual batches shows the samples (time periods) when the batch is abnormal. This is usually caused by a number of variables (such as pressure and temperature) moving away from the norm, rather than a single cause.



Drilling down further shows the primary data trends for all batches, with the selected batches highlighted. In this example, the temperature curve shows a clearly abnormal exotherm. In conjunction with other variables' deviations, this is the explanation for poor batch quality.

AJM Consulting

Europarc Innovation Centre
Europarc
Grimsby
North East Lincolnshire
DN37 9TT

Phone: 01472 500306
Fax: 01472 500307
E-mail: cathy.oswin@ajm.co.uk
www.ajm.co.uk

Creating Manufacturing Advantage

AJM Consulting is a member of:

- *Institute of Measurement and Control*
- *Humber Chemical Focus*
- *Yorkshire Chemical Focus*
- *NEPIC*
- *CPACT*

Our Clients:

- ABB
- Arthur D Little
- AEA Technology
- Aesica Pharmaceuticals
- A H Marks
- Bemis
- BP
- British Sugar
- Bridon
- British Paper & Board
- Camaxys
- Cambridge Consultants
- Carbon Trust
- ConocoPhillips
- DEFRA
- DICIDA
- DTI
- Eli Lilly
- Fibres Worldwide
- GlaxoSmithKline
- Glinojeck Sugar Refinery
- Grotech Production
- Hickson
- Holliday Pigments
- Humber Chemical Focus
- Huntsman
- ICI
- Impress/HCCTA
- Imerys
- J R Crompton
- Kemfine
- Lakeland Laboratories
- Luxus
- Manro
- Norsk Hydro
- PICME
- PSE
- Roche
- Rotork
- Scott Bader
- Servelec
- Severn Trent Water
- SIRA
- Spiritus Consulting
- Strategem Consulting
- Stepan
- STG
- Sulzer
- Syngenta
- Synthomer
- Technical Absorbents
- Tensachem
- Total
- University of Newcastle
- University of Sheffield
- University of Strathclyde

Frequently Asked Questions

When will development of on-line analysis start and can we be involved?

Soon - and yes. We are in preliminary design and prototypes are planned to be working by June. As with all our previous MS2 developments, we will involve interested companies as soon as possible, and if you're interested, now is the time to talk…

How can we use MS2 to build completely different models?

Simply by creating a blank model and importing the relevant data into it, using any rules you need from the standard MS2 rule set.

Apart from on-line, are other improvements being developed?

There are lots. We plan a completely new algorithm which will be implemented shortly - Partial Least Squares (PLS). This is particularly useful for predicting outcomes of process deviations, and is used for both on-line and off-line analysis.

We are about to release more drill-downs, giving enhanced visibility of problems. Another development will be the ability to create several sets of filtered data simultaneously, for instance to compare the performance of several reactors against each other and against known standards.

What about user documentation for MS2?

It's available, provided electronically as part of the system. Obviously, with a complex tool like MS2, functionality is changing rapidly and this is being reflected in the user guide. Each upgrade, released to users with maintenance contracts, includes an up to date user guide.

Will MS2 work with Microsoft's new Vista operating system?

Of course! As members of Microsoft's Partner Program, we tested it using a development version of Vista months ago. As soon as the final release of Vista occurs we will repeat the test and, after that, all MS2 releases will be fully compatible.

Are there any plans to simplify the upgrade process?

Yes. As we move from development to the finished product, we recognize the need to simplify upgrading. This will be performed by a simple, CD-based process, and the disks will be sent to all maintenance contract users as upgrades are released. This is necessary since many users have firewalls preventing transfer of executable files.

Do you have any plans to extend the data pre-screening rule set?

Yes. Indeed, the rule system, which is fundamental to the vital function of preparing the data for effective analysis, is being upgraded substantially. New rules, both for simple data sets and for multiway batch sets, will be available shortly, and it will also be a much simpler task for us to add to the rule set.