In the next issue

How is drinking water quality measured and maintained? How has drinking water production changed over the last few years and is there anything new to look out for? Does the average UK consumer need to worry about our drinking water? Interested in finding out the answers? Have a look at our next issue in mid-December.

BWC Business Issue 7

Blackwell Water Consultancy Ltd News

- **IChemE Tyneside Member Network**

  The first official event of the new network is due to take place on the 22nd of September 2010 at 6pm at the University of Newcastle upon Tyne. It's a presentation by Neil Foster of CatalySystems Ltd about photocatalysis for wastewater treatment.

- **BWC Analysis launches**

  We're launching a new service, BWC Analysis, which is dedicated to mathematical modelling (particularly empirical modelling) and data analysis. This service can be used in any type of industry and we'll also be launching a new website (www.bwcanalysis.co.uk). The website was launched on 17th September 2010 and will be expanded over the next two months.

Blackwell
Water Consultancy Ltd
PO Box 235
Darlington
County Durham
DL1 9GN
www.blackwellwaterconsultancy.co.uk

ISSUE
06
September
2010

BWC Business

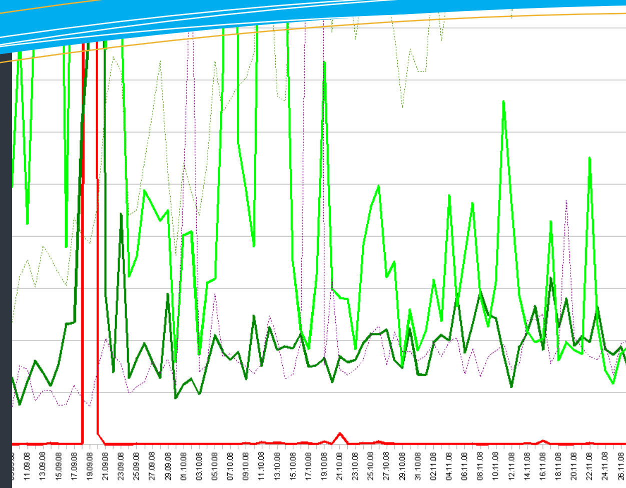
New service launches

Over the last few months we've had an increasing number of enquiries about data analysis and, more specifically, mathematical modelling. We have a considerable amount of experience in both areas and we've decided to launch a separate service dedicated solely to modelling and statistics. We've called the service BWC Analysis and its own website, www.bwcanalysis.co.uk, will be online from the 17th of September 2010.

Who we are

Blackwell Water Consultancy Ltd is based in north-east England but operates throughout the UK. We give advice to businesses in all parts of the UK economy about water supply and effluent treatment. Our sister service, BWC Analysis, offers consultancy about mathematical modelling and data analysis.

Our website has more information about what we do.



In this issue

"88.2% of statistics are made up" *Vic Reeves*

In this issue we were going to discuss the thrilling subject of regulation in the European water industry. We'd been looking forward to digesting some interesting EU legislation and we couldn't wait to give you the exciting highlights. Then we had a better idea.

Ardent readers (hello to both of you) will know BWC as a company that gives advice about the more traditional or conventional aspects of water supply or effluent treatment but we do have another key skill that we've decided to expand.

For many years, and in a number of companies, we've worked in the area of mathematical modelling and data analysis. Now, before you all disappear to find a nice treatise on economic regulation of EU trade effluent charges, hear us out. Mathematical modelling is a vitally important tool in many manufacturing industries. Without it plant design and plant supervision would be largely *ad hoc* exercises. Similarly for data analysis, which is the backbone of quality control in the manufacture of thousands of household products. In this issue, and to coincide with the launch of our new modelling and analysis service ***BWC Analysis***, we'll look briefly at the importance of data analysis and modelling.

Modelling and statistical analysis are subjects that apply to all industries where data can be collected and where quality control is important. This includes areas as diverse as wastewater treatment, the design of satellites and the production of food. But what is mathematical modelling? How do you do it? Why should you analyse data? What sort of benefits can it bring? What is "quality control"? More importantly, why should you consider it?

While the economy improves slowly (depending on which set of figures you read) manufacturing companies are taking greater care to optimise their operations. The only way this can be done meaningfully is to measure what you're doing (e.g. for manufacturers this may be measuring gas, electricity and water use and any other important data). This is where statistical analysis comes in, as it can help to define how well you're doing and how far away you are from any targets you may have.

Modelling can use this data to develop predictions of what can happen in the future. We're not talking about Nostradamus-style forecasts for the next century or two, simply predictions for the next few hours perhaps. But this is far enough ahead to optimise the use of raw materials and so help to reduce costs. A wide-ranging technique that's applicable to many industries. Let's take a look at some of the detail.

Blackwell
Water Consultancy Ltd

The Operator's Mate

Some years ago, we were involved in a project where a mathematical modelling system was referred to as the Operator's Mate. What other useful services can modelling and statistical analysis provide?

Fault detection & diagnosis – Spotting when a fault has occurred and then telling you what caused it. One approach uses a library of data patterns relating to faults and continuously scans the incoming data for these trends.

Soft sensing (see p.3) – acting like a virtual instrument to infer the value of a difficult-to-measure variable using information that is readily available

Statistical process control (SPC) – analysing the variation in your product or operations and using this to tighten control and maintain quality. Six sigma is essentially an SPC technique.

Simulation – using models to simulate operating and design scenarios (e.g. new process designs or control strategies) or for training purposes.

Of course, a model is only as valid as the data that goes into it and great care is needed when deciding if a model is valid or not. Much of the effort expended in modelling goes into making sure the data is as good as it can be.

Measure up and see the benefits

You can't improve what you do unless you measure what you do.

In the last few years a technique known as *six sigma* has become popular in UK industry. You may have come across it or met someone who is a *six sigma black belt*. It's not some form of self-defence but a variation of statistical analysis and statistics-based quality control. It's a powerful technique and we don't have the space to go into the detail of it here but it has a fundamental concept that applies to all industries where data is collected (and not just manufacturing).

That fundamental concept is that if you want to improve what you do (e.g. reduce costs, make a better quality product) then the only way you can do that is to record data about your performance and then analyse it. This defines "where you are" and tells you how far away you are from "where you want to be".

Statistical analysis has many techniques so let's see if we can give you a brief introduction to some of them and where they can be used.

The basics

Basic statistics such as the average, maximum and minimum help to define current performance and we're all familiar with these hopefully. All analysis should start of with these. The maximum and minimum define your range, i.e. the spread of what you do. Is you product always the same weight? No? Then by how much does it vary? How does this affect quality?

The next level

If you have a defined target for quality, such as product weight, the the range defines whether you're using too little raw material and making an inferior product or using too much raw material and wasting it. It's difficult to make each product exactly the same but statistical distributions can help define an acceptable variation in quality. This is the basis of six sigma. If you have the data, distribution analysis helps to define how tight, or otherwise, your quality control is. But what about achieving tighter control? How can statistics help with that?

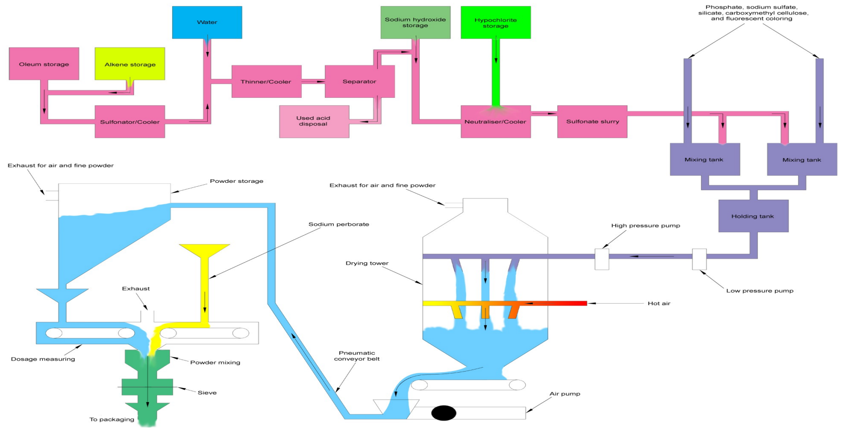
"Statistical analysis is the corner stone of process and quality improvement"

Correlation

In manufacturing processes there are a large number of variables that come together to influence product quality. The statistical relationships between these variables can help you improve your understanding of how your process really works. Correlation analysis is one method of quantifying the relationship between two variables. Understanding that interaction can help to refine process control and so improve product quality. But if you've got a complex process with lots of data, what do you do?

The complex stuff

This is where services such as our ***BWC Analysis*** come in. Techniques such as principal components analysis (PCA) can make sense of large, complex sets of data and they can help to detect and diagnose faults on-line and in real-time and so maintain product quality continually. Coupling this with predictions of future deviations in quality control, via modelling, is a powerful technique.



What do we mean by "modelling"?

Here we look at the broad categories of modelling techniques

Super models and their statistics....................

There are as many classifications of modelling techniques as there are books on he subject but here at BWC we use our own which, hopefully, we can explain in plain English!

We categorise mathematical models as either "first principles" or "empirical"

First principles models are based on the underlying chemistry, physics or biology of a process (i.e. the first principles of how it works). This type of model often includes many equations (possibly hundreds) and needs complex methods to solve them. They are best suited for "simulation", i.e. using the model to run scenarios for new designs or emergency situations.

The accuracy of first principles models varies according to what you model. For some applications the first princples are known very throughly and the models are very accurate. For others, however, the theory may have gaps which restricts the model accuracy.

Empirical models are typically a single equation that links one set of measurements with another. A simple example is a straight line regression. They're very flexible but totally dependent on the quantity and quality of the data they are derived from. Because they don't rely on complex theory then the form of the model can be varied to improve its accuracy. Having said that, it's easy to develop spurious models that seem accurate but which are actually meaningless.

Both types of model have a role to play in process and quality control but for on-line applications, particularly soft sensors (see the panel on the right), empirical models are best.

Q&A – Hose pipe bans and water resources.

Q: This all sounds complicated. Is it?

A: Like walking over hot coals or presenting live TV, mathematical modelling and statistical analysis should only really be done by those with significant experience in the field. There are many pitfalls and it is very, very easy to draw the wrong conclusions from statistics. Similarly with modelling. It's very easy to develop a model that links two seemingly unconnected events and then conclude (wrongly) there's a cause and effect relationship. Leave this one to the experts!

Q: What tools do you need?

A: If you do have some expert knowledge there are many software tools available for this type of work. MATLAB is a very powerful general purpose mathematical package but it can be costly. SciLab is a free, open source alternative. OpenStat is an open-source statistics package that also includes modelling algorithms. There are also many free add-ins for Excel that can help with data modelling (e.g. Essential Regression).

Soft sensors make difficult predictions

You may not be able to buy an instrument to measure some of the variables in your manufacturing process. For example, biochemical oxygen demand (BOD) is a standard measure of sewage strength, yet the lab test for it takes five days to complete. So, you end up with use*ful* historical data for compiling statistics but five day old data is use*less* for feedback control.

Soft sensors, sometimes known as software sensors or virtual sensors, are mathematical models that act as if they were a hardware instrument and "measure" variables for which no instrument exists.

A soft sensor is created off-line and the model relates lab data for the difficult-to-measure variable to readily available data from plant instrumentation (e.g. flows, pressures, temperatures, pH etc). If the correlation is good enough, a model can be developed that estimates the value of the difficult-to-measure variable. This type of model can be surprsingly simple and is ideal for incorporating into a computer control system. When used on-line they can give "instant" measurements of something that may take several days to quantify in a lab.

Blackwell Water Consultancy Ltd