Comments on *treNz* paper:
Biotechnology for New Zealand engineers: A review.

Author	Date
Jackson N & Medich L (2003)	1 July 2003
Max Kennedy	3 September 2003

Virtual Analysers in the Biotechnology Industry ***Jackson N & Medich L (2003)***
1 July 2003

Summary

The optimisation of product quality and plant efficiencies and the control of waste to meet environmental regulations, are vital production management objectives. However, for many biological processes it is impossible to use hardware to measure key process parameters online, and without accurate and timely online testing it is difficult to achieve these objectives.
A solution is to use virtual analysers, which use mathematical models to infer the values of critical parameters on the basis of other measurable process variables. Successful uses in biotechnology include predicting moisture in milk powder and biochemical oxygen demand (BOD) in wastewater, and using core neural network technology to identify patterns in DNA and RNA sequences.

Traditional Instruments have Limitations

For many biological processes, it is impossible to use hardware to measure key process parameters online. Often this is because instruments cannot withstand the harsh conditions present in a biological environment. Temperature and pressure extremes and chemical exposure can make the cost of instruments excessive, and in any case air bubbles, high solids concentrations and sanitary requirements can prohibit effective in-line and on-line measurement. In addition, a complex and/or lengthy chemical or biological assay is sometimes required, for which no replacement on-line sensor is available. Sometimes the parameters (such as sensory attributes) are intangible, and so cannot be measured using traditional instrumentation; and occasionally the key parameter is to too dangerous or valuable to warrant regular testing.

Historically measuring such key parameters involved estimating or making a laborious and delayed off-line measurement of the variable that had to be controlled.

Not surprisingly, the lack of accurate and timely on-line testing creates problems in monitoring and optimising product quality and plant efficiencies, reducing waste and meeting environmental regulations.

Virtual Analysers Provide Accurate and Timely Results

Virtual analysers do not have the limitations of traditional hardware instrumentation. The approach is to develop a model from the measured parameters that affect the process variable, apply experience and rules if necessary, and deploy the model in real time, so that subsequent physical measurements can be reduced or eliminated.
Virtual analysers can be used to infer the value of parameters that are otherwise difficult to measure. Other benefits include lower installation and maintenance costs, and greater reliability and accuracy. Virtual analysers can be fed directly into closed-loop control applications.

Successful in the Biotechnology Industry

The ability of virtual analysers to automatically predict process properties in real time means that they can be used in the biotechnology industry.

For example, when manufacturing large quantities of milk powder, it is important to measure and control the amount of moisture in the product, as changes as small as 0.05% can have significant financial effects. Without a reliable method of measuring moisture online, virtual analysers use online measurements such as the drying air temperature and humidity and the drier feed rate to predict the moisture content of the final product. The result is used to decide automatically whether any changes should be made to the process to achieve the moisture target. This approach is currently used by *Fonterra* in their strategic alliance with Austin, Texas-based *Pavilion Technologies, Inc*.

Another example is the prediction of BOD in wastewater, where the predicted BOD values are available within a few hours instead of the five days required for an accurate BOD test. The prediction allows a more effective response to process changes in the waste treatment plant, preventing environmental damage and the potential for exceeding the upper limit on a discharge permit. Some of the inputs to the virtual analyser are chemical oxygen demand, suspended solids, sludge bed depth and effluent colour. *Bowater Inc*, Coated Paper Division, Catawba, North Carolina, developed this approach in 1998 in conjunction with *Pavilion Technologies, Inc*.

DNA and RNA Sequence Analysis
The predictive properties of virtual analysers can also be used for DNA and RNA sequence analysis. Given large data-sets, virtual analysers can identify a pattern or structure and predict outcomes.
Therefore they can be used to identify patterns (features) from RNA and DNA sequences associated with certain diseases and prognoses, or splice junctions or ribosome binding sites. They can be useful in disease diagnosis and prognosis, medical research and pharmaceutical development. *Pacific Edge Biotechnology Ltd*, Dunedin, is developing proprietary neural network-based software to predict disease risk and outcomes.

Steps in setting up a virtual analyser
Virtual analysers can be built using neural networks technology. Neural networks are especially useful when fundamental first-order principals cannot be easily used to model the process and when a large amount of data is available.

Hamilton-based *PavTech NZ, Ltd,* A *Pavilion Technologies'* company*,* offers a neural network solution in the form of their optimisation software package *Process Insights*®.

The tool allows the user to easily build and deploy virtual analysers according to the recommended methodology below:

1. Identify key variables for prediction: which product properties need to be predicted?
2. Identify key inputs: which parameters affect which key property?
3. Collect historical data: assemble data on key inputs.
4. Pre-process data: delete data outliers, filter data, remove correlated inputs, and prepare data for modelling
5. Build a neural network model: identify the connectivity between inputs and outputs; specify proper time delays; specify portions of data to be used for training; and train the model until training and test errors are about equal.
6. Review model results: compare the results of the model with laboratory data.
7. Make sensitivity analyses: ensure results agree with process experience; delete inputs that are found to be insignificant; re-train the model accordingly.
8. Make gain analyses: check gains for each input to see that their signs and magnitudes make sense; adjust and retrain the model.

9. Validate the model: check the model against a different data set to see that it has captured the essential features of the process and is not over-trained.
10. Improve the model: if the model is not acceptable, improve it by using data for longer time periods, use higher-frequency data, add inputs and identify other measurements.
11. Put the model online: check to be sure correct information is being fed to it; watch out for process modifications.

Conclusion

As process control and optimisation become increasingly important, especially in the high-value biotechnology industry, virtual analysers can be an accurate and timely alternative to instrument hardware. Their predictive properties can also be used for gene sequencing and medical research.

References

Kasabov, Prof Nikola. New Zealand Summer School of Bioinformatics, Knowledge Engineering and Discovery Research Institute, Auckland University of Technology, Lecture Notes, 2003.

Masmoudi, Radwan A. Rapid Prediction of Effluent Biochemical Oxygen Demand for Improves Environmental Control, *Tappi Journal*, Vol 82, No 10, October 1999, p111-119.

Merritt, Rich. Hardware vs Virtual Analyzers, *Control*, February 2002.

Max Kennedy
3 September 2003

Issues of data analysis and data mining are crucial to biotech and have has gained increasing prominence in diverse areas, such as process development and bioinformatics, in recent years. "Intelligent techniques" such as fuzzy logic, neural networks and genetic algorithms are valuable tools to build on the traditional approach of statistical data analysis.

They will have a growing impact as more people become aware of their power and utility, and develop applications.

Max Kennedy
BioPharmaceutical Division, Industrial Research Ltd
3 September 2003