Please cite as Environmental Modelling & Software 22 (2007) 315-322

EXPERIENCE WITH DATA MINING FOR THE ANAEROBIC WASTEWATER TREATMENT PROCESS

M. DIXON[*], J.R. GALLOP[#], S.C. LAMBERT[#], J V HEALY[*]

[*]*Department of Computing, Communications Technology and Mathematics*
London Metropolitan University
31 Jewry Street
LONDON EC3N 2EY, UK

[#]*Business and Information Technology Department,*
CCLRC Rutherford Appleton Laboratory,
Chilton, Didcot, Oxon. OX11 0QX, UK

Abstract

Anaerobic digestion provides an effective way of disposing of organic material in wastewater. The EU-funded TELEMAC project aims at improving the reliability and efficiency of monitoring and control of this type of wastewater treatment plant. One of its special features is the idea of a telecontrol centre which monitors multiple, geographically distributed plants remotely, acts as a centre of expertise, and brings together the expertise of a network of remote experts. Data mining has been identified as a potentially useful contributing technology. Sensor data is now becoming available for some pilot, laboratory scale, and industrial sized digesters.

This paper presents the directions of work and emerging results of data mining. Particular themes considered here include:

- experience gained in the data mining exercise;
- the use of confidence and prediction intervals;
- prospects for generalisation over different sizes and types of anaerobic digester;
- relationship to the overall supervision system developed in the project.

Keywords: data mining, anaerobic waste water treatment, telemonitoring and control

1 INTRODUCTION TO ANAEROBIC DIGESTERS

Anaerobic digestion is a high-yield process for treating organic pollutants. Essentially it involves two processes taking place together in a vessel called the digester: acidogenesis, in which the organic carbon pollutant is converted to volatile fatty acids (VFAs); and methanogenesis, in which the VFAs are turned into methane and carbon dioxide (see Figure 1). Both processes are performed by bacterial populations co-existing in the digester itself.

[*] E-mails: M.Dixon@londonmet.ac.uk, J.R.Gallop@rl.ac.uk, S.C.Lambert@rl.ac.uk, J.Healy@londonmet.ac.uk

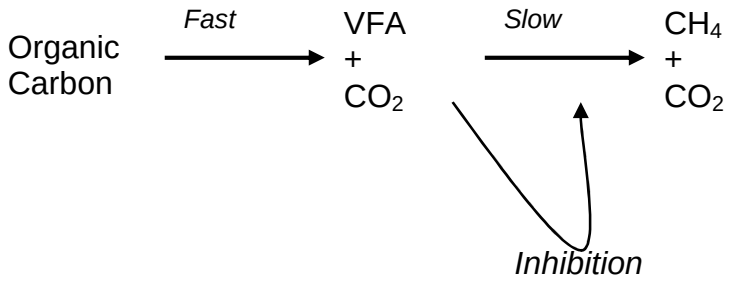

Figure 1: The anaerobic digestion process

Anaerobic digestion can be used for treating waste from the production of alcoholic beverages. This is an application of particular environmental sensitivity: a medium-sized winery generates pollution equivalent to a town of 15 000 inhabitants, and a medium-sized tequila producer in Mexico is equivalent to a city of 125 000 inhabitants. Anaerobic digestion has a number of advantages over other treatment processes. However, it is an unstable process. It is necessary to maintain the balance of the two bacterial populations and of the acidity and reaction rates. If the process is destabilised, it can lead to bacterial biomass elimination and consequent shut-down of the plant. In order to avoid this, plants are often run at low efficiency for greater stability. The ideal situation would therefore be to be able to run a plant at higher efficiency, while maintaining stability.

2 OVERVIEW OF THE TELEMAC PROJECT

The EU-funded TELEMAC project brings together 15 organisations across Europe and Latin America to improve the monitoring and control of anaerobic wastewater treatment plants for the alcoholic beverage industry. Its aim can be summarised as the development of a modular and reliable system supporting remote monitoring and control of small treatment plants with no local expertise. The innovations of the project are at several different levels, including remote monitoring through a Telecontrol Centre (TCC), which manages several plants through Internet connections, taking advantage of its own advanced control facilities and the possibility of a network of remote experts.

Another unifying aspect of the TELEMAC approach is the leveraging of knowledge. Knowledge may be extended from one source into other areas, thereby helping to raise the overall standard of management of the individual plants. It is here that data mining has promise, through allowing knowledge to be derived from past data either on a single plant or on multiple plants.

3 CHARACTERISTICS OF TELEMAC AS A DATA MINING PROBLEM

Supervising anaerobic wastewater treatment plants in general and in the TELEMAC project in particular raises problems that are promising for data mining. The TELEMAC system can be seen as an instance of an environmental decision support system, as

surveyed by Poch et al.[1], within whose general architecture TELEMAC may be positioned. Data mining fits within what they call 'Reasoning/models integration'.

Five TELEMAC project partners: (Laboratoire de Biotechnologie de l'Environment – INRA, France; University of Santiago de Compostela, Spain; Agralco S. Coop., Spain; Sauza, Mexico; National Agency for New Technologies, Energy and the Environment (ENEA), Italy) conduct experiments within the scope of this project. Although the project enables them to co-ordinate the conduct of experiments, they are independent institutions, some of whose treatment plants are in full scale industrial use. Thus several differences between digester plants can be identified.

Three methods of digestion, including continuously stirred and sludge blanket, are represented on the project partner sites and statements that apply to all methods are likely to be few. Even for those treatment plants which operate on the same design as each other, there are major differences of scale. One plant has a volume of 2 litres where another is 5 million litres.

We would expect the sequence of events to be important, since the state of a plant at one time would affect its future state. However the precise time dependence is not the same at all treatment plants. The hydraulic retention time, HRT, has been identified as one of the applicable scalings. HRT is the reciprocal of the dilution rate and is measured as the volume/inflow rate.

Instrumentation varies between digester plants, with different variables being measured for different digesters. Some variables, such as the standard process variables of temperature, flow rates, pressure, and acidity are cheap and easy to measure on line. Some key variables are hard or expensive to measure. Some variables may be measured manually offline on one plant and by automatic sensors on another. One partner, INRA, has a research digester which is equipped with an advanced collection of sensors and is capable of making multiple measurements of some variables[2]. There is therefore much interest in identifying the minimum instrumentation necessary for the stable running of the plant. Sensors fail, provide questionable values or are taken offline for cleaning, so another interest is the identification of models that allow for predictions from reduced instrumentation.

Although the differences between sites and plants make the data mining problems harder, they do offer the possibility of systematic accumulation of expertise giving rise to conclusions that may be more representative of the general population of anaerobic waste water treatment plants.

Recently, Hamed et al.[3] reviewed a range of applications of Neural Net models to environmental and water resource engineering. For their own work on wastewater treatment plants they concluded neural nets provided an efficient and a robust tool in predicting performance.

In broad terms, the questions to be answered by the data mining work in TELEMAC include the following:

- What conditions can be used to model abnormal and undesirable states in a particular digester plant? Or among plants operating under the same principle?

- Which sensors are necessary for modelling a plant to a sufficient degree of confidence?
- What statements can be made about multiple plants of different types?

4 DATA MINING

4.1 Overview

Data mining as a complete activity is recognised as an iterative process in which the understanding derived from the reports of the data mining models is fed back into data cleaning, data enrichment, data selection, and data transformation. We show in Figure 2 a schematic outline for the data mining process for the TELEMAC project. Specific data mining modelling, validation and hypothesis testing only appear in the fifth of six stages of the diagram Before data mining, it is frequently necessary to apply data pre-processing; the architecture of the GESCONDA [4] is a good example of how data filtering, knowledge discovery, knowledge management, and meta-knowledge management can be structured in a software tool for environmental databases. We give some examples of pre-processing in TELEMAC:

- In this project, there is a particular need for pre-processing on account of data being produced at several sites. Although a consistent approach to variable naming was planned and carried out, some pre-processing was still necessary. A relational database was produced to hold all the metadata, data about data, which includes for each experiment, start time and names of measured variables, which could run to 20 at one of the industrial sites and many more at the experimental plant at INRA. Descriptive information about each variable was included. Potentially the proliferation of variables gives rise to a high dimensional problem. It is necessary to find a way of reducing the set of variables to those which are significant and recorded in sufficient numbers. The relational database is proving useful in answering such questions as, given a digester method that is represented at more than one site, which variables for which experiment are reported from all those sites.
- Different software tools require different conventions for missing values and date representation.
- Some algorithms require discrete variables, where others require floating point variables.

Another important issue is the role of visual techniques. Although they are essential at the reporting phase of the iterative cycle, they are useful at the planning stage too, when deciding on which variables to focus in a given hypothesis test. For this purpose, the XMDV multidimensional tool is proving useful[5].

We now turn to the data mining algorithms themselves. The work began with neural net techniques, to enable assessment of consistency and confidence in the models of the data. Results from this work are described fully. Later on, rule induction and clustering are introduced. Rule induction has the potential of extracting rules which are easily understood. Clustering has the potential to identify sets of similar data. An introduction

to the work on clustering and rule induction is provided. In all these cases, the Clementine data mining software is being used[6].

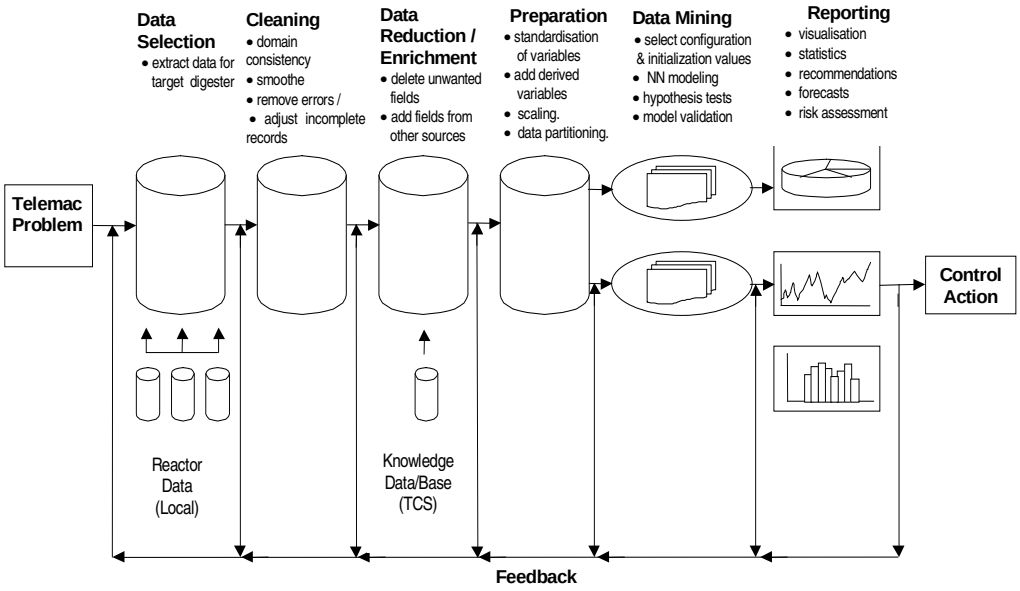

Figure 2: The data mining process

4.2 Neural nets as multivariate regression

We use neural nets to perform multivariate regression among the time sequenced data reported by the TELEMAC plant sensors.

4.2.1 Modelling issues in neural nets

Tibshirani[7] and Weigend with co-workers[8, 9] have addressed some modelling issues that need to be considered in the deployment of neural nets for regression. The issues are particularly pertinent to the TELEMAC data at this stage because of the relatively small quantity of data available. With ordinary linear regression the least squares cost function is convex and there is a well defined global minimum. For a neural net regression the surface of the cost function is unlikely to be convex so there are likely to be many local minima. Neural nets are capable of over fitting the training data with the result that a model specific to the training data points is produce but it would lack applicability to other data. A special technique called *early stopping*[10] was used in this work. The available data used in the modelling is split randomly into three data sets called training, validation, and test sets. A training set is used to fit the parameters of the model. In order to avoid over fitting the training data, the separate validation data set is used to terminate training; the minimum of the cost function is determined by the minimum on the validation set. The test set is then used to assess the quality of the fit.

Conventionally the quality of the fit to the target data set has been assessed by relatively insensitive aggregate measures such as the mean-squared-error or the square of the linear correlation known as the coefficient of determination, R^2; such measures are not

sensitive to variations in fit across the space defined by the range of values for each variable. The variance of the error term over that space, means that a wild bootstrap is required in any bootstrap approach. For confidence in predictions from regression, point estimates of the variance are desirable. Recently we have presented a robust method for estimating simultaneously the conditional mean and point estimate of variance of a target variable[11]. On a synthetic data problem defined in the literature[9] we were able to show the method gave unbiased estimates at the 5% significance level of the estimated mean and of the estimated point variance. Full details are available[12]. The heteroskedasticity consistent estimate of the prediction interval is given by

$$P \approx d^*(\mathrm{x}_i) \pm t_{(1-\alpha/2),(n-k-1)} \sqrt{\frac{n\sigma^{*2}(\mathrm{x}_i)}{(n-k-1)}} .$$

In this case, *k* is the applicable degrees of freedom, n is the total number of records used to train the neural net, *t* is the Student's t-distribution with α as the significance level. $d^*(\mathbf{x_i})$ is the estimated conditional mean and $\sigma^{*2}(\mathbf{x_i})$ is the estimated conditional variance for input tuple $\mathbf{x_i}$.

4.2.2 Example of modelling prediction intervals for TELEMAC data

Bernard et al.[13] have developed several mathematical models of the processes in an anaerobic digester plant. They calibrated the model against experimental data and then produced a dataset from a simulation of some of the plant's reactants and products over a 70 day period. We chose the simulation model as the starting point for our investigation. The data was reported in time/value pairs. A consistent data set was extracted from the simulation by aligning the time stamps and omitting incomplete records. There is particular interest in whether it is feasible to use a small set of sensor readings to produce a satisfactory neural net model of the process. We therefore choose to model one output variable, the concentration of volatile fatty acids in the digester, VFA_dig, from the input variables of the input flow rate (but not VFA concentration of influent) Q_{in}, the acidity of the reactor pH_dig, output gas flow rate Qgas, the flow rate of methane QCH_4, and the flow rate of carbon dioxide QCO_2. As with some of the experimental TELEMAC data, temperature was not reported for this simulation.

Following work of Healy et al.[11] we split the 849 data records into three data sets of approximately equal size which we label A, B, and C. Set C was designated the test set. Set A was then randomly split into 50% training and 50% validation records to generate a neural net model for the prediction of VFA_dig from { Qgas, $Q_{in,}$. pH_dig, $QCH_{4,}$ QCO_2}; we did not specifically input time as a variable. The neural net model from Set A was then applied to Set B to derive an estimate of VFA_dig together with a set of residuals. We then appended the estimates of the squared residuals to Set B. These squared residuals form the basis of the error estimate that we use in deriving the confidence intervals. An error model is needed to get confidence intervals for predictions for which the target is unknown. Set B was then randomly split 50% / 50% into a training set and a validation set. A neural net model was then fitted to both VFA_dig and the squared residuals for Set B; this model provided a simultaneous point estimate of the error and the VFA_dig concentration.

We report in Figure 3 the results of applying the model to the test set C. Student's t-tests for 5% significance (equal variance two sample paired and independent tests) were applied to the simulation and predicted values from test set C; these confirmed that values consistent with an unbiased estimate of the mean value of VFA_dig had been derived. Figure 3 shows that the estimated VFA_dig which is labelled NN_VFA_dig and shown as a solid line predicts very closely the VFA_dig data points shown as solid diamonds. This indicates that the neural net was able to predict well from the sensor variables the value of the target VFA_dig. The upper and lower prediction intervals, UPI and LPI, are shown as broken lines. The major spreads occur where there are large changes in the VFA concentration; however the alignment remains good. Calculation of the moving 20 record average of the standard deviation for VFA_dig indicates that these major spreads correspond to more volatile readings. Of course a concentration can never be negative; the error is occurring on the most steeply descending peak and appears to be due to a small modeling lag. The wider spread for data between 41 days and 44 days appears to reflect a increase in acidity or a reduced CO_2 flow rate rather than a change in $Q_{in.}$

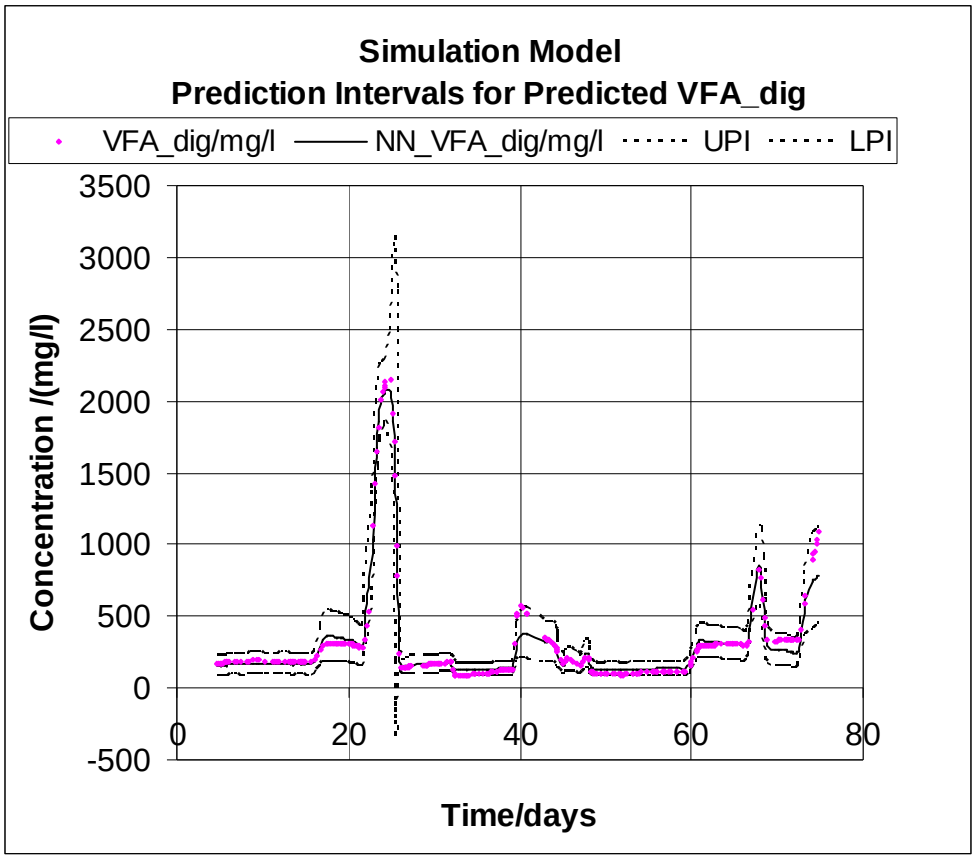

Figure 3: Results of applying neural net model of VFA_dig

Our previous investigations for this data had shown that significantly different results could be obtained from different neural modelling techniques with different target outputs so that a pruning approach was deemed advisable.

The accuracy of the overall fit to Set A data for the single variable VFA_dig was

99.42%. The most powerful of our neural net modelling tools, the exhaustive prune model, was deployed. With the exhaustive prune model the hidden node architecture and the number input nodes are varied. A very exhaustive search for the minimum of the error surface is carried out. The node architecture was determined as [5 input nodes { Qgas, $Q_{in,}$ pH_dig, $QCH_{4,}$ QCO_2}: 5 hidden_layer_1 nodes: 5 hidden_layer_2 nodes: 1 output node {VFA_dig}]. The accuracy was much higher than we found for some of the experimental reactor data which had accuracies of about 94% for predictions of VFA_dig.

The accuracy of the overall fit to Set B data for the single variable VFA_dig was 98.8%. The node architecture was determined as [5 input nodes {Qgas, Q_{in}, pH_dig, $QCH_{4,}$ QCO_2}:8 hidden_layer_1 nodes: 6 hidden_layer_2 nodes: 2 output nodes {VFA_dig, residual**2 }].

For the test set C the value of the correlation coefficient for NN_VFA_dig with VFA_dig was 0.988. For Q_{in}, pH_dig, QCH_4 , QCO_2 , it was 0.81, -0.31, 0.87, 0.9 respectively.

4.3 Rule induction

Some preliminary work on rule induction was carried out on an INRA data set. In using the algorithms, the aim is to generate rules that characterise the data to some degree of confidence. These algorithms are supervised, which means that (like neural net regression but unlike clustering) the user supervises the algorithm by nominating one or more of the variables as a target.

In this initial work, VFA_dig was chosen as a target, as it is desirable to avoid the build up of VFAs in the digester. This parallels the work with neural nets just described.

The rule induction algorithms used (unlike clustering and neural net regression) require that the target be a symbolic (or categorical) variable. Therefore the target and all the input variables were allocated to bins. There are several ways of achieving this, including linear transformations or an equalisation transformation based on equal width or equally populated bins. In this initial investigation, a linear transformation based on standard deviations was used, which had the result of transforming the variables to a small integer range.

An example of a rule that has been generated with this data is the following:

> *VFA_dig falls in a particular range of high values, if QCO_2 falls within a particular range of low values – this rule scores (431, 0.596)*

This can be interpreted as follows:

> *Assuming that we confine our examination to records with valid values, there are 431 records—the support—for which QCO_2 satisfies the condition. Of these, in 59.6% (257) of the records—the confidence—the conclusion concerning the target variable VFA_dig is satisfied.*

This rule has the advantage of being simple—only one input variable—and the number of candidate records is acceptably large. However 59.6% could be judged to be rather

low. One direction of investigation would be to determine how adding further variables to the condition would increase the confidence.

Conventionally a rule induction algorithm allows the thresholds of confidence and support to be set. However if the confidence in the target condition is also high for situations where the input condition is *not* satisfied, it is clear that the rule is not interesting. Therefore the evaluation of a rule needs to assess whether the input condition increases the likelihood of the conclusion being true—the lift.

In an application such as this one, which involves measurements on diverse plants, missing values are to be expected and the rule induction algorithms in Clementine make good use of as many records as possible. Thus if a rule involves a target variable t and input variables a, b and c, all records for which t, a, b and c are valid are used even if other variables not participating in the rule have missing values.

4.4 Using visual analysis to support rule induction

Visual analysis makes use of the human capacity to perceive patterns. It can be used in both preparation for and, as shown here, subsequent evaluation of data mining. Suppose that the following rule is generated:

> *VfaDig3_bin is close to 72.011374 if codDig3 <= 322.60001 and pressDig <= 15.366, the rule has support 156 and confidence of 79%.*

with the ranges of these variables as follows

```
vfaDig3: 3.87 - 8033.50
codDig3: 32.00 - 15214.33
pressDig: 1.20 - 46.35
```

Figure 4 shows these four variables together with ElapsedTime in a multi-way scatterplot using the visualization tool XmdvTool[5]. It consists of an array of individual scatterplots for every pair of variables. The scatterplots occupying the diagonal result from each variable plotted against itself and there is symmetry about this diagonal.

The variation of colour within each plot results from an interactive technique known as brushing. The figure illustrates this but in grey scale. This is a useful way of inspecting a portion of the n-dimensional space. First define an n-dimensional hypercube such that the range of each dimension corresponds to the range of values present for one of the variables. For each variable, we then choose a subrange. The pale grey area in each diagonally placed scatterplot portrays the subrange for the corresponding variable, is governed by interactive control and this is the process referred to as brushing. The result is, in general, a smaller n-dimensional hypercube. The n-dimensional points falling within the subset hypercube are shown in black and the remainder in a mid tone grey.

The figure illustrates the induced rule. The target variable is shown at the extreme right and bottom. Although the ElapsedTime was not used in the rule induction, the trend of each variable is often of relevance to the interpretation. The three input variables are also shown and the subranges of those variables - shown as the pale grey area - correspond to the conditions in the induced rule. Although the figure is approximate

because the subranges were selected by mouse, complete accuracy is not required for an overall view.

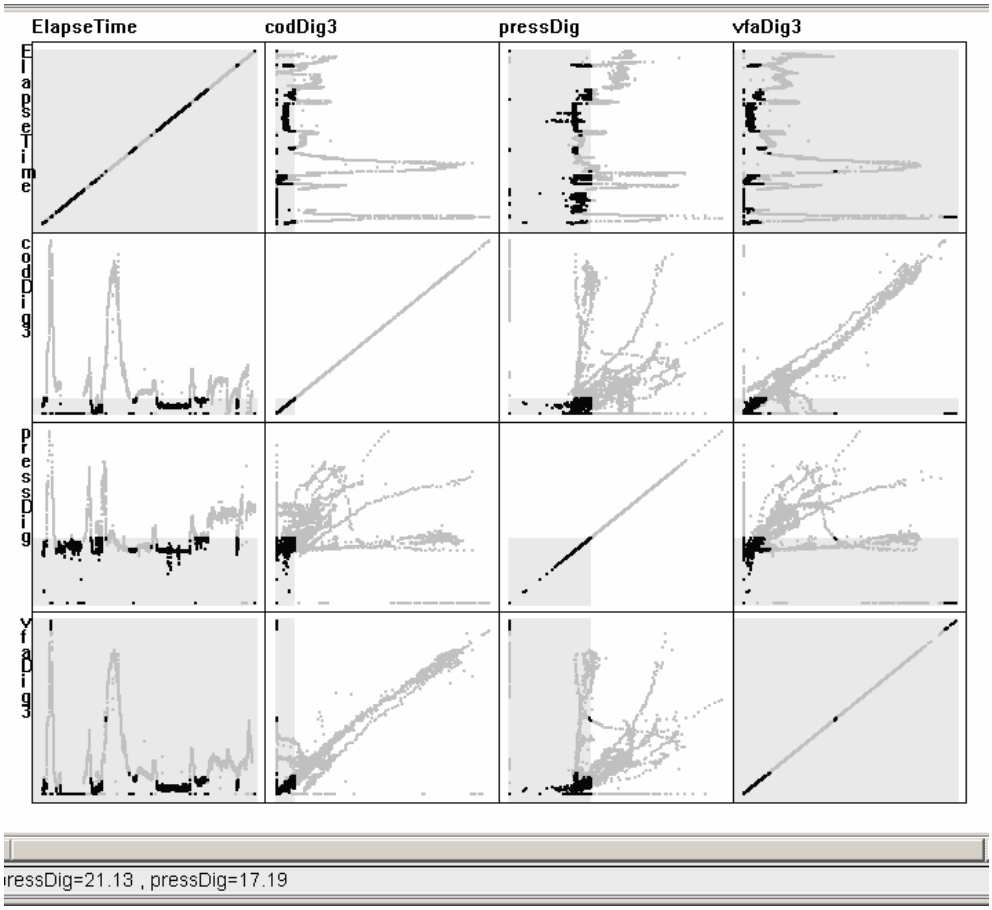

Figure 4: Using a multidimensional visualization tool to analyse rule induction

5 INCORPORATION OF KNOWLEDGE INTO THE SUPERVISION PROCESS

The TELEMAC project aims to produce a modular system as its end result. Data mining will provide one of the modules, and it is necessary to specify how it will interact with other components of the system. At present, this issue is still under investigation, but one way of approaching it is through exploring several key questions about the role of data mining.

Data are being accumulated slowly from the TELEMAC industrial sized digester sites, typically one record per day. The main sources of data at this stage are from laboratory scale digesters where the volume of the digester is several orders of magnitude less. This has direct implications for the way data mining can be introduced to assist the Telecontrol Centre in providing expertise. We have only a small coverage of the possible range space available for each for each digester. In particular industrialists are

reluctant to experiment with pushing the operating conditions into configurations that would lead to irreversible failure and consequent shut down. We have an initial indication that even for an individual digester the model constructed from one year's data is only moderately successful in predicting the volatile fatty acid content in another year when there is a restart between. It is therefore necessary that the data mining modules be based on a customisable template that is validated and enhanced over time.

Initially it will be necessary to set up a direct correspondence between an individual digester and its data mining models. It is intended that the predictions from the data mining are in the form of indicative supplementary guidance to the local controllers; the data mining output should not automatically affect the digester control without human intervention. However it is feasible to load applications containing data mining models onto the local sites at the behest of the digester operators.

The time scale for model revisions will be affected by two factors: 1) a short timescale factor arising from the extent to which the plant is operating with data beyond the data range for which the models were constructed 2) a long timescale factor arising from the steady accumulation of more data. There is some evidence on a laboratory scale that the operating characteristics of an individual digester can change over time. This is an area that will require further work in the course of the TELEMAC project.

At the time of writing, it seems that the likeliest outcome will be that data mining will be used to generate rules or models from data at the TCC that are then downloaded to the local system on the plant for execution. They will alert the local operator in the case of detecting perturbations to the system's normal operation. Periodic re-mining will almost certainly be required, due for example to changes in the operating regime, or long-term changes in the properties of the plant. The conditions under which re-mining should be performed are a subject of further work.

6 FURTHER WORK

Further data mining work within the project is planned.

Extraction of significant patterns, with their associated levels of support and confidence, will be continued and completed, using methods described in this paper.

For neural net regression in particular, we will construct a set of regression models that will allow for the systematic absence of modelling variables in order to maintain predictions during sensor non-availability. We have shown that good quality fits to VFA_dig can be more easily achieved if other variables such as partial alkalinity, and bicarbonate concentrations are fitted. Concerning variability, Le Baron and Weigend[8] have studied the effects of dataset splitting on conclusions from neural net models for time series. They warn that significant variability can be found. We intend exploring the consequences of this variability for the TELEMAC models we are constructing. Part of this extension should include an assessment of the range of suitability of linear regression models. We will also extend confidence limits work to combinations of real data since the use of summary variables are known to be inadequate in assessing modelling quality. For time prediction we will follow initially the approach suggested by Dorffner[14]

The resulting patterns will require evaluation by the project's treatment plant scientists. Results of the evaluations will then be encapsulated ready for incorporation into the supervision process.

7 CONCLUSIONS

Overall partners in the TELEMAC project have established that anaerobic digestion is an important technology to which systems modelling can contribute.

As a useful quantity of data from the industrial digesters becomes available, data mining support becomes feasible for individual digesters. Further work is being carried out to determine wider applicability.

We have demonstrated the effectiveness of a new robust method of determining confidence in predictions from non-linear regression.

8 ACKNOWLEDGEMENTS

We wish to acknowledge the support from the European Commission's IST programme under the TELEMAC project (IST-2000-28156) and useful discussions with project partners, particularly Jean-Philippe Steyer and Laurent Lardon of INRA. We also wish to acknowledge useful discussions with Brian Read and Fang Fang Cai.

9 REFERENCES

1. M. Poch, J. Comas, I. Rodríguez-Roda, M. Sànchez-Marrè, and U. Cortés (2004) Designing and building real environmental decision support systems, *Environmental Modelling and Software*, **19**(9), 857–873.
2. J.-P. Steyer, J. C. Bouvier, T. Conte, P. Gras, and P Sousbie (2002) Evaluation of a four year experience with a fully instrumented anaerobic digestion process, *Water Science Technology*, **45**(4-5), 495-502.
3. M.M. Hamed, M. G. Khalafallah, and E. A. Hassanien, E.A. (2004) Prediction of wastewater treatment plant performance using artificial neural networks, *Environmental Modelling and Software* **19**(10), 919-928.
4. K. Gibert, M. Sànchez-Marrè and I. Rodríguez-Roda (2005) GESCONDA: An intelligent data analysis system for knowledge discovery and management in environmental databases, *Environmental Modelling and Software* (In Press – available online 17 March 2005).
5. M. O. Ward (1994) XmdvTool: integrating multiple methods for visualizing multivariate data, Proceedings IEEE Visualization 1994, 326-333.
6. SPSS Inc (2003), *Clementine User Manual*.
7. R. Tibshirani (1996) A comparison of error estimates for neural net models, *Neural Computations* **8** 152-163.
8. B. Le Baron and A. S. Weigend (1998) A bootstrap evaluation of the effect of data splitting on financial time series, *IEEE Transactions on Neural Networks* 213-220.
9. D. A. Nix and A. S. Weigend (1995), Learning local error bars for non-linear regression, *Proceedings of NIPS 7*, 489-496.
10. 'What is early stopping' [www http://www.faqs.org/faqs/ai-faq/neural-nets/part3/section-5.html]

11. J. V. Healy, M. Dixon, B. J. Read, F. F. Cai (2003) Confidence in data mining predictions: a financial engineering application, *Proceedings of the 29th Conference of the IEEE Industrial Electronics Society IECON'03 Conference* 1926-1931.
12. J. V. Healy, M. Dixon, B. J. Read, F. F. Cai (2003) Confidence and prediction in generalised non linear models: an application to option pricing, International Capital Markets Discussion paper 03-6 p1-42.
13. O. Bernard, Z. Hadj-Sadok, D. Dochain, A. Genovesi, J.-P. Steyer (2001) Dynamical model development and parameter identification for an anaerobic wastewater treatment process, *Biotechnology and Bioengineering* **75**(4) 424-438.
14. G. Dorffner (1996) Neural networks for time series processing, *Neural Network World* **6**(4), 447-468.