

Selecting the Most Relevant Variables for Anaerobic Digestion Imbalances: Two Case Studies

Jordi Dalmau¹, Joaquim Comas¹, Ignasi Rodríguez-Roda^{1*}, Eric Latrille², Jean-Philippe Steyer²

ABSTRACT: In this study, a wrapper approach was applied to objectively select the most important variables related to two different anaerobic digestion imbalances, acidogenic states and foaming. This feature selection method, implemented in artificial neural networks (ANN), was performed using input and output data from a fully instrumented pilot plant (1 m³ upflow fixed bed digester). Results for acidogenic states showed that pH, volatile fatty acids, and inflow rate were the most relevant variables. Results for foaming showed that inflow rate and total organic carbon were among the relevant variables, both of which were related to the feed loading of the digester. Because there is not a complete agreement on the causes of foaming, these results highlight the role of digester feeding patterns in the development of foaming. *Water Environ. Res.*, **82**, 492 (2010).

KEYWORDS: anaerobic digestion, foaming, acidogenic states, feature selection, neural networks.

doi:10.2175/106143009X12529484815359

Introduction

Anaerobic digestion is a complex process that can be affected by several imbalances. Among them, foaming is one of the most frequent operational problems (Svensson et al., 2007). Foaming can clog gas pipes, entrap solids, cause stirring difficulties, and lead to failures in monitoring devices. There is not complete agreement on the parameters that favor conditions for foaming forming bacteria. Some authors state that proper control of feeding will prevent excessive foaming (Massart et al., 2006; Schafer et al., 2006). Others state that pretreatment of the feed is necessary to avoid foaming (Barjenbruch and Kopplow, 2003; Elliott and Mahmood, 2007). Some claim that the presence of filamentous bacteria in the activated sludge system is the main cause of foaming in the anaerobic digester during activated sludge treatment (Pagilla et al., 1997; Westlund et al., 1998). In parallel, many of the anaerobic digestion disturbances end up in acidogenic states. These entail high volatile fatty acid (VFA) concentrations associated with a decrease in pH and methanogenic inhibition, with the related decrease in methane production (Guiot, 1991).

¹ Laboratory of Chemical and Environmental Engineering (LEQUIA), University of Girona, Girona, Catalonia, Spain.

² INRA, Laboratoire de Biotechnologie de l'Environnement, Narbonne, France.

* ICRA, Catalan Institute of Water Research, Parc Científic i Tecnològic de la Universitat de Girona. Edifici H2O Emili Grahit 101, 17003 Girona, Catalonia, Spain; e-mail: irodriguezroda@icra.cat.

For example, in Carrasco et al. (2004), low hydraulic retention times (HRT) and organic overloads cause acidogenic states. Moreover, toxicants can cause VFAs to accumulate via methanogenic bacteria inhibition and inhibition from accumulation of long chain fatty acids (Hickey et al., 1987; Lalman and Bagley, 2001). But acidogenic states are not the only cause of VFA buildup and are sometimes difficult to detect (Dupla et al., 2004). For this reason, and because they are a challenging topic from a scientific and industrial point of view, both imbalances will be considered as case studies.

Because anaerobic digestion processes are complex systems, many variables have to be monitored to diagnose the state of the process. To simplify the diagnosis process, it is important to reduce the amount of information needed whenever possible.

Many techniques have been widely used to reduce amounts of information in terms of features, variables, and dimensions in complex processes (Ruiz et al., 2004). Some of them are statistical (e.g., principal component analysis [PCA] and nonlinear discriminant analysis) and others are machine learning techniques such as several neural networks methods (e.g., self-organizing maps and radial basis function networks). In Ng et al. (2008), it is stated that PCA and other transformation-based feature reduction methods do not select features from the original feature set. Those methods transform the feature set into a lower-dimensional feature vector by combining several features. Transformation-based feature reduction methods do not reduce the cost of future sample collection and storage. Moreover, the newly created feature vector typically is difficult to interpret. For example, in the physiology field, a feature vector may be composed of blood pressure times the square of body height; this feature does not help people understand the problem. Neural networks provide a new suite of nonlinear algorithms for feature extraction (using hidden layers) and classification (e.g., multilayer perceptrons). Even with seemingly different underlying principles, most of the well-known neural networks models are implicitly equivalent or similar to classical statistical pattern recognition methods. Despite these similarities, neural networks do offer several advantages, such as unified approaches for feature extraction and classification and flexible procedures for finding good, moderately nonlinear solutions (Jain et al., 2000).

There are several ways to use artificial neural networks (ANN) at wastewater treatment plants (WWTP). Belanche et al. (2000), for example, used ANN to predict bulking in an activated sludge system. It also has been used to diagnose the state of the anaerobic digestion process. Sahely and Bagley (2001), for example,

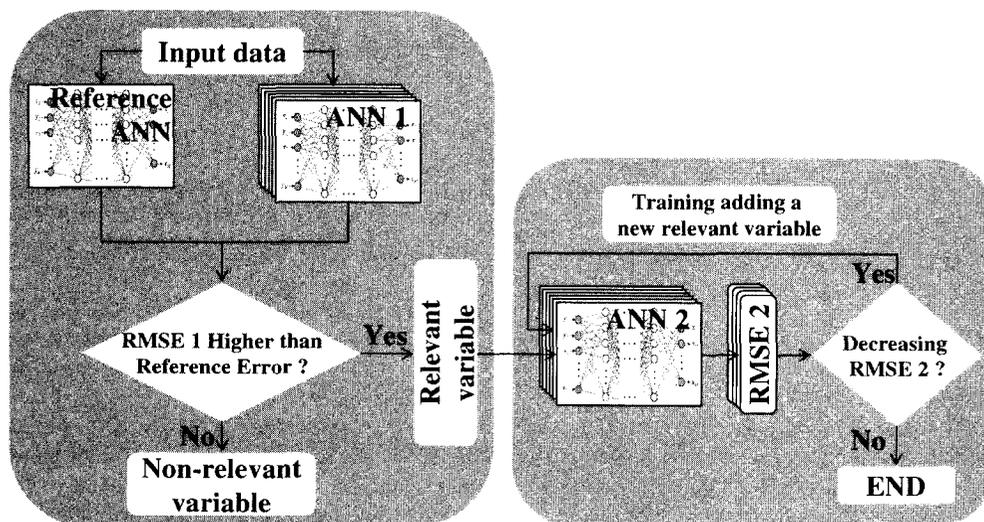


Figure 1—Detailed scheme of the method to choose the most relevant variables (RMSE = root mean square error; ANN = artificial neural networks).

developed a Bayesian belief network to infer the unbalanced states of the anaerobic digester; Steyer et al. (1997) used a hybrid ANN and fuzzy system to identify dangerous or faulty episodes. Ozkaya et al. (2007) used ANN to predict methane produced in landfills, and Strik et al. (2005) used it to model trace compounds in biogas.

The aim of this research was to select the most relevant variables for anaerobic digestion imbalances using a feature selection method on real data from a pilot plant. Feature selection improves monitoring by reducing the number of probes and other monitoring instrumentation required and minimizing the amount of information needing to be stored in fully instrumented systems. Variable selection also can be useful for ranking available information in case of probe failures. Variable reduction can be used to identify which variables are worth monitoring to prevent foaming in

anaerobic digestion because the causes are not well known. Results can then be used to choose the relevant variables as inputs to develop knowledge-based systems to assess the risk of foaming in real plants or in simulations, which Comas et al. (2008) did with activated sludge. Innovation related to the feature selection method used in this work is beyond the scope of this work.

Materials and Methods

Artificial Neural Networks. A neural network can be regarded as a nonlinear mathematical function that transforms a set of input variables into a set of output variables. The exact architecture of the ANN is determined by the number of nodes (neurons) and the layers in which they are distributed: the more neurons and layers, the more complex the ANN structure will be.

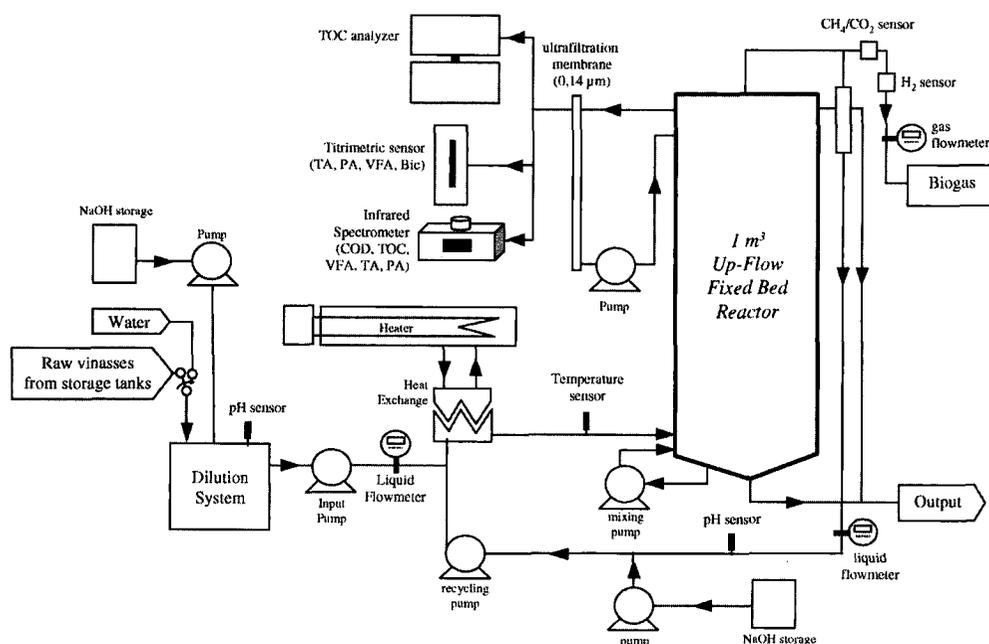


Figure 2—Layout of the upflow anaerobic fixed bed used in this study.

Table 1—Characteristic of the wastewater treated (COD = chemical oxygen demand).

Component	Range
Total COD (g/l)	[12–34]
Soluble COD (g/l)	[7.6–31]
Volatile fatty acids (g/l)	[1.5–10]
Total suspended solids (g/l)	[2.4–10]
Volatile suspended solids (g/l)	[1.2–5.4]
Phenol (mg/l)	[90–275]
pH	[4.5–5.2]

Therefore, it will always be better to have simpler structures because more complex structures require higher training and computation times. Typically, ANNs have an input layer, one or more hidden layers, and an output layer. Each node in the hidden and output layers contains a function (transfer function) that makes the node active if a certain threshold is reached. Transfer functions are governed by a set of parameters called “weights”, the values of which can be determined based on a set of examples of the required mapping. The training process can be compared to the fitting of a polynomial curve through a set of data points. Each data point consists of an input value and an output value. The output values are called “target values”. To train the ANN, an error between the calculated and desired outputs is calculated. This error function depends on the input data points and the weights. Therefore, the training process is an optimization strategy to minimize the error function by varying the weights until an optimum set of weights is found (Bishop, 1994; Jain et al., 1996).

This study used a neural network toolbox script for static models for use in MATLAB 3.0 or higher. Three layers were chosen in all the ANN architectures: an input layer, a hidden layer with sigmoid transfer functions, and an output layer with linear transfer functions for outputs (Jain et al., 1996). The exact number of nodes was determined before starting the method by training and validating the ANN with different numbers of nodes (from 1 to 10). The ANN with the best validation results was chosen. The initialization method was performed using the Nguyen-Widrow algorithm option, which initializes the weights with random values, later selecting their probability distributions to make all neurones active for the expected data ranges (Nguyen and Widrow, 1990). It also provided automatic data scaling and weights conversion. Bayesian regularization was used to prevent over-fitting.

The Wrapper Approach. The feature/variable selection method used in this study was based on a wrapper approach, which uses the leave-one-out searching approach (Kohavi and John, 1997). In each step, this method evaluates the accuracy of the learning algorithm when one of the features is left out. It then removes the feature yielding the least reduction in accuracy, such as, for example, validation error (Ng et al., 2008). The ANNs were

Table 3—Main reference artificial neural network features.

Hidden neurons	Maximum standard deviation error of the residuals	Minimum standard deviation error of the residuals	Standard deviation error of the residuals	Reference error
1	0.15	0.01	0.0545	0.1294

used as learning algorithms to select a subset of features; the root mean squared error (RMSE) was used as evaluation criteria (Figure 1).

The first step was to run the training-and-validation process ten times using the reference ANN with all the input variables and the respective output (i.e., occurrence of acidogenic states in the first case study and estimated risk of foaming in the second). Each training process was performed with 70% of the total data randomly selected. The rest of the data (i.e., 30%) was used for the respective validations. The average RMSE of the ten training-and-validation processes was calculated and stored as the reference error.

Next, one input variable was removed, and a new ANN (ANN1 in Figure 1) was trained and validated ten times without it. This last step was repeated for each input variable ending up with n —ANN1 for each removed input variable with their related average RMSE1. Whenever a relevant variable was removed, the average RMSE1 of the related ANN1 increased with respect to the average reference error. On the other hand, whenever a nonrelevant variable was removed, the RMSE1 of the related ANN1 decreased. Therefore, the variables in which RMSE1 was higher than the reference error were selected as relevant variables.

To assess the quality of the results, the second part of the method was devoted to showing that at least the same precision (an RMSE equal to the reference error) can be obtained with less input variables. Among the relevant variables, the one with the higher RMSE1 was selected first, and a new ANN (ANN2) was trained ten times using it as the only input. If the related average RMSE (RMSE2) was higher than the average reference error, then no improvement was found. In this case, the variable with the second highest average RMSE1 is selected and a new ANN2 is trained (ten times) with both variables. Again, the average RMSE2 is compared with the reference error. This iterative process is repeated until an average RMSE2 lower than the average reference RMSE is obtained.

Pilot Plant. Experimental data were obtained using an upflow anaerobic fixed-bed process of 1 m³ from INRA, France. Its schematic layout is shown in Figure 2. The reactor was a circular column 3.5 m high with a 0.6 m diameter, and a total volume of 0.982 m³. The support used, Cloisoyl with 180 m²/m³ specific surface, filled 0.0337 m³, leaving 0.948 m³ of effective volume. The support created 135 m² of surface. Three storage tanks of

Table 2—Input variables for the acidogenic states case study.

Inflow rate	pH in the digester	Gas flow rate	Hydrogen percentage in the gas phase	Methane percentage in the gas phase	Volatile fatty acids concentration in digester	Chemical oxygen demand in digester
qIn (L/h ⁻¹)	phDig	qGas (L/h ⁻¹)	h2Gas (%)	ch4Gas (%)	vfaDig (mg/L ⁻¹)	codtDig (mg/COD·L ⁻¹)

Table 4—Root mean squared error (RMSE) 2 for acidogenic states using the selected variables in relevance order (ANN = artificial neural networks; codtDig = total chemical oxygen demand in digester; qIn = inflow rate; vfaDig = volatile fatty acid concentration in digester; and phDig = pH in digester).

ANN input variables	codtDig	codtDig, qIn	codtDig, qIn, vfaDig	codtDig, qIn, vfaDig, phDig
RMSE2	0.1856	0.13106	0.13175	0.10952

27 m³ each were connected to a dilution tank (not shown) by a 60 m long piping system with a 0.1 m diameter. The wastewater came from a wine distillery and had the characteristics shown in Table 1. The anaerobic digestion process was fully instrumented with online sensors both in the liquid phase (pH, flow rates, VFA, total organic carbon [TOC], and chemical oxygen demand [COD] concentrations) and in the gas phase (i.e. CO₂, CH₄, H₂). For further details about the pilot plant's features and instrumentation, see Lardon et al. (2004).

The process was designed to be operated at an HRT of three days with input COD of 15 g/L. Over the 10 years of operation of the process, average COD removal was more than 90% during normal conditions; however, in some cases (i.e., specific experiment to evaluate the robustness of the process), HRT could vary between one day and several weeks (with very low input liquid flow rate), leading to high concentration of VFAs being produced (up to 9 g/L) and very low biogas production. Of course, in these cases, the performance of the process was very low, but this allowed the process response to be evaluated under different conditions.

Results

Case Study: Acidogenic States. The acidogenic states case study comprised three months of data corresponding to almost 400 inputs and outputs collected every six hours. Table 2 shows the seven input variables.

Outputs for the ANN training and validation (i.e., acidogenic states occurrence) were provided by a fuzzy modular expert system presented in Lardon et al. (2005). In this fuzzy system, several states of the process were diagnosed: normal, acidogenic, underload, organic overload, hydraulic overload, and toxics. For each of the 389 input vectors, Lardon's fuzzy system was used to evaluate the occurrence of acidogenic states between 0 and 1 (1 being the most likely acidogenic state occurrence). Thus, for training and validation of each ANN, an output vector with pieces of 389 data was used, indicating that the acidogenic state occurred between 0 and 1.

Reference Artificial Neural Networks. Table 3 shows the main features of the reference ANN. When the standard deviation

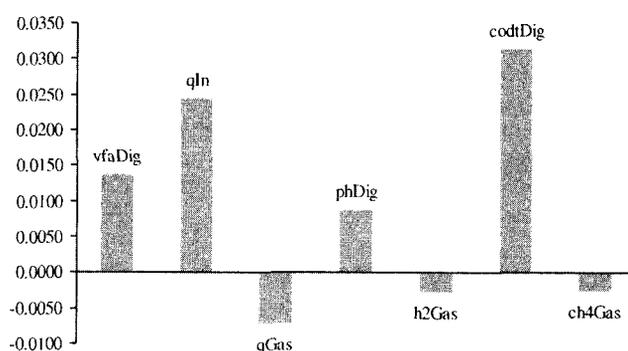


Figure 3—Average root mean square error (RMSE) minus average reference RMSE for each removed input variable (vfaDig = volatile fatty acids in digester; qIn = inflow rate; qGas = gas flow rate; phDig = pH in the digester; h2Gas = hydrogen percentage in gas phase; codtDig = total chemical oxygen demand concentration in digester; and ch4Gas = methane percentage in gas phase).

error of the residuals falls between the maximum and minimum values, which are set *a priori*, the ANN can be considered to be correctly trained. Values of the standard deviation error equal to the minimum boundary indicate over-fitting, which means that higher than necessary precision has been obtained at the expense of good generalization ability. When the standard deviation error of the residuals is equal to the maximum boundary, however, poor precision has been obtained.

The next step was to train the different ANNs. Figure 3 presents the resulting average RMSE1 for each training; each label indicates the ANN trained without the shown variable.

The highest difference corresponds to the ANN trained and validated without the COD in digester, codtDig, variable. This indicates that this variable is the most important for ANN performance. In order of decreasing relevance, the variables are: inflow rate (qIn), VFA concentration in the digester (vfaDig), and pH in the digester (phDig).

The input variable with the highest validation error was used first to train ANN2, or the ANN with the selected variables. To find out the input variable combination with the minimum RMSE 2, the variable which had the highest validation error among the remaining variables was added to the first ANN2 as an input (i.e., first ANN with vfaDig, second with phDig, and third, qIn). Results are provided in Table 4.

Because the codtDig is the most relevant variable, training of ANNs 2 starts with this variable. This variable alone, however, is not enough to diagnose acidogenic states with better precision than the reference ANN (reference error = 0.1294). When

Table 5—Input variables for foaming case study (VFA = volatile fatty acids; TOC = total organic carbon; qIn = inflow rate; vfaDig = VFA concentration in digester; tocsDig = TOC concentration in digester; phIn = pH in the inflow; co2Gas = carbon dioxide percentage in gas phase; and ch4Gas = methane percentage in gas phase).

Inflow rate	VFA concentration in digester	TOC in digester	pH in the inflow	pH in the digester	Carbon dioxide percentage in the gas phase	Methane percentage in the gas phase
qIn (L/h ⁻¹)	vfaDig (mg/L ⁻¹)	tocsDig (mgTOC/L ⁻¹)	phIn	phDig	co2Gas (%)	ch4Gas (%)

Table 6—Main reference artificial neural network features.

Hidden neurons	Maximum standard deviation error of the residuals	Minimum standard deviation error of the residuals	Standard deviation error of the residuals	Reference error
8	0.15	0.01	0.10734	0.11700

codtDig is combined with the next relevant variable (qIn), the precision of the resulting ANN (RSME2) is not lower than reference error. When vfaDig is combined with codtDig and qIn, the RSME2 of the resulting ANN2 is increased, but it is still higher than the reference error. The best result was obtained when the four variables were used together (codtDig, qIn, vfaDig, and phDig). These results—for phDig, vfaDig, and codtDig—are widely recognized as key variables when diagnosing acidogenic states (Bjornsson et al., 1997; Guiot, 1991; Tartakovsky et al., 2000). Nevertheless, it is important to consider that, in addition to codtDig, vfaDig, and phDig, other variables somehow not directly related to acidogenic states could be chosen, of which qIn is the most appropriate.

Case Study: Foam Forming. Experimental data used for the case study to detect foam forming were obtained from the same pilot plant. Some of the available variables were derivatives of the measured variables of the process. To avoid possible redundancy, these derivatives were not used. In addition, derivatives typically are not calculated in real plants. Other variables, like temperature and alkalinity, are not used because they were kept relatively constant, and it would be difficult to extract information from their profile. Table 5 presents the input variables. For this study, more than 8000 inputs and output data points were collected in total, equivalent to approximately two months of data. As output, foaming appearance in the digester was used, based on the heuristic knowledge provided by experts. When foaming appeared in the digester, high variations of the gas flow rate and pressure appeared because of the slug release of gas bubbles trapped inside the foam. As a result, fuzzy system was used to ensure a suitable foaming index between 0 and 1. Although foaming can be determined in this way, this approach typically is used to study variables influence or relation. Overall, a set of 8133 data was used for this case study, which included one foaming episode at the end of the period. The main reference ANN features for this case study are presented in Table 6.

The average RMSE of the reference ANN and the seven ANNs 1 are presented in Table 7.

Figure 4 shows the differences between the RMSE of each variable and the reference RMSE. As shown, five variables had an RMSE

Table 7—Reference artificial neural network (ANN) and ANN1 average root mean squared error (qIn = inflow rate; vfaDig = volatile fatty acids concentration in digester; tocsDig = total organic carbon concentration in digester; phIn = pH in inflow; co2Gas = carbon dioxide percentage in gas phase; and ch4Gas = methane percentage in gas phase).

qIn	vfaDig	tocsDig	phIn	phDig	co2Gas	ch4Gas
0.11934	0.11377	0.12764	0.12654	0.11328	0.11796	0.12179

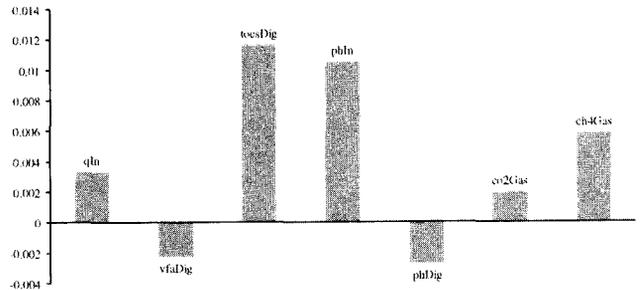


Figure 4—Average root mean square error (RMSE) minus average reference RMSE for each variable (qIn = inflow rate; qGas = gas flow rate; vfaDig = volatile fatty acids in digester; tocsDig = total organic carbon concentration in digester; phIn = pH in inflow; phDig = pH in the digester; co2Gas = carbon dioxide percentage in gas phase; and ch4Gas = methane percentage in gas phase).

higher than the reference RMSE. Therefore, the selected as relevant variables were inflow rate (qIn); TOC in the digester (tocsDig); pH in the inflow (phIn); carbon dioxide percentage in the gas phase (co2Gas); and methane percentage in the gas phase (ch4Gas).

To train ANN2s, first tocsDig was selected. Because its average RMSE2 is higher than the average reference RMSE, the second most relevant variable was added and ANN 2 was trained again. Table 8 summarizes the average RMSE2 for each ANN.

In all cases, RMSE2 was higher than the reference error. A t-test comparing the averages between each RMSE2 and the RMSE of the reference ANN, however, revealed that in the last ANN2 (last column in Table 8), the average RMSE 2 was not significantly different than the reference RMSE. This means that although the average RMSE 2 was not lower than the reference RMSE, by removing two variables from the whole input, similar results can be obtained. This can be important in terms of saving sensors and for monitoring relevance.

The higher relevance of phIn over phDig probably is because pH is controlled in the anaerobic digesters. In which case, phDig will show more constant behavior than pH in the influent; therefore, phIn will provide more information than phDig. The significance of gas-related variables (i.e., co2Gas and ch4Gas) likely is because of the approach taken to determine foaming—the fuzzy system that was used. The addition of co2Gas to the relevant variables, however, reduces RMSE2. According to Zhao and Viraraghavan (2004), high CO₂ production is representative of poor digestion and may lead to foaming. Other selected variables, such as tocsDig and qIn, also have relevance as indicated by the literature. In the study of Ross and Ellis (1992), increasing organic

Table 8—Average root mean squared error 2 (RMSE2) for the relevant variables (tocsDig = total organic carbon concentration in digester; phIn = pH in inflow; ch4Gas = methane percentage in gas phase; qIn = inflow rate; co2Gas = carbon dioxide percentage in gas phase).

	tocsDig, tocsDig	tocsDig, phIn, ch4Gas	tocsDig, phIn, ch4Gas, qIn	tocsDig, phIn, ch4Gas, qIn, co2Gas
RMSE 2	0.12666	0.12256	0.12141	0.12138

loading rates and detention times were shown to cause foaming. Massart et al. (2006) stated that inconsistent feeding in the digester is a cause of foaming.

Conclusions

Real data from a pilot plant were used to select the most relevant variables for acidogenic states and foaming using a wrapper approach. Results showed that it is possible to monitor and control the process using less probes or alternative variables. The feature selection method provided alternatives to monitor the process in case of probe failures. It also was shown that ANN and feature selection methods successfully can be used in complex environmental systems to select relevant information.

For foaming, two out of seven variables were shown not to be relevant because of the complexity of this microbiology-related problem, indicating that many variables are involved. Nevertheless, when diagnosing or controlling the conditions, the selected variables are helpful indicators of what is worth controlling to avoid excessive foaming. As shown, results are supported by some of the current literature on this topic from an operational point of view. The information extracted with this approach also can be used to develop knowledge based-systems when applied to complex problems by helping, for example, to decide which variables to choose as inputs.

Acknowledgments

The authors wish to thank Dr. Cristian Trelea from the Institut National Agronomique Paris-Grignon for developing and providing us with the Neural Network toolbox. The work was partially supported by the Spanish Ministry of Education and Sciences projects DPI2006-15707-C02-01 and NOVEDAR_Consolider-CSD2007-00055 and by a mobility grant from the University of Girona under the Catalan DRAC program.

Submitted for publication November 5, 2008; revised manuscript submitted September 14, 2009; accepted for publication September 22, 2009.

References

- Barjenbruch, M.; Kopplow, O. (2003) Enzymatic, Mechanical and Thermal Pre-Treatment of Surplus Sludge. *Adv. Environ. Res.*, **7**, 715–720.
- Belanche, L.; Valdés, J. J.; Comas, J.; Rodríguez-Roda, I.; Poch, M. (2000). Prediction of Teh Bulking Phenomenon in Wastewater Treatment Plants. *Artif. Intell. Eng.*, **14** (4), 307–317.
- Bishop, C. M. (1994) Neural Networks and Their Applications. *Rev. Sci. Instr.*, **65** (6), 1803–1832.
- Björnsson, L.; Mattiasson, B.; Henrysson, T. (1997) Effects of Support Material on The Pattern of Volatile Fatty Acid Accumulation at Overload in Anaerobic Digestion of Semi-Solid Waste. *Appl. Microbiol. Biotechnol.*, **47**, 640–644.
- Carrasco, E. F.; Rodríguez, J.; Puñal, A.; Roca E.; Lema, J. M. (2004) Diagnosis of Acidification Status in an Anaerobic Wastewater Treatment Plant Using a Fuzzy-Based Expert System. *Control Eng. Pract.*, **12**, 59–64.
- Comas, J.; Rodríguez-Roda, I.; Gernaey, K. V.; Rosen, C.; Jeppsson U.; Poch, M. (2008) Risk Assessment Modelling of Microbiology-related Solids Separation Problems in Activated Sludge Systems. *Environ. Modell. Softw.*, **23**, 1250–1261.
- Duplá, M.; Conte, T.; Bouvier, J. C.; Bernet, N.; Steyer, J. P. (2004) Dynamic Evaluation of a Fixed Bed Anaerobic Digestion Process in Response to Toxic Shocks. *Water Sci. Technol.*, **49** (1), 61–68.
- Elliot, A.; Mahmood, Y. (2007) Pretreatment Technologies for Advancing Anaerobic Digestion of Pulp and Paper Biotreatment Residues. *Water Res.*, **41**, 4273–4286.
- Guiot, S. R. (1991) Modeling of the Upflow Anaerobic Sludge Bed-Filter System: A Case With Hysteresis. *Water Res.*, **25** (3), 251–262.
- Hickey, R. F.; Vanderwielen, J.; Switzenbaum, M. S. (1987) The Effects of Organic Toxicants on Methane Production and Hydrogen Gas Levels during the Anaerobic Digestion of Waste Activated Sludge. *Water Res.*, **21** (11), 1417–1427.
- Jain, A. K.; Duin, R. P. W.; Mao, J., (2000) Statistical Pattern Recognition: A Review. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22** (1), 4–37.
- Kohavi, R.; John, G. H. (1997) Wrappers for Feature Subset Selection. *Art. Intell.*, **97**, 273–324.
- Lalman, J. A.; Bagley, D. M. (2001) Anaerobic Degradation and Methanogenic Inhibitory Effects of Oleic and Stearic Acids. *Water Res.*, **35** (12), 2975–2983.
- Lardon, L.; Puñal, A.; Martínez, J. A.; Steyer, J. P. (2005) Modular expert system for the diagnosis of operating conditions of industrial anaerobic digestion plants. *Water Sci. and Technol.*, **52** (1–2), 427–433.
- Lardon L.; Punal A.; Steyer J.-P. (2004) On-Line Diagnosis and Uncertainty Management in Biological Wastewater Treatment Processes. *J. Process Contr.*, **14**, 747–763.
- Massart, N.; Bates, R.; Corning, B.; Neun, G. (2006) Design and Operational Considerations to Avoid Excessive Anaerobic Digester Foaming. *Proceedings of the 80th Annual Water Environment Federation Technical Exposition and Conference* [CD-ROM]; Dallas, Texas, Oct 21–25; Water Environment Federation: Alexandria, Virginia, 2553–2575.
- Ng, W. W. Y.; Yeung, D. S.; Firth, M.; Tsang, E. C. C.; Wang, X. Z. (2008) Feature Selection Using Localized Generalization Error for Supervised Classification Problems Using RBFNN. *Pattern Recognit.*, **41**, 3706–3719.
- Nguyen, D.; Widrow, B. (1990) Improving The Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of The Adaptive Weights. *Proceedings of International Joint Conference on Neural Networks*; International Neural Network Society; Washington D.C., January 15–19, 21–26.
- Ozkaya, B.; Demir, A.; Bilgili, M. S. (2007) Neural Network Prediction Model for the Methane Fraction in Biogas from Field-Scale Landfill Bioreactors. *Environ. Modell. Softw.*, **22**, 815–822
- Pagilla, K. R.; Kraney, K. C.; Kido, W. H. (1997) Causes and Effects of Foaming in Anaerobic Sludge Digesters. *Water Sci. Technol.*, **36** (6–7), 463–470.
- Ross, D. R.; Ellis, L.-A.M. (1992) Laboratory-Scale Investigation of Foaming in Anaerobic Digesters. *Water Environ. Res.*, **64** (2), 154–162.
- Ruiz, G.; Castellano, M.; González, W.; Roca, E.; Lema, J. M. (2004) Algorithm for Steady State Detection of Multivariate Process: Application to Anaerobic Wastewater Digestion Process. *Proceedings of The 2nd International IWA Conference on Automation in Water Quality Monitoring (Autmonet)*; Langergraber, G, Winkler, S., Fleischmann, N., Press, A., Haberl, R., Eds.; Vienna, Austria, April 19–20; IWA Publishing: London, England, United Kingdom.
- Sahely, B. S. G. E.; Bagley, D. M. (2001) Diagnosing Upsets in Anaerobic Wastewater Treatment Using Bayesian Belief Networks. *J. of Environ. Eng.*, **127** (4), 302–310.
- Schafer, P.; Uhte, W.; Newman, G. (2006) Improved Anaerobic Digestion Feed Systems and Concepts. *Proceedings of the 80th Annual Water Environment Federation Technical Exposition and Conference* [CD-ROM]; Dallas, Texas, Oct 21–25; Water Environment Federation: Alexandria, Virginia.
- Steyer, J. P.; Rolland, D.; Bouvier, J. C.; Moletta, R. (1997) Hybrid Fuzzy Neural Network for Diagnosis-Application to the Anaerobic Treatment of Wine Distillery Wastewater in a Fluidized Bed Reactor. *Water Sci. Technol.*, **26** (6–7), 209–217.
- Strik, D. P. B. T. B.; Domnanovich, A. M.; Zani, L.; Braun, R.; Holubar, P. (2005) Prediction of Trace Compounds in Biogas from Anaerobic Digestion Using the MATLAB Neural Network Toolbox. *Environ. Modell. Softw.*, **20** (6), 803–810.

- Svensson, L. M.; Björnsson, L.; Mattiasson, B. (2007) Enhancing Performance in Anaerobic High-Solids Stratified Bed Digesters by Straw Bed Implementation. *Biores. Technol.*, **98**, 46–52.
- Tartakovsky B.; Morel E.; Steyer J. P.; Guiot S. R. (2002) Application of a Variable Structure Model in Observation and Control of an Anaerobic Digester. *Biotechnol. Prog.*, **18**, 898–903.
- Westlund, A. D.; Hagland, E.; Rothman, M. (1998) Foaming in Anaerobic Digesters Caused By *Microthrix Parvicella*. *Water Sci. Technol.*, **37** (4–5), 51–55.
- Zhao, H. W; Viraraghavan, T. (2004) Analysis of the Performance of an Anaerobic Digestion System at the Regina Wastewater Treatment Plant. *Biores. Technol.*, **95**, 301–307.