International Congress on Environmental Modelling and Software

7th International Congress on Environmental Modelling and Software - San Diego, California, USA - June 2014

Jun 17th, 2:00 PM - 3:20 PM

Where Are We In Wastewater Treatment Plants Data Management? A Review and a Proposal

Manel Poch
University of Girona, manuel@lequia.udg.cat

Joaquim Comas
University of Girona, quim@lequia.udg.cat

José Porro
University of Girona, jose.porro@lequia.udg.cat

Manel Garrido-Baserba
Urban Water Research Center, University of California, manelg@uci.edu

Lluis Corominas
Catalan Institute for Water Research (ICRA), Scientific and Techological Park of the University of Girona, lcorominas@icra.cat

Poch, Manel; Comas, Joaquim; Porro, José; Garrido-Baserba, Manel; Corominas, Lluis; and Pijuan, Maite, "Where Are We In Wastewater Treatment Plants Data Management? A Review and a Proposal" (2014). *International Congress on Environmental Modelling and Software*. 21.
https://scholarsarchive.byu.edu/iemssconference/2014/Stream-G/21

Presenter/Author Information

Manel Poch, Joaquim Comas, José Porro, Manel Garrido-Baserba, Lluis Corominas, and Maite Pijuan

International Environmental Modelling and Software Society (iEMSs)
7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA,
Daniel P. Ames, Nigel W.T. Quinn and Andrea E. Rizzoli (Eds).
http://www.iemss.org/society/index.php/iemss-2014-proceedings

Where Are We In Wastewater Treatment Plants Data Management? A Review and a Proposal

Manel Poch[a], Joaquim Comas[a], José Porro[a], Manel Garrido-Baserba[b], Lluis Corominas[c], Maite Pijuan[c]

[a] LEQUIA. Institute of the Environment. University of Girona (UdG) Universitat de Girona. E-170171 Girona. Spain (E-mail: manuel@lequia.udg.cat; quim@lequia.udg.cat;jose.porro@lequia.udg.cat)

[b] Urban Water Research Center. 844G Engineering Tower. University of California, Irvine. Irvine, CA. 92697-2175. USA. (E-mail: manelg@uci.edu)

[c] Catalan Institute for Water Research (ICRA). Scientific and Technological Park of the University of Girona. Emili Grahit 101. 17003. Girona. Spain (E-mail: lcorominas@icra.cat; mpijuan@icra.cat)

Abstract: Wastewater treatment plants (WWTP) are comprised of complex processes that need to be optimally managed. To attain that, in the last years an impressive effort has been made to incorporate monitoring devices able to provide from several hundred to more than ten thousand signals. With the aim to take benefit of those data, different data mining techniques have been applied to transform them into information and knowledge in order to help WWTP's managers. Furthermore, several mathematical models have been developed intending to simulate process behaviour including biomass and pollutants transformation. However, it is recognized that this it is not enough to cope with all the operational problems, but rather it is necessary to integrate heuristic knowledge to better manage specific situations. In this context, it is clear that hybrid systems integrating ontologies, statistical techniques and mathematical models should be developed. In this communication, an overview of some of the main tools applied to obtain information and knowledge from raw data in WWTP's is presented. This overview is complemented by the work developed by the authors of managing two important specific problems: sludge bulking and greenhouse gas emissions, which are good examples of situations that require integration of different types of knowledge.

***Keywords**:* WWTP management, data mining, heuristic knowledge

1. ANALYSIS OF THE PROBLEM

1.1 Co-evolution of objectives and monitoring capacity

Wastewater Treatment Plants (WWTPs) embrace a wide range of interrelated biological, physical and chemical processes. Thus, complexity in WWTPs stems from various sources (ill-structured and non-linear domain, high dimensionality, heterogeneity of data, non-stationary, multi-granularity, cyclic behaviour, etc.)

Several factors (economic, technical and environmental) are increasingly used to assess the performance of WWTPs. Moreover, the occurrence of new objectives and a new environmental and sustainability paradigm would even increase the need for further control of the plant as new parameters are foreseen to be taken into account in next years (e.g. sensors for greenhouse gas emissions, control of odours substances, priority pollutants…). Thus, as the number of objectives and considerations increase, the complexity in the operation and management consequently raises corresponding to the implementation of new concepts, new relations, or even for the consideration of new details and insights that current technology unravels over the conventional processes. Simultaneously, the number of signals that a WWTP generates has increased exponentially. In the last years an impressive effort has been made to incorporate new monitoring devices able to give up to thousands of signals. And similar advances can be described in relation to WWTP automation systems, where modern devices make use of the latest software and hardware standards in the field of communication technologies.

Our proposal is that the co-evolution of these two factors (new objectives in the sustainability paradigm and development of advanced monitoring systems) may be described according with the scheme presented in figure 1 where the results obtained from data acquisition have evolved along the time producing crisp data, information and knowledge according with the tools applied (statistics and data mining).

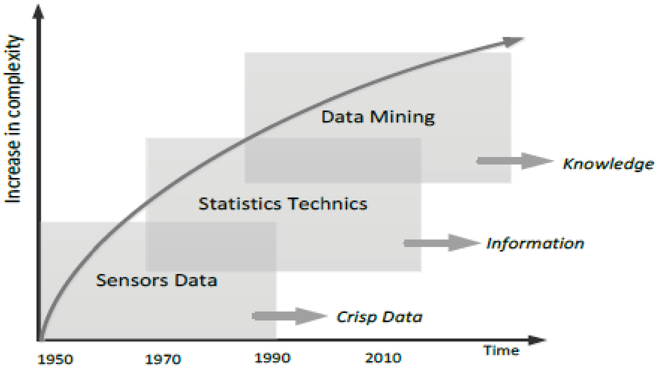

Figure 1. Evolution of results obtained from data acquisition in WWTP

1.2 Evolution of tools for data processing

Already in the 1970s researchers realised that WWTPs were designed assuming steady-state conditions even if the system never seems to be. The challenge to consider dynamic changes and disturbances leverage the control of WWTP and the implementation of sensors intended to provide more information of the system variability. The processing of signals from WWTP started and in the 1970s, when the use of on line dissolved oxygen (DO) sensors were well established (Olsson, 2012).

Early research on this topic was focused on data handling, error correction and basic signal filtering. The idea of applying parameter estimation and process identification in wastewater treatment appeared first in the 70's and the estimation methods have been gradually more become more sophisticated.

In the 80s the amount of data was increased and new methodologies applying basic and advanced statistics to the databases improved the information acquisition process in WWTP management. Statistics applied to the environmental engineering field (using control algorithms) progressively became a new source of fundamental information revealing the insights of the treatment process, optimizing capacities of existing treatment plants and enhancing the predictive control, building the pillars of knowledge acquisition.

In this sense, the so-called Knowledge Discovery from Databases (KDD) process and several data mining techniques have already shown to provide successful tools for dealing with this complexity. There is an important work to be done between getting the results of the data mining techniques and using them to support decision-making. This step is highly dependent on both data mining techniques used and the problem goals, as has been identified by Gibert et al. (2005, 2010).

The capacity to predict (modelling) and control WWTP using more advanced techniques did not stop in the following years. Thus, knowledge steaming from new analytical techniques was creating a large amount of rich knowledge requiring serious attention. In this respect, in the 90's the first approaches to integrate the aforementioned expert systems, rule-based systems and the other analytical methodologies (even, approaches able to collect experience from operators and designers) were presented under the structure of the so-called Knowledge-based reasoning systems (Serra et al. (1994) that were followed along the decade as pointed out by Poch et al. (2004) with the development of Environmental Decision Support Systems (EDSS).

In the XXI century innovative knowledge-based systems evolved to deal with multi- criteria decision in view of plant-wide control perspective using new knowledge management approaches based on knowledge-based systems (using numerical models and heuristic knowledge) coupled with classical and innovative knowledge acquisition techniques. But, there is still much work to do for the definitive approach for EDSS design and it will probably be the dominating development for the next decade. It is now recognized by experts that there is a need to integrate and complement the complex WWTP system with other interdependent and equally complex systems, as the sewer system and the receiving media. Therefore, a plant-wide perspective that seeks for the integrated control not only the various unit processes of the WWTP, but also the relationship between the sewer and receiving media is required. As was previously stated, a successful plant-wide control (multi- criteria optimization) cannot be based on a single technique. In this respect, decision support systems arise to fill the missing gap in the integration of multi- criteria, interdependent systems integration (sewer and river basin), modelling, and qualitative operational experience. Thus, EDSS are key elements to make the plant-wide control possible as presented in figure 2. It shows the methodologies applied in practice, the results obtained after applying these methodologies, the type of implementation, the end product and the relationship with the different objectives which have been evolving along the years from single process control, to plant-wide control and up to system-wide control.

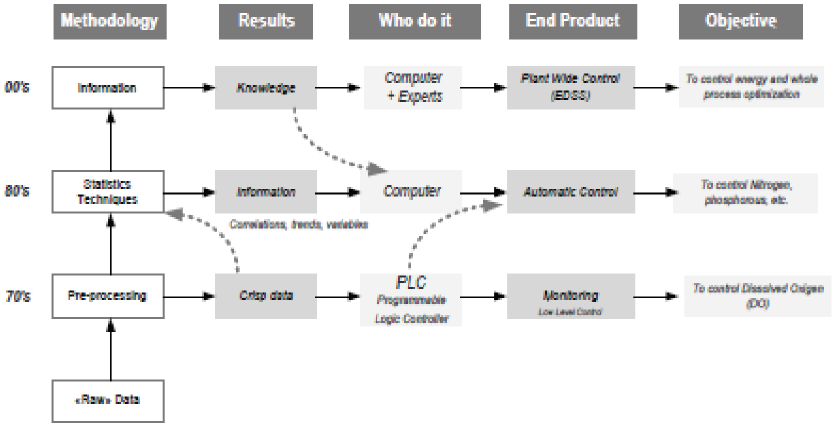

Figure 2. A proposal of relationship between methodologies, results, and objectives

2. KNOWLEDGE-BASED DECISION SUPPORT TOOLS FOR ASSESSING RISK OF WWTP SEPARATION AND N2O PRODUCTION PROBLEMS

2.1 Activated sludge solids separation risk model

descriptions, but common operational issues related to population dynamics are still not well understood. Many of these problems relate to population imbalances between filamentous and floc-forming bacteria leading to filamentous bulking, filamentous foaming and deflocculation. It is not uncommon to combine heuristic knowledge from the WWTP operators with specific data streams to identify these problems. However, the predictive power of this approach is limited as these data tend to appear only after the onset of the problem. The presence of heuristics and qualitative knowledge on complex phenomena such as filamentous bulking, foaming and rising sludge stands in sharp contrast with the lack of basic mechanistic knowledge on the population dynamics of the microorganisms causing these phenomena. As a consequence, although a limited number of attempts have been made to explain the development of filamentous bacteria by means of mathematical modelling, none of these attempts has led to a general and experimentally validated model. As an alternative, a qualitative risk assessment model, which integrates empirical knowledge with the mechanisms of standard deterministic models, has been reported to infer solids separation problems of microbiological origin, as a function of influent composition and operational conditions (Comas *et al.*, 2008). Initially, the knowledge related to the risk of filamentous bulking proliferation was synthesised into a decision tree. Each branch of the tree evaluates one assumed cause. The mathematical representation of the decision tree is captured using the principles of fuzzy decision theory. The output of this model (AS risk model) is a qualitative risk score between 0 (low risk) and 1 (high risk), indicating risk of developing microbiological operational problems, considering a score of 0.8 as the threshold for high risk, as well as the percentage of time (%) under high risk.

This AS risk model has been already applied for various case studies to provide a third dimension for evaluation along with the more conventional effluent quality and economic criteria. The consideration of this third dimension enables to reach more robust decision support, since in efforts for identifying optimization strategies could lead to conditions causing operational problems, if only quality and costs criteria were assessed.

To highlight the decision support provided by this knowledge-based approach, the AS risk model was recently applied in a study identifying MBR process optimization opportunities in a full-scale MBR plant (Terrassa WWTP, Terrassa, Catalonia, Spain), through a model-based approach (Gabarrón et al., 2014). Specifically, Table 1 illustrates the AS risk model results for two different control strategies (scenario 1 with DO current set point, 1.2 mg/L O_2, and scenario 2 with DO set point decreased to 0.8 mg/L O_2), together with effluent quality and aeration energy costs (AEC). Specifically, the decrease of the DO set point to 0.8 mg/L O_2 did not increase the severe risk (higher than 80%) due to the low DO, while showing a significant reduction in AEC and improvement of effluent quality. On the contrary, if the DO had been decreased to 0.5 mg/L O_2, despite of achieving the highest energy savings, the severe risk of bulking episodes due to low DO values would have increased from 0 to 3 % (results not shown), thus deteriorating the sludge properties.

2.2 N_2O production risk model

As a result of the new environmental and sustainability paradigm previously discussed, WWTP operators have been moving towards energy neutrality and more eco-friendly operations through activated sludge process optimization.

However, actions to save energy and carbon dioxide (CO_2) emissions include lowering dissolved oxygen (DO) set points, as aeration typically comprises a significant portion of the WWTP energy costs. Although it may not negatively impact effluent water quality, lowering the DO can cause oxygen limited conditions in the biological reactors, which can promote N_2O production via denitrification by ammonia oxidizing bacteria (AOB) because AOB can use nitrite (NO_2) as the terminal electron acceptor as opposed to oxygen, and hence, lead to increased N_2O production as the NO_2 is converted to N_2O (Kampschreur et al., 2009). This process, also known as AOB denitrification, can significantly increase the overall carbon footrpint of a WWTP given that the global warming potential of N_2O is significantly higher than that of CO_2. Therefore, a tool is needed to assess the risk of N_2O production from proposed WWTP control optimization in energy neutrality schemes. Building on the efforts and motivations of Comas et al. (2008) in integrating mathematical modelling with heuristic and empirical knowledge from literature and domain experts, to qualitatively assess the risk of WWTP solids separation problems, as described in the previous section, Porro et al. (2014) proposed a similar knowledge-based approach for qualitatively assessing the risk of N_2O production in WWTPs. As in the case of filamentous bulking, filamentous foaming and deflocculation, N_2O production is caused by complex mechanisms, and it is produced in several pathways: i) the previously mentioned AOB denitrification, ii) through the oxidation of hydroxylamine by AOB (hydroxylamine pathway (Ni et al., 2013)), and via heterotrophic denitrification (Schulthess et al., 1994). At the moment, there are no widely validated mechanistic models, but it exists a significant amount of knowledge in the literature with regard to risk of N_2O production associated with various operational parameters (Porro et al., 2014).

Extending the case of the Terrassa WWTP, the knowledge-based risk assessment model proposed by Porro et al. (2014) (N2O risk model) was implemented in Excel and applied using the Terrassa WWTP mathematical model output data to dynamically asses the risk of N_2O production due to low DO for the same two control scenarios, current operation DO and maintaining a constant DO of 0.8
$mg{\cdot}L^{-1}$.

Figure 3 shows the dynamic risk results for Scenarios 1 and 2. As the current DO control results in a fluctuating DO concentration (0.8 – 1.6 mg·L^{-1}) based upon influent load, the N_2O risk shown for Scenario 1 in Figure 3 varies substantially; however, it does not exceed 0.64 on a risk scale of 0 to 1, as defined for the AS risk model. As Scenario 2 maintains a constant DO of 0.8 mg·L^{-1}, the N_2O risk is constant at 0.64.

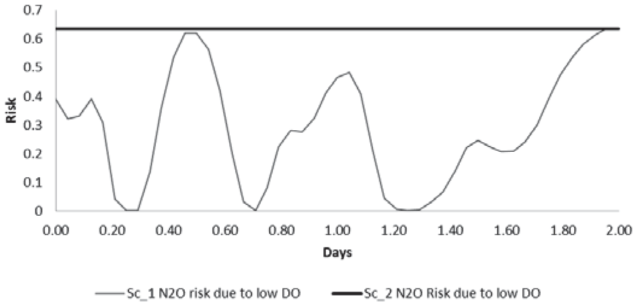

Figure 3. Dynamic N_2O Risk Results for Terrassa WWTP Dynamic Simulation Scenario Analysis

Table 1 also includes average N_2O risk, along with AS risk, effluent quality, and AEC results discussed in the previous section. As expected, lowering the DO would increase the risk of N_2O production due to low DO (AOB denitrification); however, it only increases by approximately 0.3 and is still not under severe risk, (above 0.8). Depending on the utility, one can look at the results and think that since there was only a slight increase in N_2O risk, then implementing the Scenario 2 control strategy would not significantly increase environmental impact and would lead to cost savings, better effluent water quality, and no cause solids separation problems. From a more environmental perspective, one could also look at the results and conclude that the savings were not substantial, and hence, not worth the slight increase in N_2O risk. Therefore, the results demonstrate how a fourth (N_2O risk) dimension could add an even greater decision support than already provided by the AS risk model. As the N_2O risk model was implemented in Excel, the approach could also be applied with online data to review risk of current and historical operations and not necessarily with mathematical model output data. This may be of interest to WWTP operators who do not have a developed and calibrated process model of their plant.

Table 1. Summary of control scenario benchmarking results including AS risk and N_2O risk due to low DO

	Scenario1 DO_current	Scenario2 DO_0.8
Average N_2O Risk due to low DO	0.3	0.6
AS risk due to low DO (% time under severe risk)	0	0
Effluent quality (kg poll·d^{-1})	3000	2940
Energy costs (€·d^{-1})	226	218

3. CONCLUSIONS

Wastewater treatment plants are a good example of co-evolution of sustainable objectives and data processing tools application. From the seminal works of the 70's with the use of dissolved oxygen probes until the use, nowadays, of advanced data mining techniques and environmental decision support systems there has been a long way.

However, although some important successes has been obtained using mathematical models or in simulation frameworks, more work is needed to cope with the complex processes that take place in real WWTP.

In this context, this manuscript illustrates that the use of integrating of heuristic and qualitative knowledge may be useful to manage two important operational problems: bulking phenomena and N_2O emissions. In both cases, dynamic risk is evaluated under different control strategies, which may support WWTP manager's decisions.

4. ACKOWLEDGEMENTS

The Spanish Government is acknowledged for funding provided via Projects CTM 2011-27163 and CTM 2012-38314-C02-01. M. Pijuan acknowledges the Ramon y Cajal research fellowship (RYC-2009-04959) also provided by the Spanish Government and M. Pijuan and L. Corominas the European Comission for their FP7-PEOPLE--CIG projects.

5. REFERENCES

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.

Gabarrón et al., 2014 Bioresource Technology (submitted)

Gibert, K., Sànchez-Marrè, M., Flores, X., 2005. Cluster discovery in environmental databases using GESCONDA: The added value of comparisons AI Communications 18 (4).

Gibert, K., Rodríguez-Silva, G., Rodríguez-Roda, I., 2010. Knowledge discovery with clustering based on rules by states: A water treatment application. Environ Modell Softw 25 (6), 712-723

Kampschreur, M.J.; Temmink, H.; Kleerebezem, R.; Jetten, M.S.M.; van Loosdrecht, M.C.M., 2009. Nitrous Oxide Emission During Wastewater Treatment. Water Res*.,* 43(17), 4093-4103.

Ni, B.J.,Ye, L., Law, Y., Byers, C., Yuan, Z., 2013. Mathematical Modeling of Nitrous Oxide (N_2O) Emissions from Full-Scale Wastewater Treatment Plants. Environ. Sci. Technol., *47* (14), 7795–7803.

Olsson, G., 2012. ICA and me – a subjective review. Wat. Res., 46,6, 1585-1624. Poch, M., Comas, J., Rodríguez-Roda, I., Sànchez-Marrè, M., Cortés, U., 2004. Designing and building real environmental decision support systems. Environ. Modell. Softw 19, 857-873.

Porro, J., Milleri, C., Comas, J., Rodriguez-Roda, I., Pijuan, M., Corominas, L., Guo, L., Daelman, M., Volcke, E.I.P., van Loosdrecht, M.C.M., Vanrolleghem, P.A., Nopens, I. (Accepted). Risk assessment modelling of nitrous oxide in activated sludge systems: Quality not Quantity. 4th IWA/WEF Wastewater Treatment Modelling Seminar, WWTmod 2014, Spa, Belgium, 30 March – 3 April 2014.

Schulthess, R.V., Wild, D., Gujer, W., 1994. Nitric and nitrous oxides from denitrifying activated sludge at low oxygen concentration. Water Science and Technology 30 (6), 123–132.

Serra P., Sánchez M., Lafuente J., Cortés U., Poch M., 1994. DEPUR: A knowledge-based tool for wastewater treatment plants. Engineering Applications of Artificial Intelligence. 7, 1-23.