Advanced data management for optimising the operation of a full-scale WWTP

Sergio Beltrán, Mikel Maiza, Alejandro de la Sota, José María Villanueva and Eduardo Ayesa

ABSTRACT

The lack of appropriate data management tools is presently a limiting factor for a broader implementation and a more efficient use of sensors and analysers, monitoring systems and process controllers in wastewater treatment plants (WWTPs). This paper presents a technical solution for advanced data management of a full-scale WWTP. The solution is based on an efficient and intelligent use of the plant data by a standard centralisation of the heterogeneous data acquired from different sources, effective data processing to extract adequate information, and a straightforward connection to other emerging tools focused on the operational optimisation of the plant such as advanced monitoring and control or dynamic simulators. A pilot study of the advanced data manager tool was designed and implemented in the Galindo-Bilbao WWTP. The results of the pilot study showed its potential for agile and intelligent plant data management by generating new enriched information combining data from different plant sources, facilitating the connection of operational support systems, and developing automatic plots and trends of simulated results and actual data for plant performance and diagnosis.

Key words | advanced monitoring, data management, data to information, enriched information, managing utilities, WWTP

Sergio Beltrán (corresponding author)
Mikel Maiza
Eduardo Ayesa
CEIT and Tecnun (University of Navarra),
Manuel de Lardizábal 15,
20018 San Sebastián,
Spain
E-mail: *sbeltran@ceit.es*

Alejandro de la Sota
José María Villanueva
Consorcio de Aguas Bilbao Bizkaia,
Calle Maestro José s/n,
48910 Sestao,
Spain

INTRODUCTION

New and advanced management facilities for wastewater treatment plants (WWTPs) have been already designed and implemented over recent years. Their main objective is the plant achieving its optimal operation, considering water quality and energy sustainability. For this ambitious objective, the synthesis of process information is the crucial point. However, the lack of appropriate data management tools is presently a limiting factor for a broader implementation and a more efficient use of sensors and analysers, monitoring systems and process controllers in WWTPs. The real challenge is synthesising information from the processing and analysis of combined data gathered from several interrelated processes (usually measured on-line, analysed in the laboratory and provided by plant operators), dealing with their heterogeneous characteristics and storage decentralisation. In addition, the vast amount of data collected is not always reliable due to the fact that they are frequently incorrect or incomplete. Therefore, the overload of decentralised data, their heterogeneity and unreliability is demanding an effort to synthesise key-process indicators and reliable data for monitoring, diagnosis, operation and control.

This paper proposes a technical solution for the advanced data management of a full-scale WWTP. This software (SW) tool is based on a database which centralises and stores in a standard way the different kind of data available in a plant (e.g. other databases –DBs, programmable logic controllers –PLCs, on-line data analysers and sensors, lab samples). In addition, the database includes a set of algorithms and tools for structuring, filtering, reconstructing and combining all the existing data. The tool has been adapted for a real plant and integrated with three advanced plant operation support systems. This tool is the first step of a research line aimed at optimising the operation of wastewater treatment systems by adequately managing all the available information.

doi: 10.2166/wst.2012.175

THE ADVANCED DATA MANAGER TOOL

The proposed technical solution addresses the design, implementation and validation of a set of algorithms and tools for optimising the operation of WWTPs based on an efficient and intelligent use of data. The Advanced Data Manager tool plays two main roles. Firstly, the role of processing the experimental data acquired from different sources by performing sampling, filtering, fault and/or inconsistency detection and reconstruction, and storing it in a centralised manner. And secondly, that of extracting synthesised information or generating new, non-measurable information. In order to achieve this objective, several basic signal processing algorithms for data filtering and data reconstruction (Montgomery 2009) have been already implemented. The Advanced Data Manager tool provides standardisation and condensation of experimental data which can be directly used by plant operators for assessing the overall status of the system or extracting relevant information for its follow-up and is, therefore, an important plant operation support tool by itself. However, its utility in this research work can go further, by providing pre-processed, adequate and condensed information for the integration with complementary advanced algorithms and tools aimed to optimally operate a WWTP.

Figure 1 summarises the conceptual idea of the Advanced Data Manager tool, its SW architecture, the main SW components, functionalities already implemented and some other that can be potentially implemented. Thus, it can be seen that the advanced data manager tool is supported on a Central DB which comprises all the data sources coming from different points of the plant (PLCs, laboratory analysis, supervisory control and data acquisition systems –SCADAs), properly collected and centralised by the Data Collection and Centralization component through the corresponding interfaces. These raw data, stored according to a given data structure, can then go through a set of processing and adequation algorithms (filtering, fault detection, reconstruction), if required, implemented by the Data Processing and Adequation component. A fourth component (Enriched Information Provider) is implemented by the Advanced Data Manager tool which allows extracting and generating enriched information, mainly combining on-line and off-line (laboratory samples) data and automatically monitoring it. All the information, from the raw data to the enriched information, is duly checked and validated, reliable and complete (and condensed) and is made readily accessible, through a common access point and the corresponding interfaces, to operators and/or third parties' SW applications. The integration with third party SW applications being able to extract information in an easy and agile manner is one of the potential functionalities provided by the advanced data manager tool. Advanced Monitoring Systems, Model-Based Processes Simulation and Advanced Controllers are proven powerful plant operation support tools whose efficiency very much depends on the consistency of the input data, which can be now be achieved by interfacing with the advanced data manager tool in an integrated manner. A full-scale example of this is herein given whereby the integration of an advanced data manager with a solids inventory observer, a plant simulator and a supervisory controller is implemented.

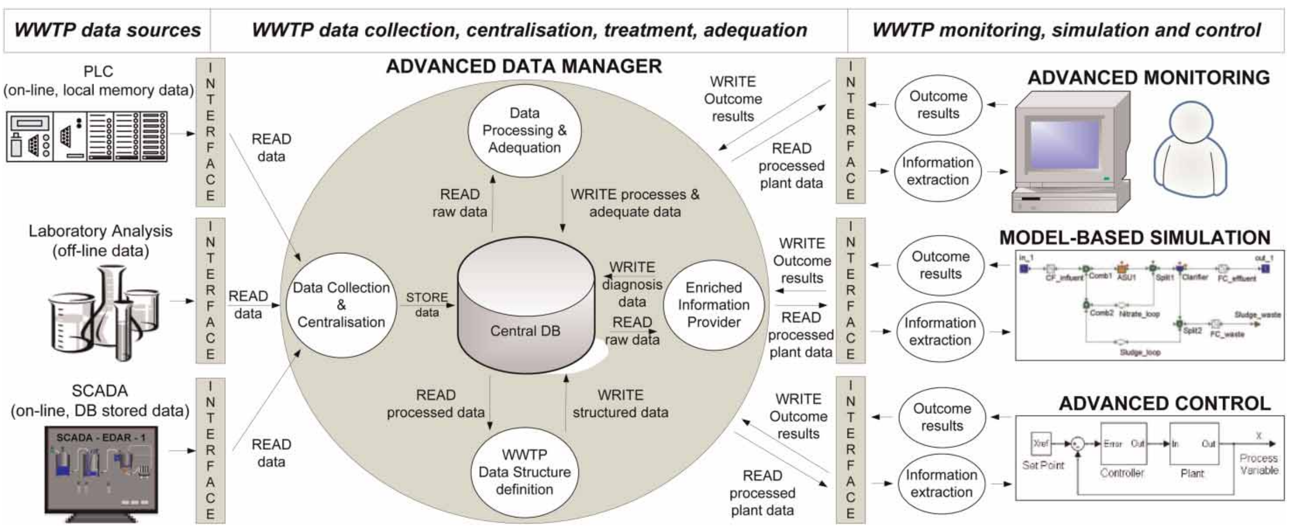

Figure 1 | The Advanced Data Manager tool and its architecture.

Next, the adaptation and application of the Advanced Data Manager tool to a real plant and its integration with some given plant monitoring, simulation and automatic control systems is presented. Finally, although the Advanced Data Manager tool is potentially able to deal with other technologies (e.g. (sequential) batch reactors, fluidised and fixed bed reactors, etc.), the technical solution must be specifically tailored in order to be implemented.

THE GALINDO-BILBAO ADVANCED DATA MANAGER TOOL

This is the case of the Galindo-Bilbao WWTP, whose managers have clearly shown their concern about appropriate plant data management and have given support to this work by technically collaborating and economically financing the development of the Galindo-Bilbao Advanced Data Manager tool. The Galindo-Bilbao WWTP was designed for the carbon and nitrogen removal of an urban wastewater flow of 345.600 m^3/d. A more complete description of the plant configuration and reactors dimensioning can be found in Galarza *et al.* (2001) and Rivas *et al.* (2001).

The Galindo-Bilbao Advanced Data Manager tool has been divided into two graphical user interface applications, which are respectively used to design the plant data-model and monitor raw plant data, as well as processed and synthesised plant data (Figure 2).

The Administrator application allows definition of the data-model/structure according to which the data will be acquired, stored and created into a centralised DB. This DB is plant-specific, but implemented according to an adaptation of the standard advanced data manager DB. The Client application allows monitoring, according to the data-model defined by the Administrator tool and/or to the data structure of the source data, raw data, filtered data, reconstructed data and calculated data. Simple and conventional tools for data filtering (low-pass filters, elimination of negative and zero values) and data reconstruction (interpolation and extrapolation of blanks) are integrated here. The interfacing with other operation support tools such as advanced monitoring and control systems and dynamic simulators is also implemented.

Next, the integration of the Galindo-Bilbao Advanced Data Manager tool with three SW applications designed and developed for improving the operation of the plant (advanced monitoring, dynamic plant simulation and advanced control) is presented. Figure 3 shows the conceptual idea of the architecture of the Galindo-Bilbao Advanced Data Manager tool, adapted from the standard Advanced Data Manager tool (Figure 1).

Advanced monitoring: a solids inventory observer integrated with the Advanced Data Manager tool

The secondary treatment of the Galindo-Bilbao WWTP is composed of six parallel lines, each of them in a regeneration–denitrification–nitrification configuration. This biological configuration shows a significant solids gradient between the regeneration and denitrification–nitrification

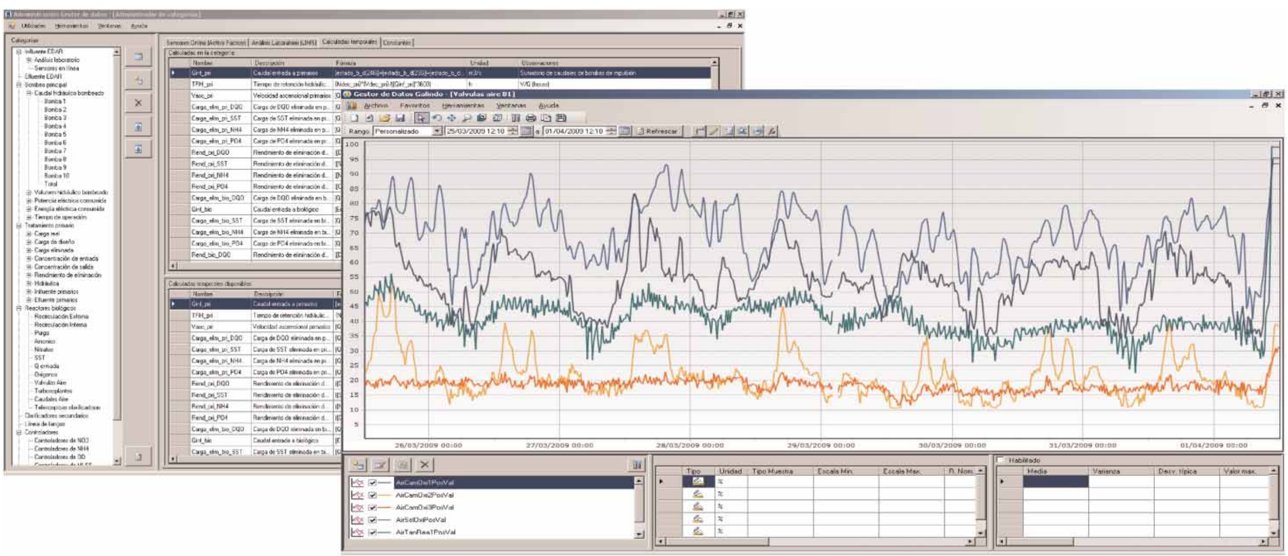

Figure 2 | The Galindo-Bilbao Advanced Data Manager Tool: visual appearance of the Administrator and Client applications.

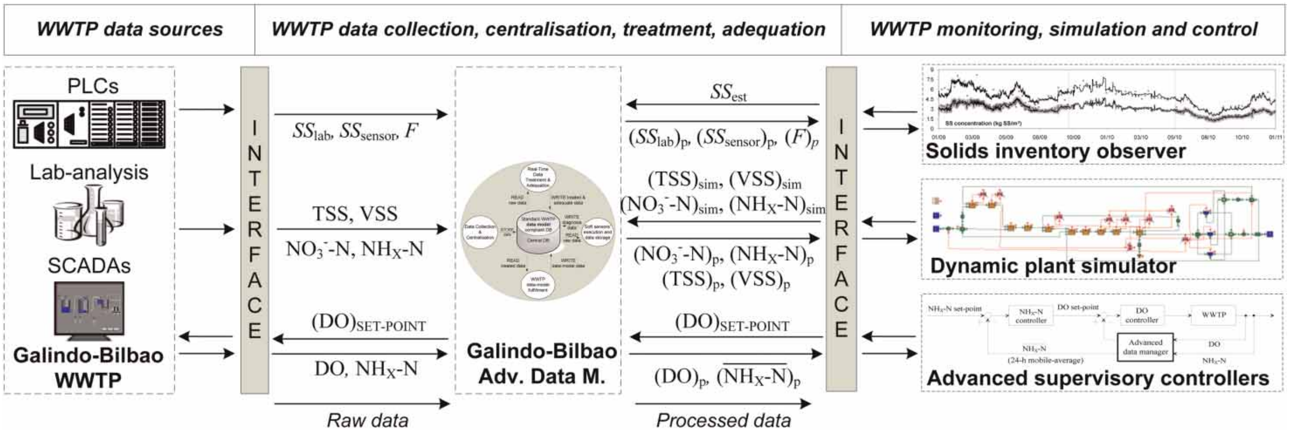

Figure 3 | The Galindo-Bilbao Advanced Data Manager tool and its architecture.

reactors. In addition, the hydraulic dynamics of the plant provoke significant variations in the solids concentrations. It is important to note that the total amount of solids in WWTPs and their distribution among the different elements and lines play a crucial role in the stability, performance and operational costs of the process. Thus, considering the biological configuration and the hydraulic dynamics, an appropriate estimation of the solids concentration at each point of the plant is especially interesting for operational purposes as a closely approximation of biomass. However, on-line solids concentrations are only measured in the nitrification reactors. For the on-line estimation of the solids concentration in all the of the plant's tank reactors, a state observer was designed, developed and integrated within the Galindo-Bilbao Advanced Data Manager tool. A specific estimation of the active biomass within the reactors (Ayesa *et al.* 1991) would be very useful for plant operation but, from the experience of the users over several years, this is still not practically attainable with the incomplete and heterogeneous data existing in full-scale plants. In fact, the development and full-scale validation of Data Manager tools are the first steps to cope with this ambitious objective.

The state observer has been based on a recursive 'prediction–correction' structure (Kalman Filter algorithm, Bar-Shalom & Li 1993) with a simplified hydraulic description of the plant lines as in-series tank reactors (Beltrán *et al.* 2009). In the predictive step, the observer dynamically estimates the suspended solids (SS_{est}) distribution among the reactors (regulated by faster dynamics) using the information provided by flow-rate measurements (F). On the other hand, in the corrective step, the observer obtains updated experimental information about the total amount of solids in the system (regulated by relatively slow dynamics) from both analytical measurements of total suspended solids SS_{lab} and on-line suspended solids sensors SS_{sensor} in some of the tanks. Previously, those measurements have been processed by filtering, fault detection and simple reconstruction algorithms (interpolation methods), leading to $(F)_p$, $(SS_{lab})_p$ and $(SS_{sensor})_p$, respectively. Thus, an on-line estimation of the solids concentrations at any point in the plant is given by using the information supplied by solids measurements in just some of the tanks and flow-rate measurements.

Figure 4 shows the suspended solids concentration in the regeneration tank (above) and in the nitrification tank

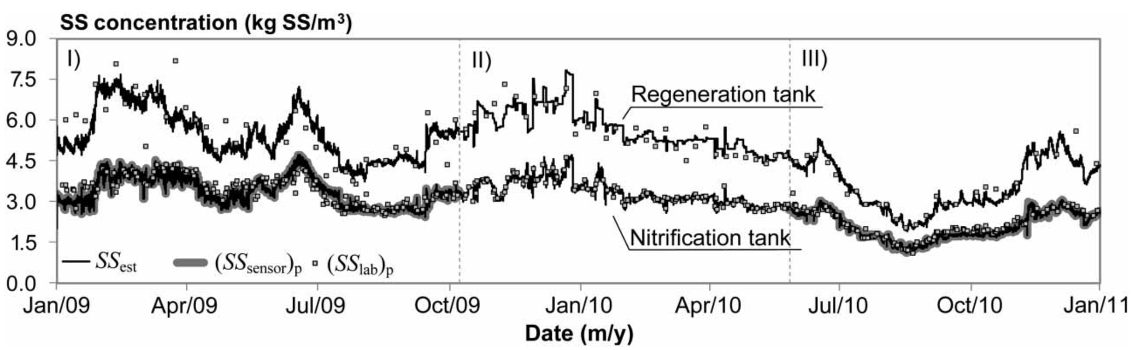

Figure 4 | On-line estimation of the solids concentration in the Galindo-Bilbao WWTP during a 2-year period.

(below) for one of six parallel lines during 2 years of plant operation. The suspended solids concentration is measured in the regeneration tank by only lab samples (2540 D, Standard Methods 2005), and in the nitrification tank also by an on-line sensor. During time periods (I) and (III), the laboratory samples and the on-line measurements of the nitrification tank were used for the estimation of the suspended solids concentrations in both tanks. In these periods, the state observer was able to reasonably estimate the suspended solids cosncentrations in the nitrification reactor and in the regeneration reactor. Moreover, in time period (II), even though the on-line suspended solids sensor of the nitrification reactor was unavailable, the state observer provided a reasonable on-line estimation of the suspended solids in both tanks by using only the lab samples in the nitrification tank. The estimation results (Figure 4) corroborated the capacity of the Kalman filter for an appropriate estimation of SS at different points of the plant. It is interesting to remark that the Kalman filter approaches are especially suited for on-line estimation when the mechanisms described by the model predictions are relatively known (Bastin & Dochain 1990) as, for example, the dynamic solid flux among stirred tanks.

The solids inventory observer integrated with the Galindo-Bilbao Advanced Data Manager tool showed its capacity for the dynamic estimation of the solids concentration combining data from on-line measurements and lab-analysis. In this respect, it must be noted that flow-rate measurements were considered to be known. To take into account uncertainty in flow measurements, a possible solution is the inclusion in the on-line estimation of both suspended solids concentrations and all the flows within the plant. This way, the resulting estimation will estimate simultaneously the hydraulic pattern and the solids' flux within the plant. This combined estimation of hydraulic pattern and solids' distribution is currently under development.

Model-based simulation: a dynamic plant simulator integrated with the Advanced Data Manager tool

Dynamic plant simulators need as input data time-discrete values from different sources with dissimilar time bases, such as on-line flow-rates and lab samples of nutrient concentrations. Thus, a clear example of processing heterogeneous data and extracting automatic synthesised information is the integration of a dynamic plant simulator with the Galindo-Bilbao Advanced Data Manager tool.

A mathematical model of the Galindo-Bilbao WWTP was designed based on the Activated Sludge Model n° 1 (Henze *et al.* 2000), developed in the simulation platform WEST® (mikebydhi.com), experimentally calibrated and validated beforehand (Grau *et al.* 2007), and integrated within the Galindo-Bilbao Advanced Data Manager tool as an optimisation tool to support both the decision-making and performance evaluation and diagnosis of the plant. Using time-discrete data provided by the Advanced Data Manager tool, the computer simulation of the model was used on an ongoing basis by successive interaction between predicting responses and monitoring of actual system responses by developing automatic plots and trends of simulated results and actual measurements. As an example, Figure 5 shows two plots generated automatically of a simulation study of 1 year of plant operation. The first plot (above) shows the simulation of the effluent nitrates $(NO_3^-\text{-}N)_{sim}$ and 24 h average ammonium $(NH_X\text{-}N)_{sim}$ concentrations (by continuous lines) compared with their processed actual laboratory samples (by dots), $(NO_3^-\text{-}N)_p$ and $(NH_X\text{-}N)_p$, respectively. The second plot (below) shows

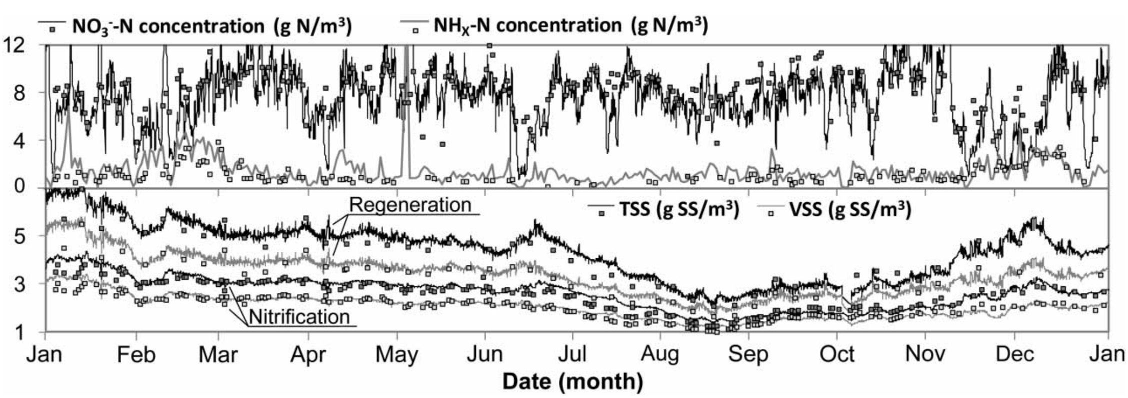

Figure 5 | Simulation study of effluent nitrates and ammonium and mixed liquor solids concentration.

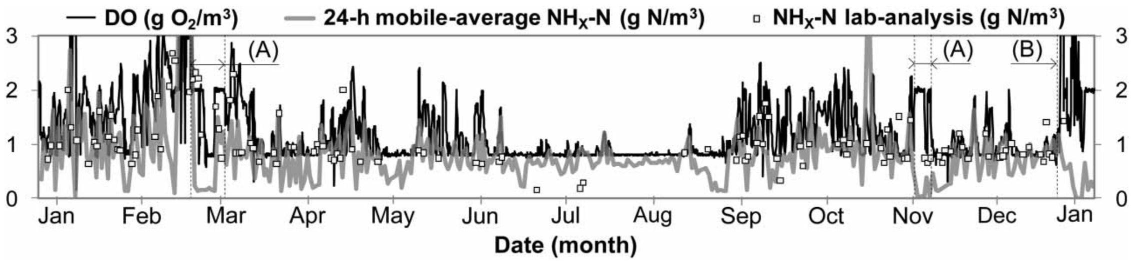

Figure 6 | Aeration control performance in the Galindo-Bilbao WWTP.

the simulation of the total and volatile suspended solids concentration in the regeneration tanks and in the nitrification tanks (by continuous lines), $(TSS)_{sim}$ and $(VSS)_{sim}$, compared with their actual laboratory samples being previously processed (by dots), $(TSS)_p$ and $(VSS)_p$, respectively.

It is interesting to note how the dynamic plant simulator was directly connected to the Galindo-Bilbao Advanced Data Manager tool and showed its usefulness for plant performance evaluation and diagnosis, providing an automatic generation of plots and trends from different data sources.

Advanced control: a supervisory controller integrated with the Advanced Data Manager tool

Important enriched information for monitoring, operation and control of WWTPs often have to be obtained from the processing of on-line signals (Irizar *et al*. 2008). For example, in the case of control purposes, raw signals from sensors and analysers have to be processed before being used by automatic controllers. In this way, an on-line management of raw plant signals from sensors and analysers for automatic control is accomplished by the Galindo-Bilbao Advanced Data Manager tool. The processes take into account sampling frequency, filtering, time-frame specified for every control objective and technical characteristics of every sensor such as time responses and accuracy. A clear example of on-line data management for control purposes is the information required by the aeration controller of the secondary treatment reactors in the plant.

The overall aeration control architecture of each line of the Galindo-Bilbao WWTP consists of two SISO PID loops in cascade where the first controller dynamically determines the dissolved oxygen set-point $(DO)_{SET\text{-}POINT}$ in the aerated basins with the objective of maintaining the 24 h-average concentration of ammonium in the effluent (NH_X-N controller), and the second controller sets the DO value in each basin to the required DO set-point (DO controller) (Ayesa *et al*. 2006).

In this plant, as the quality of the effluent is defined in terms of daily average samples, for the measurements of effluent NH_X-N, a time-horizon of 24 h is defined, while current NH_X-N analysers need about 10–15 min to get new measurements. Thus, the calculated NH_X-N 24 h moving-average values $(\overline{NH_X\text{-}N})_p$, provide more useful information for monitoring and control. Finally, the DO controller does not require a post-processing of the DO measurements and the only prerequisite is to remove the high frequency noise from the raw sampled data, resulting in the processed signal $(DO)_p$. Figure 6 shows the experimental results of DO processed measurements in the aerobic basins and the 24 h mobile-average value of effluent NH_X-N used for the aeration control of 1 year of operation. Neglecting maintenance periods (A) and idle air valve problems (B), the data show how the supervisory control moves the DO concentration according to the averaged effluent quality, thus, producing a net reduction in air flow and energy consumption during low-load periods while ensuring complete nitrification during high-load periods.

As a result, the advanced supervisory controller was directly connected to the Galindo-Bilbao Advanced Data Manager tool and showed its ability to process raw NH_X-N and DO measurements for optimising effluent quality and energy consumption under medium and long-term perturbations. With the presented methodology the Advanced Data Manager tool would be able to adapt other control algorithms, although the new solution should be specifically tailored in order to be implemented.

CONCLUSIONS

This paper proposes a technical solution for the advanced data management of a full-scale WWTP. The solution is based on an efficient and intelligent use of plant data by a standard centralisation of the heterogeneous data acquired from different sources, an effective data processing to extract adequate information, and a straightforward connection to other emerging tools focused on the operational optimisation of the plant such as advanced monitoring and control systems or dynamic simulators. A pilot study of the Advanced Data Manager at full-scale was designed and implemented in the Galindo-Bilbao WWTP.

The results obtained from the integration with the Galindo-Bilbao Advanced Data Manager tool with three operation support systems demonstrated its potential for agile and intelligent data management of WWTPs through the implementation of data standard centralisation, effective data processing and easy connection to support systems:

- The solids inventory observer integrated with the Galindo-Bilbao Advanced Data Manager tool showed its capacity for dynamic estimation of solids concentration combining data from on-line measurements and lab-analysis.
- The dynamic plant simulator integrated with the Galindo-Bilbao Advanced Data Manager tool demonstrated its usefulness for the agile extraction of synthesised information from heterogeneous data, generating automatic plots and trends.
- The advanced supervisory controller integrated with the Galindo-Bilbao Advanced Data Manager tool showed its ability to enrich data (e.g. effluent NH_X-N mobile-average values) for optimising plant operation under medium and long-term perturbations.

Finally, the flexibility and adaptability of the proposed technical solution will allow further implementation of advanced data management algorithms and tools, the integration with third party wastewater treatment operation SW tools such as advanced monitoring and control systems and the extension to sewer systems or the impact on water receptors.

ACKNOWLEDGEMENT

The authors would like to thank the financial support from the Spanish Ministry of Science and Innovation (NOVEDAR_Consolider CDS2007–00055 and OPTIMAR PSE-310000-2009-0005) and the water authority Consorcio de Aguas Bilbao Bizkaia. Thanks are also due to Alain Castro (CEIT) for his technical collaboration in this R&D project.

REFERENCES

Ayesa, E., Flórez, J., García de las Heras, J. L. & Larrea, L. 1991 State and coefficients estimation for the activated sludge process using a modified Kalman filter algorithm. *Water Sci. Technol.* **24** (6), 235–247.

Ayesa, E., De la Sota, A., Grau, P., Sagarna, J. M., Salterain, A. & Suescun, J. 2006 Supervisory control strategies for the new Galindo-Bilbao WWTP: the long run from the conceptual design to the full-scale experimental validation. *Water Sci. Technol.* **53** (4–5), 193–201.

Bar-Shalom, Y. & Li, X.-R. 1993 *Estimation and Tracking: Principles, Techniques, and Software*. Artech House, Boston.

Bastin, G. & Dochain, D. 1990 *On-line Estimation and Adaptive Control of Bioreactors*. Elsevier, Amsterdam.

Beltrán, S., Irizar, I., Monclús, H, Rodríguez-Roda, I. & Ayesa, E. 2009 On-line estimation of suspended solids in biological reactors of WWTPs using a Kalman observer. *Water Sci. Technol.* **60** (3), 567–574.

Galarza, A., Ayesa, E., Linaza, M. T., Rivas, A. & Salterain, A. 2001 Application of mathematical tools to improve the design and operation of activated sludge plants. Case study: the new WWTP of Galindo-Bilbao Part II: operational strategies and automatic controllers. *Water Sci. Technol.* **43** (7), 167–174.

Grau, P., Beltrán, S., de Gracia, M. & Ayesa, E. 2007 New mathematical procedure for the automatic estimation of influent characteristics in WWTPs. *Water Sci. Technol.* **56** (8), 95–106.

Henze, M., Gujer, W., Mino, T. & van Loosdrecht, M. C. M. 2000 *Activated Sludge Models: ASM1, ASM2, ASM2d and ASM3*. Scientific and Technical Report No. 9. IWA Publishing, London.

Irizar, I., Alferes, J., Larrea, L. & Ayesa, E. 2008 Standard signal processing using enriched sensor information for WWTP monitoring and control. *Water Sci. Technol.* **57** (7), 1053–1060.

Montgomery, D. C. 2009 *Introduction to Statistical Quality Control*. John Wiley & Sons, New York.

Rivas, A., Ayesa, E., Galarza, A. & Salterain, A. 2001 Application of mathematical tools to improve the design and operation of activated sludge plants. Case study: the new WWTP of Galindo-Bilbao Part I: Optimum design. *Water Sci. Technol.* **43** (7), 157–165.

Standard Methods 2005 *Standard Methods for the Examination of Water and Wastewater*, 21st edition. American Public Health Association/American Water Works Association/Water Environment Federation, Washington DC, USA.

First received 3 August 2011; accepted in revised form 24 February 2012