1. FINAL PUBLISHABLE SUMMARY REPORT

1.1. Executive summary

Wastewater treatment plants (WWTPs) deal today with extremely large volumes of heterogeneous, redundant, incorrect and sometimes incoherent data that come from very different sources (online sensors and analysers, laboratory, manually collected, etc.), which makes it difficult for operators to know the plant status, and to close the loop with advanced controllers. This way, the lack of appropriate data management is one of the main bottlenecks for a more efficient use of sensors, monitoring systems and process controllers in WWTPs.

The objective of the DIAMOND project is to optimise the global operation of WWTPs (in terms of water quality, resources and energy) by achieving a model of excellence in the management of all the available data. This has been done with the design, development, implementation, testing and validation at three full scale WWTPs of four exploitable systems that are now technically ready to be commercialized:

(1) A new advanced data management platform (the ADAM Tool), which accesses, samples, acquires and stores in a standard way the experimental data, centralising them into a single Central Database; validates the stored data, guaranteeing that the data is reliable, coherent, fault free and complete; and syntheses useful information and generates plant-wide key performance indicators from the stored, reliable data. The ADAM tool is supported by:

(2) The Online Data Pre-processing Tool (the ODP Tool), which provides plant operators with information about single sensor diagnostics and alarms activated by faulty data; and assesses the quality signal of several sensors using information from all of them at the same time.

(3) The Advanced Monitoring System (the AMS), which monitors historical and real-time measured data and synthesised new useful data (including KPIs) that are necessary for showing plant status; and provides the future plant status based on current plant conditions, making easier and more efficient the operational decision-making processes.

(4) The Advanced Control System (the ACS), which defines the optimal control strategy in terms of nitrogen removal efficiency, while minimizing the operational costs of the plant; and prepares the plant data for the Optimal Process Control Module and defines the process supervision tools.

The results obtained at the Mekolalde WWTP (Spain), the Bromma WWTP (Sweden) and the Kymen Vesi Oy Mussalo WWTP (Finland) demonstrate that the DIAMOND Systems are a solution for the advanced data management in WWTPs, allowing plant operators to optimise the global operation of the plants in terms of water quality, resources and energy. By achieving a model of excellence in the management of all the available data, it has been shown that with DIAMOND Systems the lack of appropriate data management is no longer a bottleneck for a more efficient use of sensors, monitoring systems and process controllers in WWTPs.

Due to the significance of the results, the DIAMOND project and its results have been effectively and widely disseminated to a broad public through different channels and languages: a DIAMOND web page, 39 pieces of news, 7 printed and electronic press releases, 5 publications, 13 participations in industrial and academic seminars, conferences and presentations, a video clip and a Wikipedia webpage.

1.2. Summary description of project context and objectives

Nowadays most EU countries treat practically the totality of the generated volume of urban wastewater. However, an in depth look at the performance of most of the existing wastewater treatment plants (WWTPs) evidences that significant improvements can still be made in terms of effluent quality, process robustness and operational costs. More efficient operation of WWTPs could significantly improve the stability of the process and the quality of the effluent, while at the same time achieving a reduction in operational costs, with a special impact on energy consumption and recovery.

In this respect, the new and more stringent quality requirements for water treatment systems have led to major advances in WWTPs: on the one hand, new treatment technologies based on more and more complex plant configurations have emerged; on the other hand, online instrumentation has experienced significant improvement.

In fact, the above advances have brought about important changes with respect to the availability of data in WWTPs. Some years ago one of the main technical limitations to optimising plant operation was the scarcity of data and therefore the very limited information about a plant, mainly due to the lack of reliable measuring devices. However, WWTPs now have to deal with a very large volume of heterogeneous data separately collected from different sources at the plant. Data logging and supervisory control and data acquisition software tools used by WWTPs on a daily basis manage thousands of points of data from all sources of the plant, which means the likelihood that errors and/or faults are present in the data has greatly increased. The appropriate processing and proper management of all these heterogeneous, incomplete and frequently inconsistent data points normally beyond the capacity WWTP staffs. Consequently, although now it for reasons of excess rather than shortage, access to valuable plant information for diagnosis and carrying out operation activities continues to be limited.

In conclusion, in order to arrive at the optimum operation of WWTPs, there is presently a clear and urgent need for advanced data management algorithms and tools.

1.2.1. Objectives

The overall objective of the project is:

"to optimize the global operation of wastewater systems by adequately managing and using all the information available in the plant at every moment".

It is well known that process monitoring systems and automatic controllers play a very important role in wastewater treatment plants (WWTPs). On the one hand, process monitoring software tools are in general meant to facilitate the daily decision making of the WWTP operators, where the information needed by these monitoring systems mostly comes from heterogeneous data sets made up by various sources of data in the plant (online sensors and analysers, laboratory analysis, programmable logic controllers, etc.). On the other hand, process controllers become a must when complex decisions must be based on a large number of operational variables in order to maintain the optimum operational point of a plant.

However, the successful implementation of existing process monitoring software tools and control systems very much depends on the reliability and completeness of the data collected, the quality of the information extracted and how quickly and easily it is accessed. On the one hand, although online sensors and analysers as well as data acquisition and data base systems have greatly evolved, they do not guarantee reliable data, high quality information and easy access: sensors have failures, delays, drift and noise, and data acquisition and data base systems frequently are incompatible.

On the other hand, not all the data are measured online; that is, some of the data used to analyse the behaviour of a WWTP can only be obtained in the laboratory. This is the case, for example, with some crucial parameters such as COD, biodegradable COD, nitrogen, or phosphorous: due to the lack of appropriate online sensors for measuring these parameters, there is no other solution than analysing them in the laboratory. This obliges WWTPs to use sparse and delayed laboratory results (some laboratory experimental data might not be available for days or weeks).

These limitations greatly complicate proficient process online monitoring and control, and thus optimally operating the plant is complicated; that is, the plant might continue working at a conservative, standard operational point but without exploiting all of its potential in terms of effluent quality, process robustness and operational costs. In fact, many automatic control strategies that have shown their usefulness at simulation scale, or even pilot plant scale, are experiencing a limited full-scale plant applicability because they are particularly sensitive to the quality of the experimental data needed (errors, failures, noise, delays, loss of data, etc.), which means they should incorporate data pretreatment algorithms that are capable of analysing the measures be-fore sending them back to the controller.

In other cases, especially for supervisory control loops, optimum utilization of the controllers requires information from different elements of the system and not only from those local elements on which they are directly acting; for example, controlling a sludge line of a given WWTP by taking into account the status of the water line, and vice versa, will be much more efficient and optimum than designing a controller using only local information (sludge line-related or water line related, but not both), which re-quires advanced controllers to have common access to reliable and complete data not only regarding the status of the system/process that they are acting on but also regarding the status of adjacent and interrelated systems/processes.

In this respect, DIAMOND addresses the development of a software platform (the ADAM Tool) in order to overcome all these drawbacks and limitations providing duly processed, condensed and enriched, reliable and high quality plant data of interest, as well as straight forward and immediate access to valuable supplementary information about the status of other elements of the WWTP. The ADAM Tool, together with all of its auxiliary algorithms and tools, facilitates the successful implementation of process monitoring and automatic control systems for existing and future wastewater treatment systems, thus allowing the main objective of this project to be met, which is optimising the global operation of WWTPs.

1.2.2. Partial objectives

To meet the main objective, the project addresses four main secondary objectives:

1. The design, development, implementation and validation of a new advanced data management platform (**The ADAM Tool**), which allows for the centralisation and processing of all the heterogeneous, incomplete and frequently inconsistent data of WWTPs, as well as the extraction of valuable and validated information about the status of the entire plant.

The ADAM Tool plays the three important roles:

- Firstly, it allows for the sampling/selecting of relevant experimental data (online + offline) from different sources within a plant, as well as the centralisation of data into a Central database (DB).

- Secondly, it allows for the validation of the data stored in the Central DB by properly filtering, reconciling, correcting and completing it.

- Thirdly, it allows for the extraction of synthesized and enriched information or the creation of new, non-measureable plant-wide information from the existing, measurable data already checked and validated.

In order to do all the above, among other things, the ADAM Tool incorporates signal processing, data reconciliation and data reconstruction algorithms and tools capable of checking the coherence of the data, reconstructing incomplete signals and building new information which effectively condenses the experimental data so that they are useful for the plant operation. The ADAM Tool implements the above algorithms and tools in the following modules:

(a) Data Collection & Centralisation Module: this consists of collecting (sampling/selecting) all the raw data available (online + offline) in the WWTP and storing them in a Central Data Base (DB). This central DB is designed in order to give easy and direct access to all plant data for any SW system.

(b) Data Validation Module: its main functionalities are to guarantee the reliability and high quality of the collected plant data. It mainly consists of processing the data stored by the Data Collection & Centralisation module, checking the coherence of the data (data correlation) and, correcting and/or reconstructing the data when applicable.

(c) Synthesised & Enriched Information Generation Module: its main functionality is to provide online plant status information by extracting valuable, synthesised and enriched, plant-wide information and generating new, non-measurable information from the measurable data that has already been stored, checked and validated by the other two (Data Collection & Centralisation Module and Data Validation Module) software modules.

2. The ADAM Tool is primarily supported by **The Online Data Pre-processing Tool**, which consists of implementing Sensor Data Models. This are software- or sensor-implemented models that estimate the status of and/or data from sensors by using fault detection and model-based prediction algorithms and tools, thus providing the ADAM Tool with reliable and robust online data.

These tools facilitate monitoring and decision-making processes for operators, as well as the implementation of plant-wide automatic controllers. In order to prove this, two additional advanced plant operation support systems are developed:

3. **The Advanced Monitoring System**, which consists of the design, implementation (calculation and storage) and monitoring of plant-wide Key Performance Indicators (KPIs) based on the information provided by the ADAM Tool and are necessary and sufficient for indicating the status of different aspects of a WWTP, such as treatment performance and treated water quality rates, energetic performance and generated energy rates (when applicable), economic performance, environmental impact, and so on.

The Advanced Monitoring System, like the ADAM Tool, has a direct utility in of itself as it provides and monitors overall plant status information for plant operators and their decision-making processes. Additionally, it is complementary with and integral to the ADAM Tool and the Advanced Control System: i) on the one hand, all the information needed by this Advanced Monitoring System must come from experimental data that has been efficiently collected, processed and validated by the ADAM Tool; ii) on the other hand, in addition to monitoring, the Advanced Monitoring System stores the calculated KPIs in the Central DB of the ADAM Tool so they can be used by other plant operation support systems.

4. **The Advanced Control System**, which consists of the design, development and implementation of plant-wide control algorithms based on the information provided by the ADAM Tool (duly processed and condensed in order to guarantee the quality of the feedback data). It is aimed at optimising the operation of the plant according to certain economic and/or environmental criteria, for example energy consumption minimisation and/or energy generation maximization.

Although the Advanced Monitoring System proposed above is designed for facilitating the daily decision-making of WWTP operators, wastewater treatment systems are systems of growing complexity, where the number of operational variables gradually increases. Consequently, an increasing number of decisions needs to be taken in order to keep the plant at its optimum operational point.

For this reason, it seems logical to leave the strictly necessary aspects of operational decisions (duly supported by the Advanced Monitoring System) to plant operators and to try to carry out the remaining actions according to properly designed and automatically controlled operational criteria.

The latter type of actions will be under the purview of the Advanced Control System. Similar to the ADAM Tool and the Advanced Monitoring System, it has a direct utility in of itself, and it is complementary with and integral to the ADAM Tool and the Advanced Monitoring System as it makes use of the efficiently collected, processed and validated experimental data, together with the condensed, enriched and/or newly constructed information from the ADAM Tool, in addition to the KPIs calculated by the Advanced Monitoring System. Analogously, the Advanced Control System has an area available for storing controller data (set-points, parameters, outputs) in the Central DB of the ADAM Tool so that these data can be monitored or used for any other purpose.

The entire SW framework is a modular SW solution where every software module is independent and has a utility in of itself, and at the same time, all software modules are complementary and integral to each other.

1.3. Description of the main scientific & technical results/foregrounds

1.3.1. Technical solution

The proposed technical solution is based on the development, testing and validation of a set of advanced data management tools to optimize the operation of WWTPs (see **Figure 1**):

The ADAM Tool implements the following modules: a Data Collection & Centralisation Module for sampling relevant experimental data (online + offline) from the plant and centralising them into a Central Database; a Data Validation Module for validating the data stored in the Central Database; and a Synthesised & Enriched Information Generation Module for extracting synthesized and enriched plant-wide information from the measurable data. The ADAM Tool incorporates signal processing, data reconciliation and data reconstruction algorithms and tools for checking data, reconstructing signals and building new information that is useful for the operation of the plant.

The ADAM tool is supported by three tools:

- The Online Data Pre-processing Tool (ODP Tool) is based on sensor data pre-processing models for calculating sensor diagnostic data; the models use available sensor data and calculate estimated values of the parameters that are not measured.

- The Advanced Monitoring System (AMS) calculates and monitors plant-wide key performance indicators; these are based on the information provided by the ADAM Tool and are necessary and sufficient for indicating WWTP status.

- And the Advanced Control System (ACS) runs plant-wide control algorithms that, based on the information provided by the ADAM Tool, aimed at optimising plant operation according to economic and/or environmental criteria.

The software framework is a modular software solution, where every module has its own utility and all software modules are complementary and integrated.

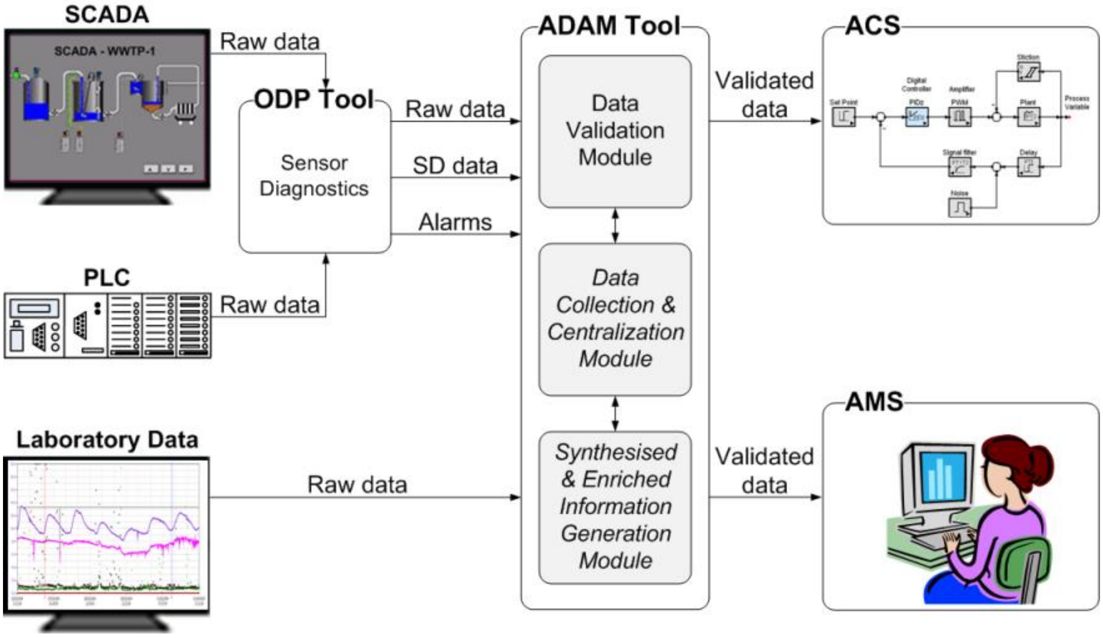


Figure 1 Schematic depiction of DIAMOND project's technical solution

1.3.2. Testing and full-scale validation

DIAMOND partners worked during the first period (from 1st of September of 2012 to the 31st of May of 2013) on the individual design of the tools and the overall design of the integrated solution, based on the interfacing and data exchange of the tools. In parallel, DIAMOND partners carried out data collection activities at the full-scale demonstration plants where project results have been validated. A first final version of the platform was ready by December of 2013.

Subsequent to preparing the first version of the final platform, a definition of a testing and full-scale validation procedure of the tools and algorithms developed in the project was completed. Three full-scale WWTPs were used to test and validate the individual design of the tools and the overall design of the integrated solution (**Figure 2**):

1. The Mekolalde WWTP (Spain), a small to medium size, under-loaded WWTP for treating urban wastewater (7200 m^3 aeration basin), has been the plant used to test and validate the ADAM Tool and the Advanced Monitoring System.

2. The Bromma WWTP (Sweden), a medium size, highly loaded WWTP for treating urban and industrial wastewater (23,400 m^3 aeration basin), has been used to test and validate the Online Data Pre-processing Tool integrated with the ADAM Tool. In addition, the Bromma WWTP has been also used to validate the ADAM Tool and the Advanced Monitoring System.

3. The Kymen Vesi Oy Mussalo WWTP (Finland), the fourth biggest WWTP in Finland (23,800 m^3 aeration basin), has been used to test and validate the Advanced Control System integrated with the ADAM Tool. In addition, the Kymen Vesi Oy Mussalo WWTP has been also used to validate the ADAM Tool and the Advanced Monitoring System.

Figure 2 The Bromma WWTP (left), the Mekolalde WWTP (centre) and the Musslao WWTP (right), plants were DIAMOND Systems have been tested and validated

1.3.3. Scientific & technical results

Due to this project, four systems have been designed, developed, implemented, tested and validated at full-scale. The main scientific and technical results of each product are subsequently analysed.

<u>The Advanced Data Manager Tool (The ADAM Tool)</u>

The ADAM Tool consists of a set of standard, advanced data management algorithms and tools for supporting the optimum operation of WWTPs. This is done by collecting, storing, correcting/completing, validating and serving data to the project's other complementary plant operation optimisation support tools (the Advanced Monitoring System and the Advanced Control System) and/or other third party software systems.

For this, three modules have been developed:

1) The Data Collection & Centralization Module, which includes the algorithms and tools to access, sample, acquire and store in a central database all the data of a WWTP that are needed for monitoring its status. This module also includes the task related with the user administration of the ADAM Tool.

For this, a Central Database has been developed in an instance of SQL Server 2012 running on a Microsoft Windows 2008 R2 64 bit server operating system. The aim of the Central Database is to store the following information: a) the configuration of the data source interfaces; b) the configuration of the raw and enriched variables; c) the information received by the data source interfaces: raw data and reliability data provided by the Data Validation Module and the Online Data Pre-processing Tool; d) several configuration tables for converting between the units provided in the data sources into the stored units; e) the configuration tables with the enriched variable configurations and data; f) the authentication and authorization tables to control which users can access each variable; g) the classification tables which are designed to make it easier to find the stored variables; and h) several tables for the definition of sheets and scoreboards which have been defined in the Advanced Monitoring System. In order to create a standard procedure for assigning names to the tables, fields, triggers and the definition of tables used in the Central Database, a nomenclature has been also created.

The structure of the ADAM Tool software was developed in Microsoft Visual Studio 2010, using .Net framework 4.0 and the C# programming language. Four different layers have been created to develop the ADAM Tool structure (**Figure 3**): a) ADAMDAL: this layer defines the Data Access Layer; b) ADAMBL: this layer defines the business layer, which contains the different logic modules of the ADAM Tool, such as the Data Collection & Centralization Module, the Data Validation Module, the Synthesised & Enriched Information Generation module, as well as other less important modules; c) ADAMEntities: this layer defines the entities or objects that the ADAM Tool shares among the different layers of the application; and d) ADAMWS: this layer defines the interface layer.

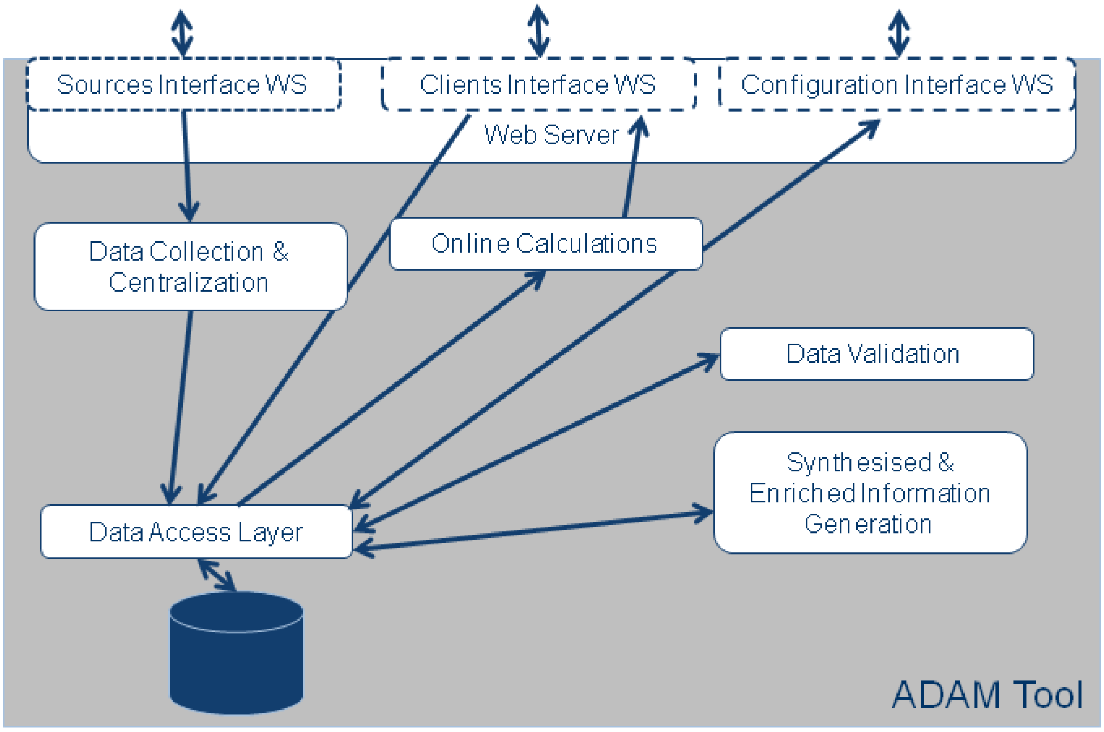


Figure 3 Structure of the ADAM Tool software

In this case, the interface of the ADAM Tool is defined by SOAP Web Services, which are grouped into three web service libraries: 1) Source Web Services, which are used by the interfaces that send the data and information from the WWTPs' data sources to the ADAM Tool. 2) Configuration Web Services, which are used by the ADAM Tool configuration tool when configuring the system. 3) Clients Web Services, which are used by the application's various clients such as the Advanced Monitoring System and the Advanced Control System.

Finally, a Windows service called "ADAM Service" has been created to generate two types of new data: reconstructed data and synthesized and enriched information. This service uses two functions that are developed in the ADAMBL layer and are started every 30 minutes: one for the reconstruction of raw data, and the other one for the generation of synthesized and enriched information.

A Configuration Tool for administrating the ADAM Tool was developed in Microsoft Visual Studio 2010 (**Figure 4**). This application is focused on setting up all the configurations and properties of the ADAM Tool such as the data source interfaces, variable properties and equations used for calculating the enriched variables. In addition, this configuration tool can manage the configuration tables of the Central DB and the tree of classifications of variables. The functionalities of the Configuration Tool can be divided into five groups: 1) roles and users, 2) variables, 3) units, 4) classifications, and 5) options.

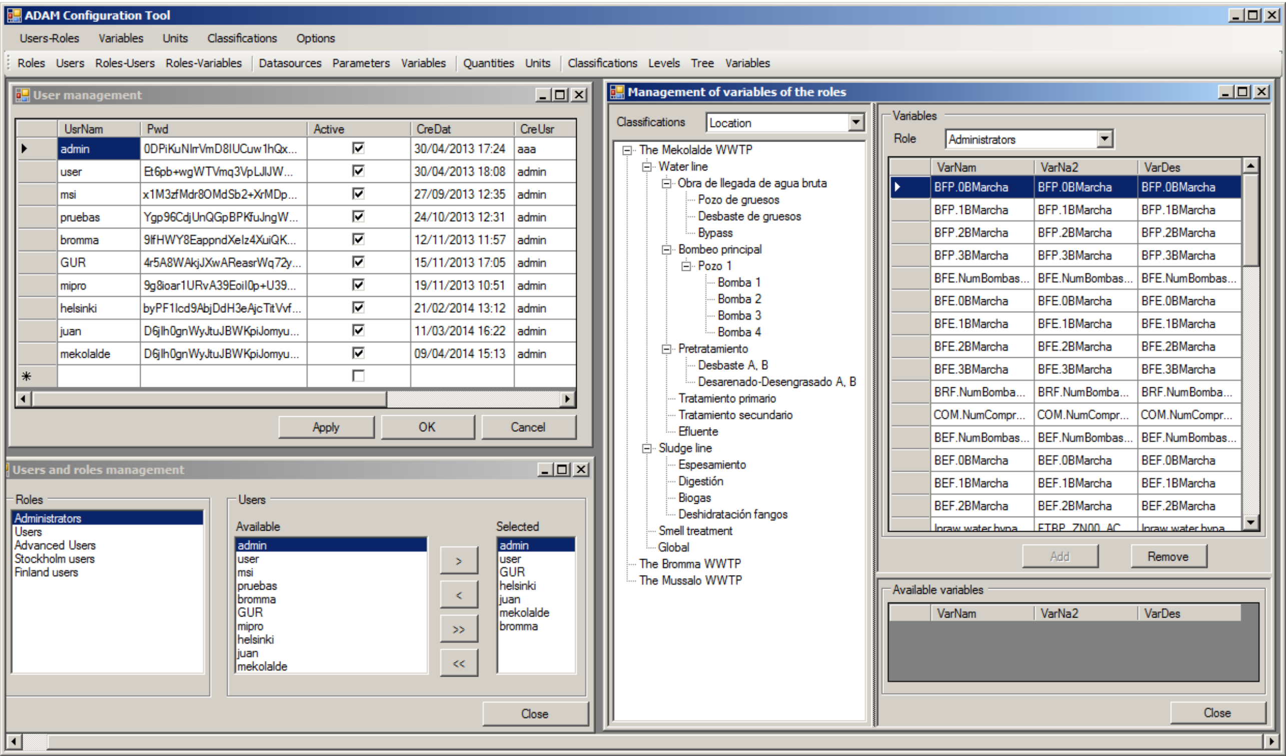


Figure 4 Configuration tool used for the administration of the ADAM Tool

A Data Collection Tool has been developed with the aim of collecting data by means of the different Data Source Interfaces (DSI) and their corresponding configuration. Thus, the Data Collection Tool has been developed with two main functionalities: a) it sends to each DSI the data frequency that it expects to receive from them, the list of the variables, and other information; b) it stores the data received by each DSI, after checking the security and the consistency of the data.

The Data Collection Tool checks the security and the consistency of the data by: a) checking whether the DSI is authenticated correctly; and b) checking whether the date ranges of the values do not overlap.

The DSI is the software that bridges the different data sources (programmable logic controllers, laboratory databases, maintenance databases, etc.) available in the WWTP and the Data Collection Tool of the ADAM Tool (**Figure 5**). Therefore, a DSI is in charge of reading the data coming from a data source and provides those data to the ADAM Tool. Since the ADAM Tool has a defined data format, each DSI is responsible for the following three functions: a) reading the original data; b) transforming the original data into the standard format of the ADAM Tool; and c) storing the transformed data inside the ADAM Tool through the Data Collection Tool.

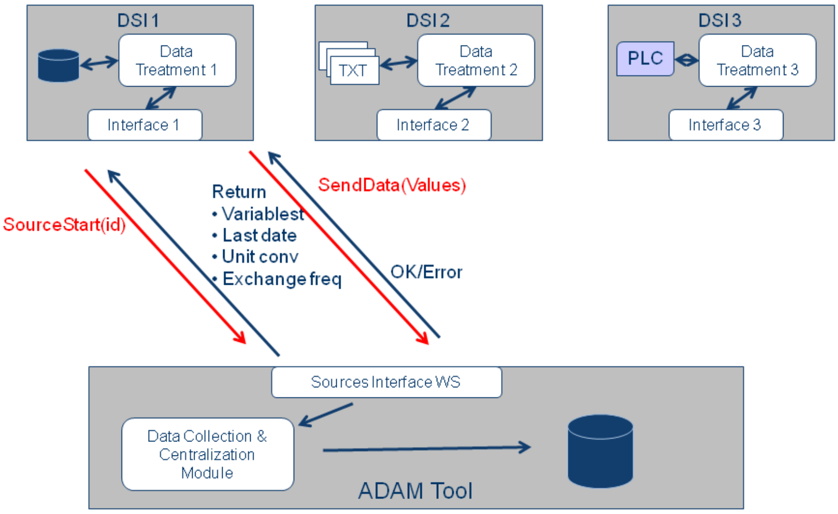


Figure 5 Interaction between the DSIs and the Data Collection Module

Raw data are reconstructed when a gap is detected. To reconstruct data, two algorithms were developed, each depending on the type of variable (sample or composite piece of data). In the case of a sample variable, linear interpolation is used between initial dates. However, in the case of composite variables, a linear interpolation is used between final dates.

With the purpose of testing the Data Collection and Centralisation Module, several data source interfaces (DSIs) were created to be stored in the Central DB data of the WWTPs where the ADAM Tool was tested and validated:

- The Mekolalde WWTP (60,000 m^3 d^{-1}, Spain). Four data source interfaces (DSIs) were developed in order to integrate the plant data from the Mekolalde WWTP in the ADAM Tool. The following data sources are being read: 1) online data, stored in the SCADA system. These data include not only data coming from online sensors and analysers but also advanced control data such as data from the ammonia control, dissolved oxygen control, nitrates control and suspended solids control (set-points and errors). 2) Laboratory data, managed by the software xPlims. 3) Energy data, provided by the energy supplier. 4) Environmental data, from an official weather webpage (air temperature, humidity and rain).

- The Bromma WWTP (126,000 m^3 d^{-1}, Sweden). Two data source interfaces (DSIs) were developed in order to integrate the plant data from the Bromma WWTP in the ADAM Tool. The following data sources are being read: 1) online data stored in the SCADA system; 2) Laboratory data.

- The Kymen Vesi Oy Mussalo WWTP (40,000 m^3 d^{-1}, Finland). Two data source interfaces (DSIs) were developed in order to integrate the plant data from the Kymen Vesi Oy Mussalo WWTP in the ADAM Tool. The following data sources are being read: 1) online data stored in the SCADA system; 2) Laboratory data.

- In addition, three more DSIs were developed in order to integrate the DIAMOND Systems with the ADAM Tool: a) two diagnostic tools for single sensor data (ODSSD), developed by IVL, UU, CERLIC, SVVAAB and AALTO; and b) the Advanced Control System (ACS), developed by AALTO, MIPRO and UU.

No DSI was developed for the integration of the Advanced Monitoring System (AMS) since the AMS only reads data from the Central DB. The Key Performance Indicators (KPIs) are calculated inside the ADAM Tool. All of the DSIs were developed in Microsoft Visual Studio 2010 and they share the library ADAMDscLib. The communication between the DSIs and the ADAM Tool was done through Web Services.

2) The Data Validation Module, which includes the algorithms and tools to guarantee the reliability and high quality of the collected plant data. It mainly consists of processing the data stored by the Data Collection & Centralisation module, checking the coherence of the data (i.e. detecting inconsistencies), and reconstructing data if necessary (i.e. a faulty data point has been detected).access, sample, acquire and store in a central database all the data of a WWTP that are needed for monitoring its status.

The problem of data validation of the laboratory analyses of WWTPs consists in the study of the consistency of three vectors of measurements: the vector of raw wastewater analyses (Y_{RAW}), the vector of settled wastewater analyses (Y_{SET}) and the vector of treated wastewater analyses (Y_{TRE}). Since the three vectors are not independent, the proposed methodology is based on the analytical redundancy approach.

First of all, the raw data is introduced in an outlier detection module for a preliminary check of faulty measurements using mass balances. If any outlier is detected, then the data will be sent to the fault isolation and data reconstruction module; otherwise it will be sent to the consistency check module where a more complex data analysis is carried out in order to detect any inconsistency in the data set.

This module is based on the mathematical modelling of the processes that correlate the three vectors of measurements among them. This can be described in a general way as: $Y_{SET} = f_1(Y_{RAW},\theta_1,\rho_1)$ and $Y_{TRE} = f_2(Y_{SET},\theta_2,\rho_2)$, where f_1 represents the function that correlates the set of input variables Y_{RAW} with the set of output variables Y_{SET}, θ_1 a set of parameters that are fixed and ρ_1 a set of uncertain parameters that can be estimated. For the second equation the definitions are analogous but applied to the correlation between the settled measurements Y_{SET} and the treated measurements Y_{TRE}.

In general, it can be stated that if the measurements comply with the process equations, then, the data can be considered consistent and it can be used for monitoring the plant performance or for control purposes; otherwise, that would mean that there is a faulty measurement in the data set.

One important thing to note is that some of the parameters of the models can vary overtime. Hence, if these fluctuations are not taken into account, even in the absence of faults, the equations of the process might be violated. Thereby, the set of uncertain parameters ρ needs to be adjusted each time a new data set arrives. For this calibration, an automatic optimization algorithm is used. This algorithm tries to minimize the discrepancies between the model f and the measurements Y^* while meeting some constraints, such as permissible ranges.

It must be noted that Y^* and Y represent different things. The first one represents the mathematical variable of a system while the second one is the measurement of that variable, which is affected by the noise of the sensors or, even worse, by faults. Thereby, while the variables always meet the equations of the process, this does not apply to the measurements.

The calibration procedure can lead to three possible outcomes. The optimisation algorithm might be unable to find a possible solution in order to adjust the input and output measurements. This means that no matter the value of the uncertain parameters of the model, the measurements just do not fit among them. Thereby, the explanation for this behaviour is that there is a fault in the data set. Another possible outcome is that the optimisation problem finds a solution but the values of the parameters that achieve this are not reasonable or even the evolution of this parameters does not make sense. The last possible outcome is that the algorithm finds a suitable solution. In this case, the measurement set is considered consistent since the mathematical model is unable to detect any discrepancy.

If the data set is inconsistent, the data validation procedure continues with the fault isolation and data reconstruction module; otherwise, the data have been validated.

In the fault isolation and data reconstruction module, an algorithm for detecting which is the faulty measurement is used. This algorithm also proposes a new reconstructed value for the faulty measurement, and, with it, the measurement validation procedure is finally completed. **Figure 6** shows the flowchart for the Data Validation Tool.

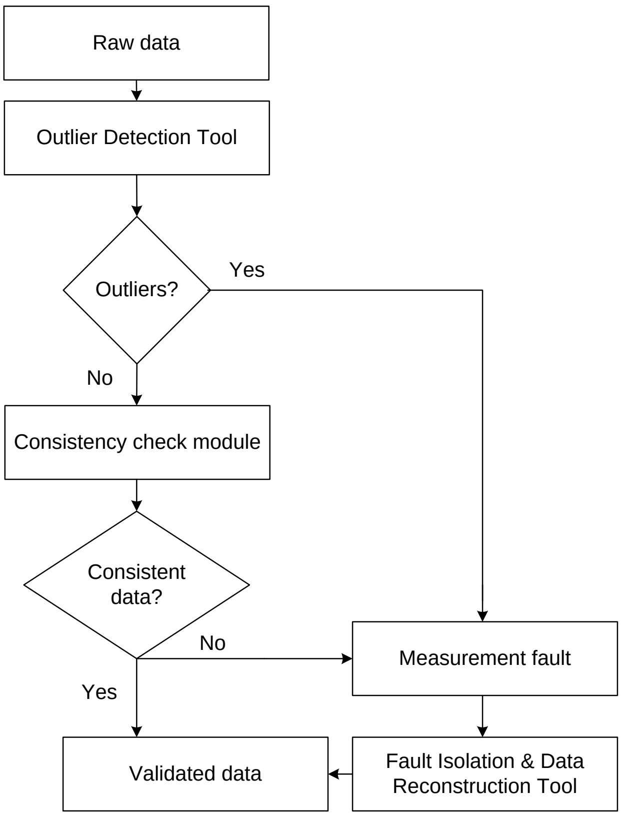


Figure 6 Flowchart for the Data Validation Module

3) The Synthesised & Enriched Information Generation Module, which consists of different algorithms and tools to provide plant status information by extracting valuable, synthesised and enriched, plant-wide information and generating new non-measurable information from the measurable data that has already been stored, checked and validated by the other two (Data Collection & Centralisation Module and Data Validation Module) software modules.

From the point of view of the programming implementation of enriched variables, there two types of enriched variables: 1) Formula enriched variables: they are defined by the administrator of the ADAM Tool by writing their full formula in a text field. For example: Influent NH_4 Load = Influent Flow rate·x NH_4. (**Figure 7**, left); 2) Special enriched variables: when the equation is too complex to be written in a single text field, some programming code must be used (e.g. when it is necessary to use computational loops). For example, an algorithm called "Time that a variable is higher than a value" can be used to calculate the period of time in which a flow rate exceeds an upper limit, or for the period of time during which the effluent $N\text{-}NO_3^-$ concentration is higher than the requirements. In this case, the algorithm for the calculation of the special enriched variable is previously written in the ADAM Tool (**Figure 7**, right).

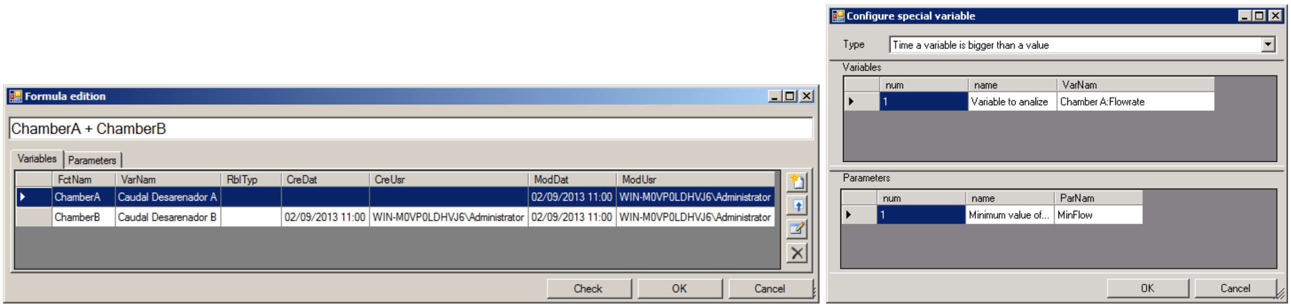


Figure 7 Visual appearance of the formula editor used to define an enriched variable and a special enriched variable

In order to combine measurable plant data of different types and frequencies, the data need to have the same number of data points and they have to coincide at the same time. To that end, several mathematical techniques were designed, developed and implemented for the generation of enriched variables (calculated from measurable plant data). The calculation of enriched values is developed as a Windows service (i.e. it runs automatically): it checks periodically at a previously set time if there are new raw data in order to calculate the current values of each enriched variable.

General conclusions for the Advanced Data Manager Tool

In summary, it can be said that the integration in the ADAM Tool of different plant data from the three WWTPs and from the DIAMOND Tools has been possible due to the standard design of the ADAM Tool. Moreover, the results obtained from the integration with the ADAM Tool with the 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, adequate information, and a straightforward connection to other emerging tools focused on the operational optimisation of the plant. In addition, the results of the developed ADAM Tool have showed its potential for generating new enriched information combining data from different plant sources, facilitating the connection of operational support systems, and developing automatic plots and trends of actual data for plant performance and diagnosis. This way, the flexibility and adaptability of the proposed technical solution will allow further implementation of advanced data management algorithms and tools and the integration with third party wastewater treatment operation software tools.

The Online Data Pre-processing Tool (The ODP Tool)

The Online Data Pre-processing tool (ODP Tool) has been developed including algorithms and methods for fault detection in sensors. The tool contains two main features: online diagnostics using single sensor data (ODSSD) and online diagnostics using multiple sensor data (ODMSD).

The ODSSD includes pre-processing of sensor signals, together with fault decision methods, and calculations of indication parameters describing the process status. The multi-sensor diagnostics contain both fault detection and indication methods sending out an alarm if a fault is detected and a suggestion of where in the process the fault occurred. The methods have been tested and validated in the case study at the Bromma WWTP, Sweden. The sensor faults tested in full-scale were drift due to clogging of a sensor and faulty measurements with a constant offset in the signal (bias). Fault detection were tested on sensors in one of the operating lines in order to cover sensor and process behaviour in situations when faulty data are used in control activities. The sensors used in testing and validation for the single sensor diagnostics were two types of dissolved oxygen (DO) sensors (optical and membrane), suspended solids sensors (SS) and a sludge level meter.

The multi-sensor fault detection methods (ODMSD) were tested on DO sensors installed in different aerated zones and working in closed-loop control.

Main results of the ODP Tool are presented below.

1. Online diagnostics using single sensor data (ODSSD)

Spectral analysis and likelihood ratio test – performing diagnostics on DO and SS sensors

This section was focused on the development of algorithms able to detect any significant change in the signal dynamics. For the fault detection method design, either change in the spectra or in the noise was assumed as fault detection indicators. It was assumed that a certain residual (ε) of the signal belongs to one out of two different hypotheses: H_0 (normal condition) and H_1 (faulty condition). Then, a fault is decided if $\varepsilon > h$, where h is a predefined threshold. The aim is to decide if the system has changed between H_0 and H_1 when changes in the dynamic of the process are presented. It was also assumed that H_0 and H_1 are equally likely. The method was tested for optical and membrane-type DO sensors (see the installation of DO sensors at the Bromma WWTP in **Figure 8**) and SS sensor.

Figure 8 Installation of the four DO sensors at the Bromma WWTP

Faulty conditions were detected after 10 days for optical sensors, after 15 days for membrane sensors, and after 10 days for SS sensor. For optical sensors, drift and noisy behaviour were observed. For membrane sensors, noisy behaviour was mainly observed. For SS sensor, drift behaviour was observed. An illustration of these responses is given in **Figure 9**.

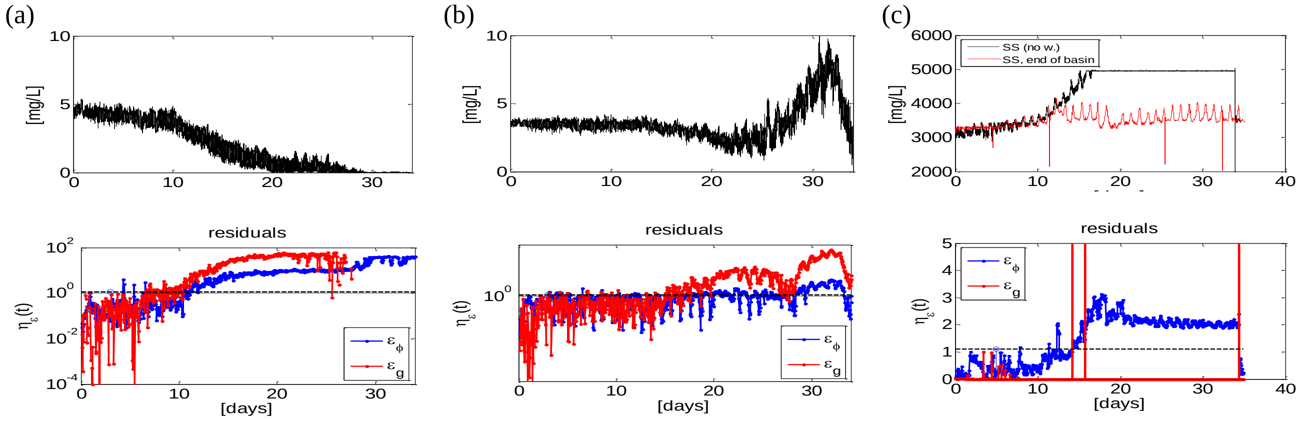


Figure 9 DO sensor and fault detection method response for: (a) Optical sensor; (b) Membrane sensor; (c) SS sensor. Signal profiles (top); residual responses (bottom)

Time constant of the cleaning response – performing diagnostics on DO sensors

This simple method is based on the sensors response to automatic cleaning with air. When cleaned with air, the dissolved oxygen is temporarily increased around the sensor. The time constant is the time it takes from the cleaning is initiated until the response has reached 63% of the total amplitude, i.e. maximum measured DO concentration as a result of the cleaning. The diagnostics was binary (non-faulty/faulty) based on a threshold value for the time constant. The method was tested and validated for an optical DO sensor during periods with different DO concentrations in the basin.
With the final algorithm, clogging of the DO sensor was detected on day 21, from manual cleaning, with a drift of -0.6 to -0.8 mgO_2/l compared to the reference DO measurements. Out of the tested methods, the time constant was considered to be the most reliable for specific fault detection of DO sensor, since process disturbances will not influence the fault decision. An example of the diagnostics using the time constant is given in **Figure 10**.

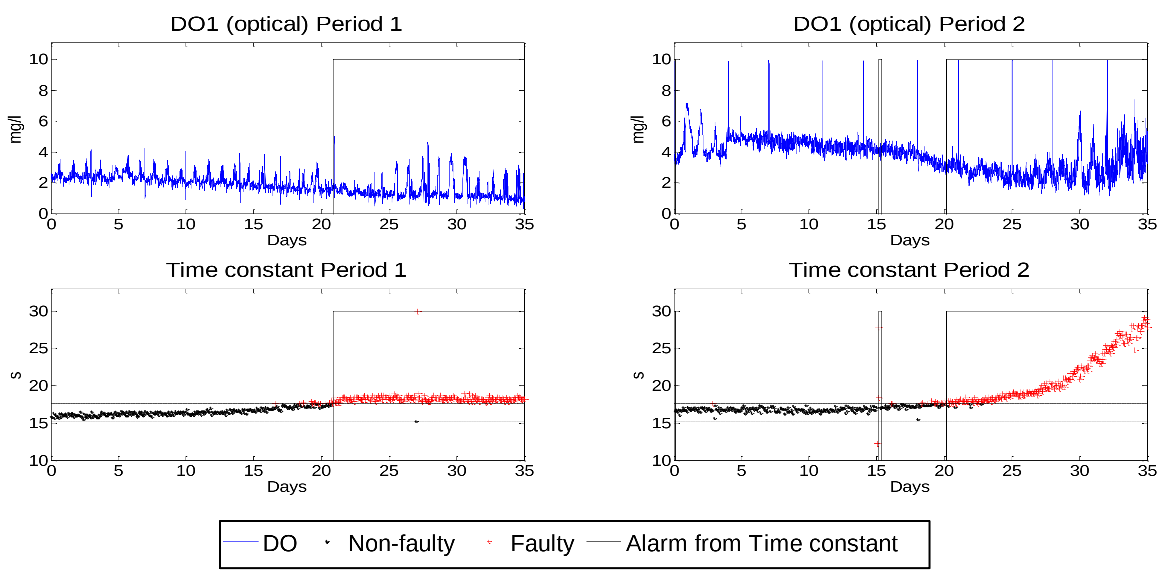
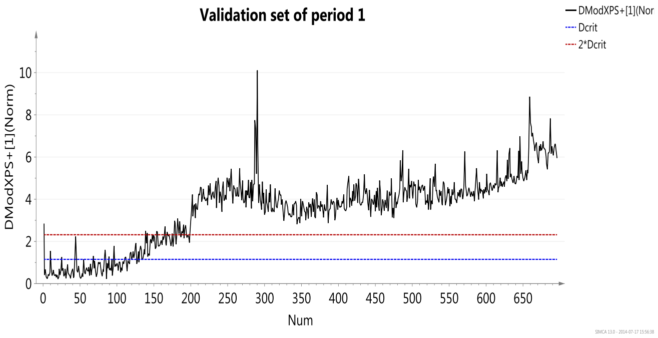


Figure 10 Example of the time constant method for diagnostics of DO sensor during two different periods

PCA of the cleaning response - performing diagnostics on DO sensors

The other fault detection method using the response to cleaning was based on principal component analysis (PCA) of the response. In general, using data in non-faulty condition (referred as training data), a PCA model can be developed, and certain confidence limits statistics such as DModX (residual distance, root mean square) can be calculated. Then, new data is evaluated during possible faulty conditions. If any fault occurs, it is expected that the value of the DModX will be larger than those given in non-faulty conditions. A threshold value (Dcrit, critical distance) is used to define the maximum tolerable distance of DModX.

The final PCA model uses the data from the full response and the method gives one DModX value for each response. In **Figure 11** DModX values are presented for the same periods as presented for the time constant. Each number (Num) corresponds to a cleaning event. The threshold Dcrit and 2*Dcits is marked with dashed lines. With the PCA method, the clogging was detected 15-17 days after manual cleaning, and the sensor drift was in the range of -0.2 mgO_2/l to -0.7 mgO_2/l compared to the reference.



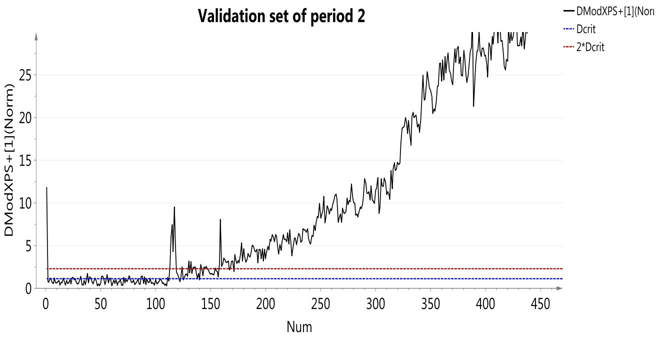


Figure 11 DModX values from the testing and validation of the PCA model for DO sensor fault detection using the cleaning response

Indication parameters for sedimentation process status – using data from the sludge level meter

A CBX sludge level meter was installed in one of the settlers at the Bromma WWTP (see **Figure 12**). It was measuring the SS concentration and level along the depth of the settler, which resulted in a sludge concentration profile as shown in **Figure 12**. There is no general definition in terms of concentration of SS for clear phase and sludge phase. For this evaluation a critical profile has been used to evaluate the sludge concentration profile. This critical profile consists of a set of maximum allowed concentrations at certain depths in the settler.

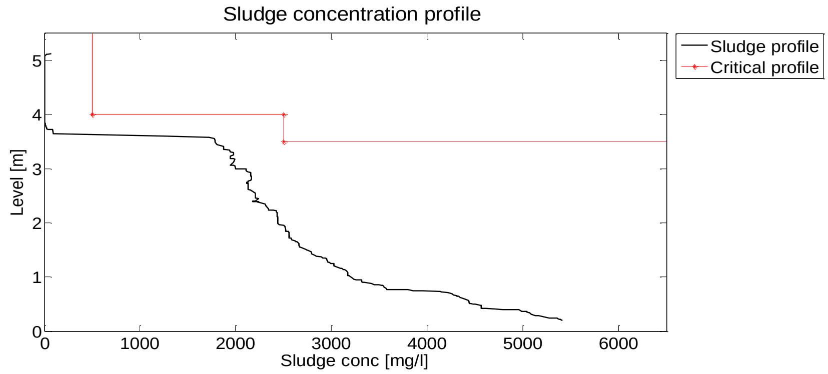



Figure 12 Left: Installation of the CBX sludge level meter at Bromma WWTP. Right: Example of a sludge concentration profile in the secondary sedimentation tank. The red line is an example of a critical profile.

Three indication parameters, calculated from the sludge concentration profile, were included in the ODP Tool: plateau level, thickness of the fluff layer and an estimated amount of sludge per square meter of the sedimentation tank. The sludge concentration profile generally has a plateau level where the concentration suddenly increases (in **Figure 12** the plateau level is found at about 3.7 meter). During the period when the sensor was automatically cleaned, the plateau level gave the same information as the reference sensor (see comparison in **Figure 13**).

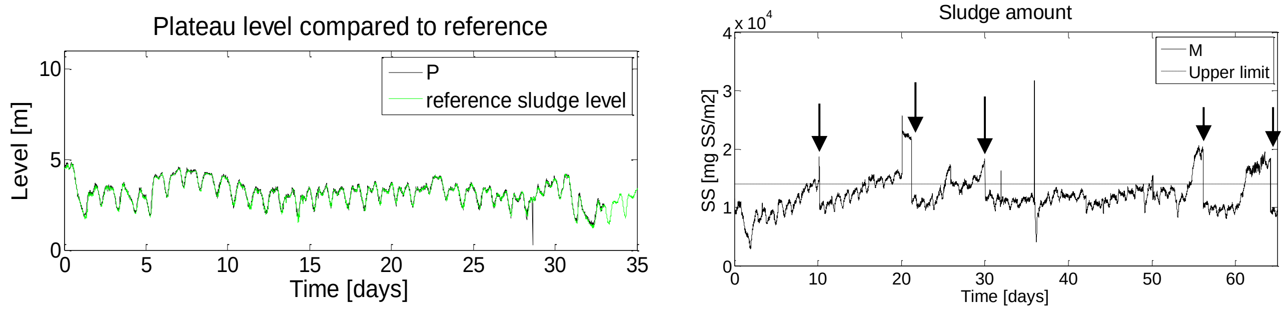
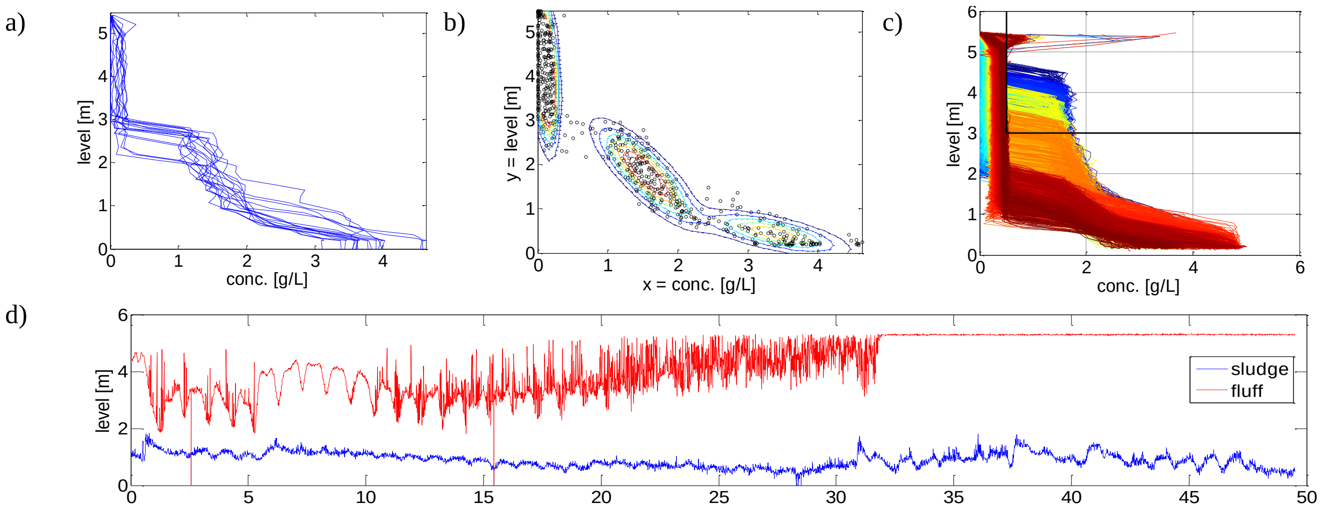


Figure 13 Left: Plateau level (P) compared to reference sludge level during period with automatic cleaning of the sensor. Right: Sludge amount (M) during test period with only manual cleanings. Arrows indicate when the sensor was manually cleaned. The sludge amount instantly decreases as the sensor was manually cleaned, which is indicating that the sensor indeed was dirty or clogged before cleaning

The thickness of the fluff layer was defined as the layer where the SS concentration was between 500 mgSS/l and 2500 mgSS/l. This parameter gives an indication of the settling properties of the sludge: a thick fluff layer is a sign of poor settling properties. The amount of sludge was estimated by an approximation of the integral of the sludge concentration profile. This indication parameter was showing a drifty or step-change behaviour due to clogging, during tests with no cleaning of the sensor (see **Figure 13**).

Gaussian mixture models – performing diagnostics using data from the sludge level meter

An alternative approach to get information from the sludge blanket meter is based on Gaussian Mixture Model (GMM). GMM is based on a collection of several sludge concentration profiles during a period in non-faulty conditions H_0, where a probability density function (p) is calculated and a residual (ε) is defined. After period H_0, a threshold (h) is computed based on the maximum value obtained for ε during H_0. For a new profile, the value ε is computed and compared to h. Then, a fault is decided if $\varepsilon > h$. An illustration of this method is given in **Figure 14**.



e)

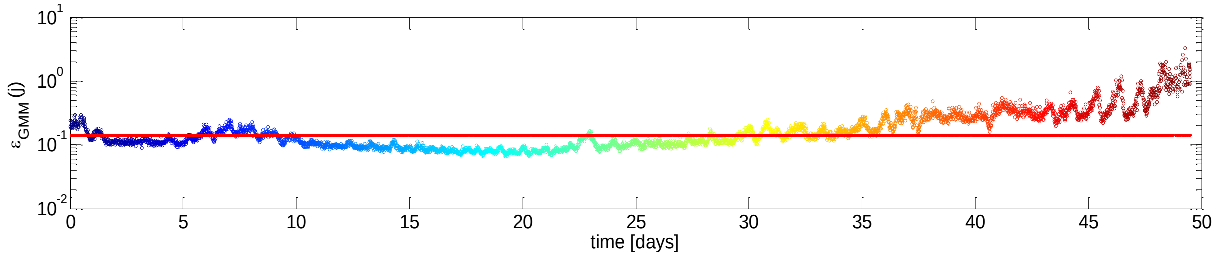


Figure 14 (a) Profiles in non-faulty conditions; (b) Data fitted with the GMM, showing the contours of constant probability; (c) Total of profiles, showing critical profile (black line); (d) Sludge and fluff levels; (e) Residual given by the GMM, showing threshold value (red line)

2. Online diagnostics using multiple sensor data (ODSSD)

PCA models - for DO sensor bias and clogging detection

The PCA method, as described for the single sensor fault detection, was also used for multi-sensor fault detection. Specific models were built and trained for the bias and clogging experiments respectively, taking in to account multiple variables such as the DO concentrations, valve positions and air flow rates in all aerated zones of the studied treatment line. Although the PCA method was looking promising when tested using simulated data, when tested at the case study plant, it was found that this method was sensitive to process disturbances such as abrupt changes in the control or weather changes. If process conditions were kept constant it was possible to detect a -0.5 mgO_2/l bias for a DO sensor, however false alarm was estimated to occur 15 % - 40 % of the time. A more general model was tested and gave only false alarm on one occasion, however this model did not detect all periods with bias (**Figure 15**). Moreover, this method required pre-processing and selection of data in order to reduce the variation in the data set, which resulted in a daylong interval between the diagnostics.

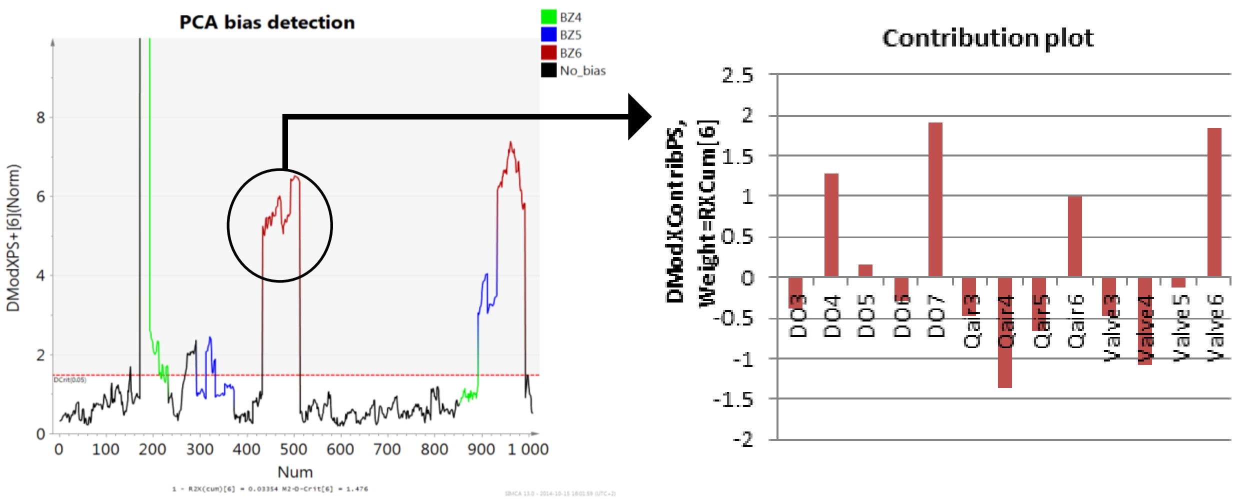


Figure 15 Results from PCA for bias detection. Left: DModX values, coloured according to bias (green for bias in zone 4, blue for bias in zone 5, red for bias in zone 6 and black for no bias). Right: contribution plot for the evaluated variables during one of the periods with bias in zone 6 (period indicated with a circle)

Both PCA models for detection of clogging were clearly capturing the fault during the sensor clogging periods; however the deviation from set point for the reference sensor was about 2.2 mgO_2/l – 3.7 mgO_2/l at the time for fault detection (see example in **Figure 16**). By using the mode-specific model, the fault was detected 2.7 days earlier, compared to the more general model. The pre-processing of data resulted in a diagnosis performed once every hour.

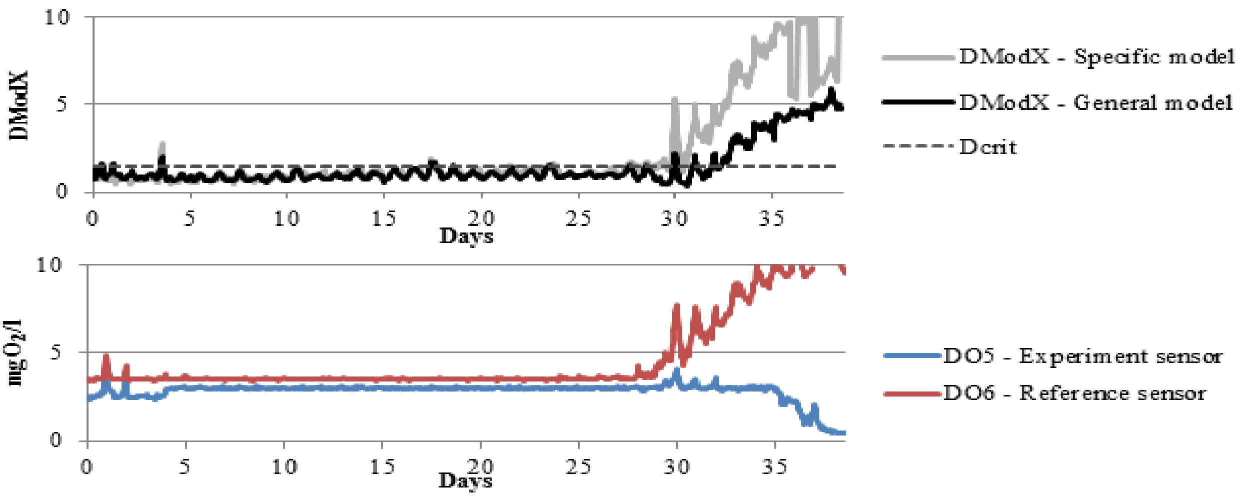


Figure 16 Results from PCA for detection of DO sensor clogging. Upper plot; DModX for the two tested models. Lower plot; DO5 – experiment sensor exposed to clogging, DO6 – reference sensor

AM and ARM - for DO sensor bias and clogging detection

An alternative method tested in real plant was the Airflow Ratio Method (ARM), which is based on the ratio between airflows (or valve positions) of different aerated zones. Experiments for bias and clogging detection were performed. For comparison, another simple method, airflow method (AM), was also tested. AM is based on monitoring the airflow (or valve position) of each aerated zone.

For bias experiment, AM and ARM used the valve position of three aerated zones for detecting faults. Results showed that both methods were sensitive to weather changes, giving several false alarms and wrong fault isolations. Compared to AM, the amount of false alarms given by ARM was lower (around 25% of false alarms for ARM, and around 30% of false alarms for AM). However, during stable conditions, it was possible to detect a -0.5 mgO2/l bias for a DO sensor. An illustration of the response of these methods for bias detection is given in **Figure 18**.

a)

b)

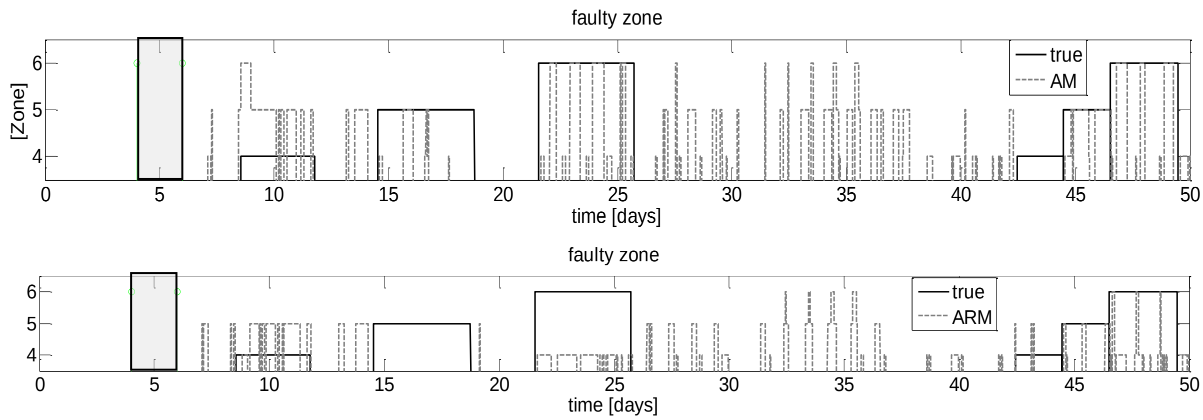


Figure 17 Response to bias experiment. (a) Applying AM. (b) Applying ARM. Period for threshold calculation marked with grey box

Similarly, valve position of three aerated zones was used for clogging detection. Compared to AM, this experiment shown better results in terms of fault detection and isolation when ARM was implemented. For ARM, the time delay for fault detection was around 4.5 days, which was approximately 1 day earlier when compared to AM performance. An illustration of the response of these methods for clogging detection is given in **Figure 18**.

a)

b)

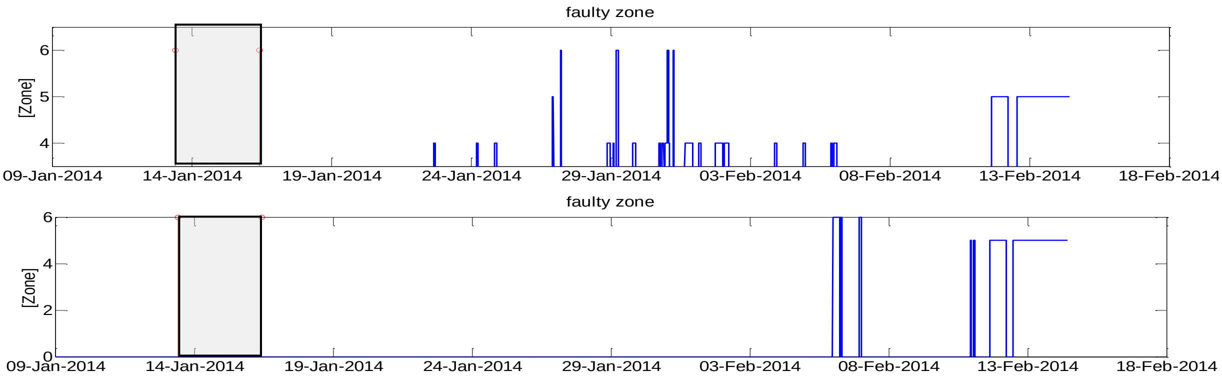


Figure 18 Response to clogging experiment. (a) Applying AM. (b) Applying ARM. Period for threshold calculation marked with grey box

General conclusions for the Online Data Pre-processing Tool

The development and testing of the Online Data Pre-processing Tool showed the potential of different fault detection methods in giving relevant information about sensor status and process conditions, enhancing the energy savings of the plant and the effluent quality.

Regarding single sensor scenario, it is important to remark that the results of this project show that using data from just a single sensor in full-scale operation will, in general, not make it possible to distinguish between small sensor faults (like a slow drift) or an abnormal process change. The methods should therefore be used as giving an alert to the operators that something may be wrong rather than giving an alarm that a specific sensor is faulty.

For all multi-sensor fault detection methods, the project has shown that it is expected to get indications of fault if the process conditions changes much from the situation the methods are trained for. Nonetheless, any such alarms could be of use to the operators, notifying them that something is deviating and needs to be investigated.

The Advanced Monitoring System (AMS)

In order to provide overall plant status information for plant operators and their decision-making processes, a set of plant-wide Key Performance Indicators (KPIs) that is necessary and sufficient for indicating the status of different aspects of a treatment plant have been defined and implemented. All the information needed to calculate KPIs come from experimental data that are efficiently collected and processed by the ADAM Tool. The calculated KPIs are also stored in the central database of the ADAM Tool so they can be used by other plant operation support systems.

In addition, a monitoring tool has been created so that plant operators can monitor in a single application overall plant status information, independently from the source of the data. This chapter describes the design, development, implementation and results of this monitoring tool.

The innovation of the AMS is twofold: on the one hand, plant operators and managers can see on a single application all the data measured at the plant independently from the source of the data and the database where they are stored. So far, each datasource in a WWTP (e.g. programmable logic controllers, laboratory analyses, energy provider databases, etc.) has its own software application to show its stored data to its managers. On the other hand, the AMS allows plant managers to easily create new information that is important for knowing the status of the whole plant (e.g. performance, efficiency, requirements, stability, etc.). Up to now, the generation of new non-measurable information was not straightforward since it requires the processing of different types of data (sample frequency, noise level, etc.) stored in different datasources.

Two modules have been developed with the objective of supporting the optimum operation of WWTPs:

1. The Real-time Monitoring Module

In order to monitor historical and real-time measured data, and synthesised new useful data (including KPIs) that are necessary for showing plant status, a technical document containing the KPIs of the AMS for the Mekolalde WWTP has been developed. The technical document is divided into five sections.

The first section includes the list of the parameters (i.e. of constant value) of the WWTP, such as the different volumes of the reactors, the design load or the effluent requirements. These parameters are used for the calculation of the KPIs.

The second section contains the list of variables (or tags) registered from the plant, which include data obtained in all the following sources of information: analytical laboratory, supervisory control and data acquisition (SCADA) system (i.e. online sensors and analysers), , manually collected databases and databases provided by the electricity energy supplier.

The third section is comprised of the list of equations that are used for the calculation of KPIs. Due to the large number of KPIs, the entire list is not shown; however, a brief example is presented in **Table 1**.

Table 1 Several examples of KPIs for the Mekolalde WWTP

Raw, settled and effluent load of TCOD	Load removed of TCOD	Removal efficiency of TCOD
Raw, settled and effluent load of FCOD	Load removed of FCOD	Removal efficiency of FCOD
Raw, settled and effluent load of BOD_5	Load removed of BOD_5	Removal efficiency of BOD_5
Raw, settled and effluent load of TKN	Load removed of TKN	Removal efficiency of TKN
Raw, settled and effluent load of NH_X-N	Load removed of NH_X-N	Removal efficiency of NH_X-N
Raw, settled and effluent load of TSS	Load removed of TSS	Removal efficiency of TSS

The forth section of the document presents a list of KPIs that are specially calculated (i.e. some short of computational code is used for their calculation). For example: mass of influent, settled and effluent TCOD, ratio kWh/TCOD removed, operating time of pumping, pumped volume, sludge retention time, hydraulic retention time or maximum level of tanks.

The fifth (and final) section explains the mathematical methodology that is used to combine measurable plant data of different type and frequency, needed for the calculation of the KPIs.

An Advanced Monitoring System (AMS) Tool has been developed as a graphical analysis environment of the stored plant-wide process variables (raw, enriched variables and KPIs). The objective of the AMS Tool is to allow users to monitor and analyse variables with the aim of speeding up the decision-making of plant operation.

The AMS Tool has been developed based on the requirements (i.e. needs, characteristics and uses to be considered) of the project SMEs and End-users. The AMS Tool has been developed in Microsoft Visual Studio 2010 and .Net Framework 4.0. The AMS Tool has multilingual support, being Spanish and English the current languages that the Tool supports.

A group of sheets is called a scoreboard and the AMS Tool can have opened several scoreboards at the same time. Each user (e.g. plant operators) can store in the ADAM Tool their own scoreboards for personal use so that they could be accessible for the user independently on the computer they are logged in. This way, the visualization of plant information becomes faster. **Figure 19** shows the visual appearance of a scoreboard of the AMS Tool.

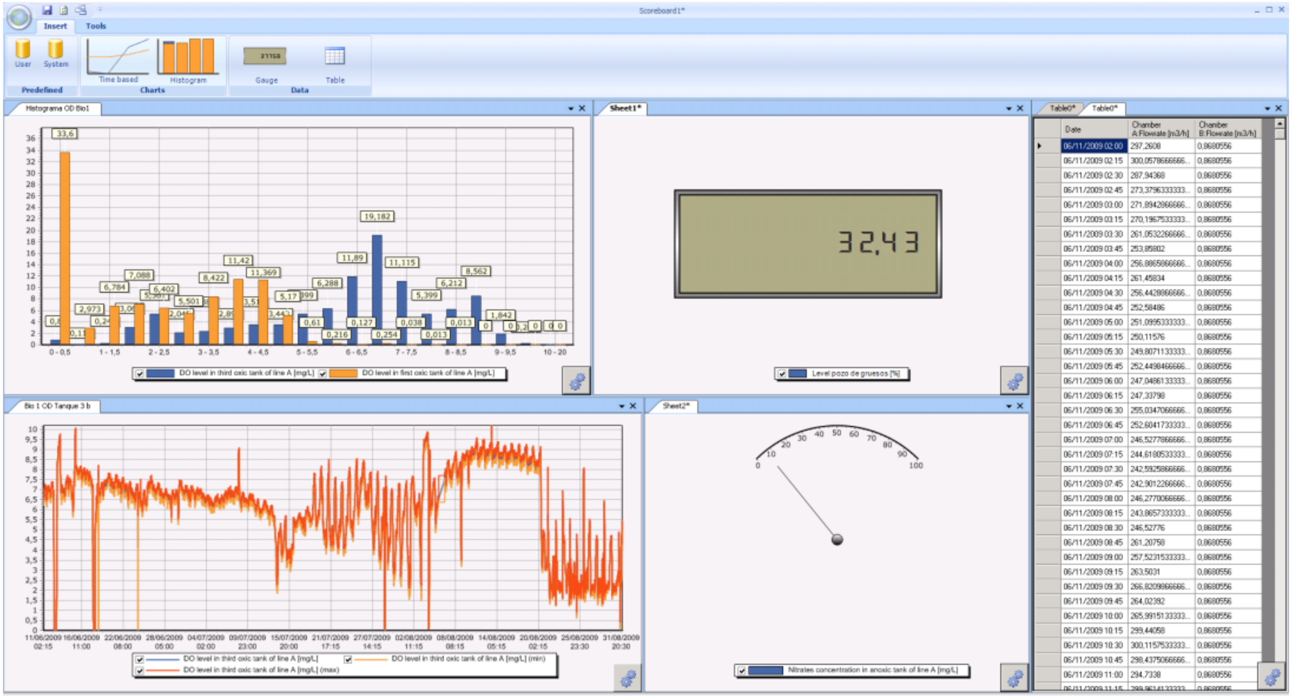


Figure 19 Visual appearance of a scoreboard of the AMS Tool

Users can also open system scoreboards and sheets that have previously been configured by an administrator user (e.g. plant managers). Administrator users can convert his personal sheets and scoreboards in system ones in order to them be shared with other users (e.g. plant operators).

The application layout of the AMS Tool is defined by dockable sheets that contain one single chart. The charts can be divided into four different groups:

a) Time based chart: displays one or more signals based on time (i.e. the X-axis is a timeframe and the Y-axis displays one or more variables). The objective of this chart type is to show the time evolution of variables so that the user can easily monitor the operation of the WWTP.

b) Histogram chart: displays the distribution of the data (one or more variables) against time, against another selected variable or against some ranges of values. The objective of this chart type is to allow users to analyse de distribution of the data, changes in the plant operation tendencies and the relationship between different variables.

c) Gauge chart: displays only one value for each variable, which can be the last registered value or the average value for a period of time. Thus, the objective is to display a single value of one variable which is important for users.

d) Table: displays the values of several variables in a rows and columns format. The objective is to provide a summarised overview about the values of important plant variables for further analysis.

The developed AMS can be used to detect improvements that could be done to the plant, and it can quantify in energy consumption, effluent quality, etc. the changes that have been previously done to a plant. Next, an example is presented.

One important indicator for plant performance is the active energy consumed by a plant per volume of wastewater that is treated (i.e. the ratio kWh/m3). This ratio is used to compare the performance among WWTPs of similar technologies and sizes.

Figure 20 shows the ratio of the energy consumption per volume of treated wastewater in the Mekolalde WWTP in each month of the year 2010. For analysis purposes, the temperature of the wastewater is also shown. The figure is an actual screenshot of the AMS Tool and it was created using few steps, showing the potential of the tool for an easily development of plots. The figure is a combination between time-based charts (lines and bars plots) and gauge charts. The standard centralisation of the data done by the ADAM Tool allows plant operators to see in one single chart KPIs (ratio kWh/m^3) and raw plant data (water temperature).

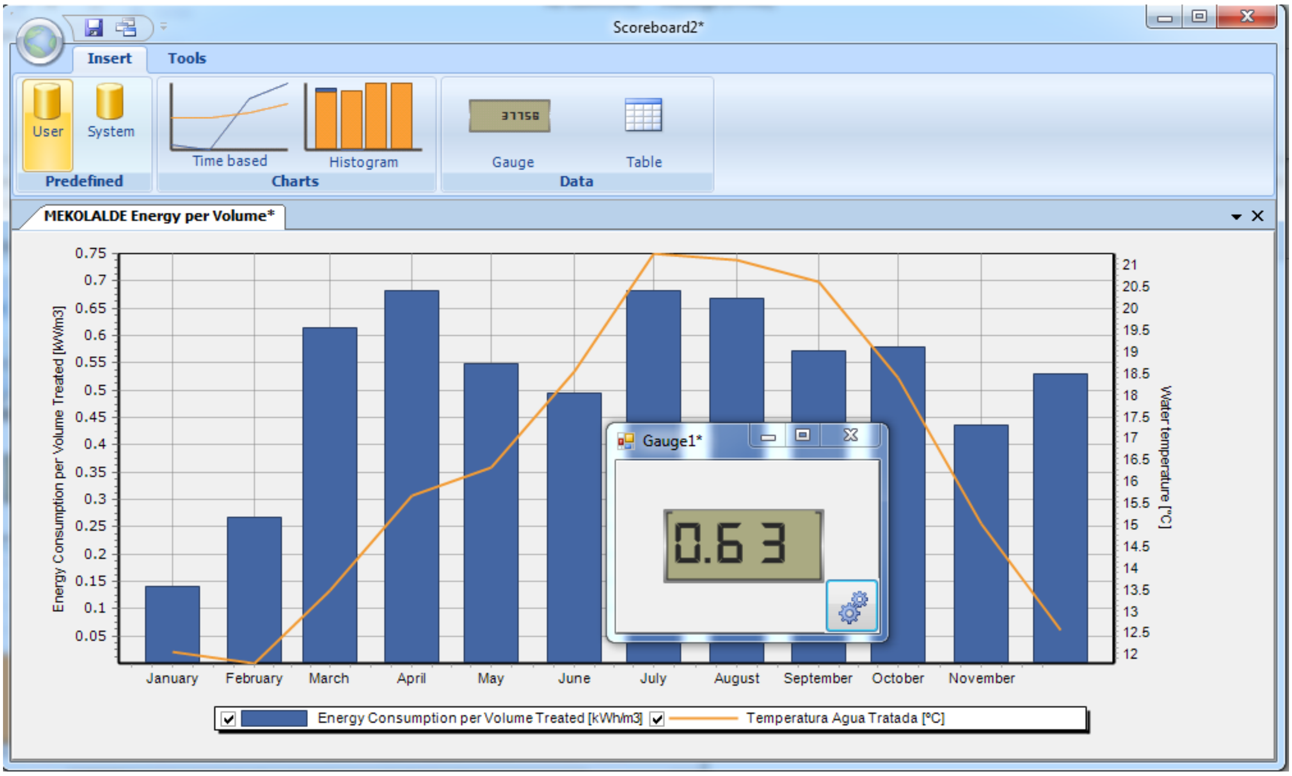


Figure 20 Ratio of the energy consumption per volume of treated wastewater in the Mekolalde WWTP (AMS)

Plant operators can point out several important facts based on the figure. The obtained specific value, 0.63 kWh/m^3 (gauge value), is a typical value for Spanish WWTPs: 0.67 kWh/m^3. During the first three months of the year (January, February and March) the heavy rain affected largely the ratio kWh/m^3, pulling down the value from 0.7 kWh/m^3 to 0.15 kWh/m^3 (**Figure 21**). The influent wastewater volume was also high during January and February due to ice melting, variable that can also be seen in the decrease of the water temperature.

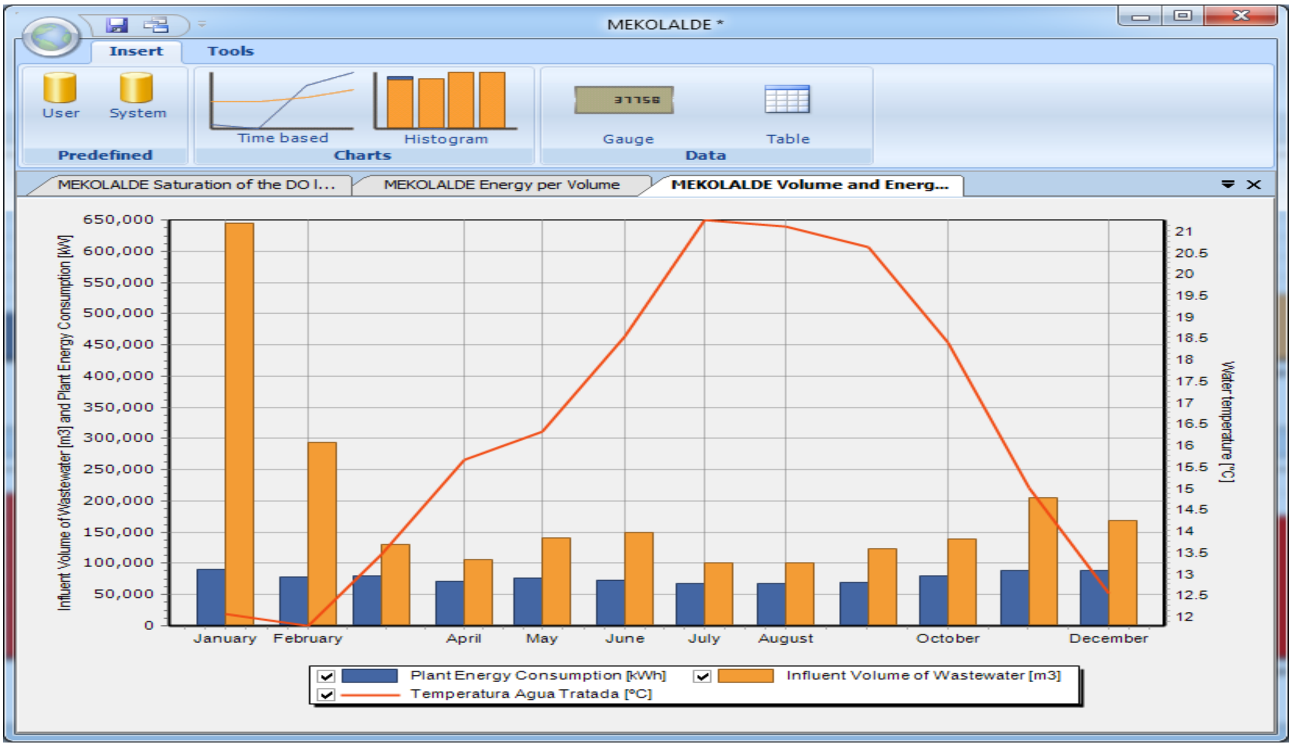


Figure 21 Influent volume of wastewater (m^3, blue) and total plant energy consumption (kWh, red) in the Mekolalde WWTP visualized in the AMS Tool

2. The Predictive Monitoring Module

The objective of this module is to provide the future plant status based on current plant conditions, making easier and more efficient the operational decision-making processes.

In the previous section, KPIs have been defined in order to monitor the past and current status of the Mekolalde WWTP. On the other hand, due to the slow dynamics of wastewater treatment plants, changes in some of their operational variables cause effects in the status of the plants several days later. Thus, it is crucial for plant operators not only to assess the past and present performance of a plant, but it is also important that plant operators should know the future status of a plant given the current operation. This way, in case the predicted KPIs show that the future performance of the plant is not suitable, plant operators could take preventive actions. For solving this issue, and as part of the AMS, a model-based plant performance prediction tool has been developed with the aim of predicting KPIs to monitor the future status of the plant.

One of the most important parts of a plant is the secondary treatment. This way, the prediction of KPIs of the Mekolalde WWTP is focused on the prediction of the behaviour of the activated sludge process in the secondary treatment. This process has been modelled based on a modified Activated Sludge Model No. 1 (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, UK). Two main modifications have been implemented. The first one consists of the exclusion of the particulate and soluble biodegradable organic nitrogen. Instead of them, for adjusting the nitrogen cycle, the ammonia concentration is used just as in the ASM2d (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, UK). The second one consists of the inclusion of the inorganic particulate matter so that the total suspended solids concentration can be adjusted more easily.

The ordinary differential equations of the model have been introduced in a dynamic library and they are solved using the numerical integration CVODE algorithm (http://computation.llnl.gov/casc/sundials/main.html). This algorithm is able to solve the set of ordinary differential equations at steady state within a second. Thus, the results are obtained almost immediately.

It has been decided to use Microsoft Office Excel® as the algorithm interface due to its user friendliness characteristics. The interface is divided into two worksheets. The first worksheet is used for introducing the model parameters so that it can be easily calibrated. For the current conditions, a set of parameters has been calibrated using historical data of the Mekolalde WWTP.

Figure 22 shows the visual appearance of the input and output interface of the plant performance prediction algorithm of the Mekolalde WWTP. The influent wastewater measurements are first introduced. In the case of the Mekolalde WWTP, these measurements are the total and dissolved chemical oxygen demand (COD_T and COD_S), the total suspended solids concentration (TSS) and the ammonia concentration (NH_X). With these measurements, the influent is fully characterised but operational plant variables are also needed for the simulations. Hence, the following operational variables are introduced: the temperature of the wastewater (T), the effluent total solids concentration (TSS_{eff}), the influent (Q_{inf}), internal (Q_{int}) and external (Q_{ext}) flow rates and the dissolved oxygen set point of the aerated reactors (DO).

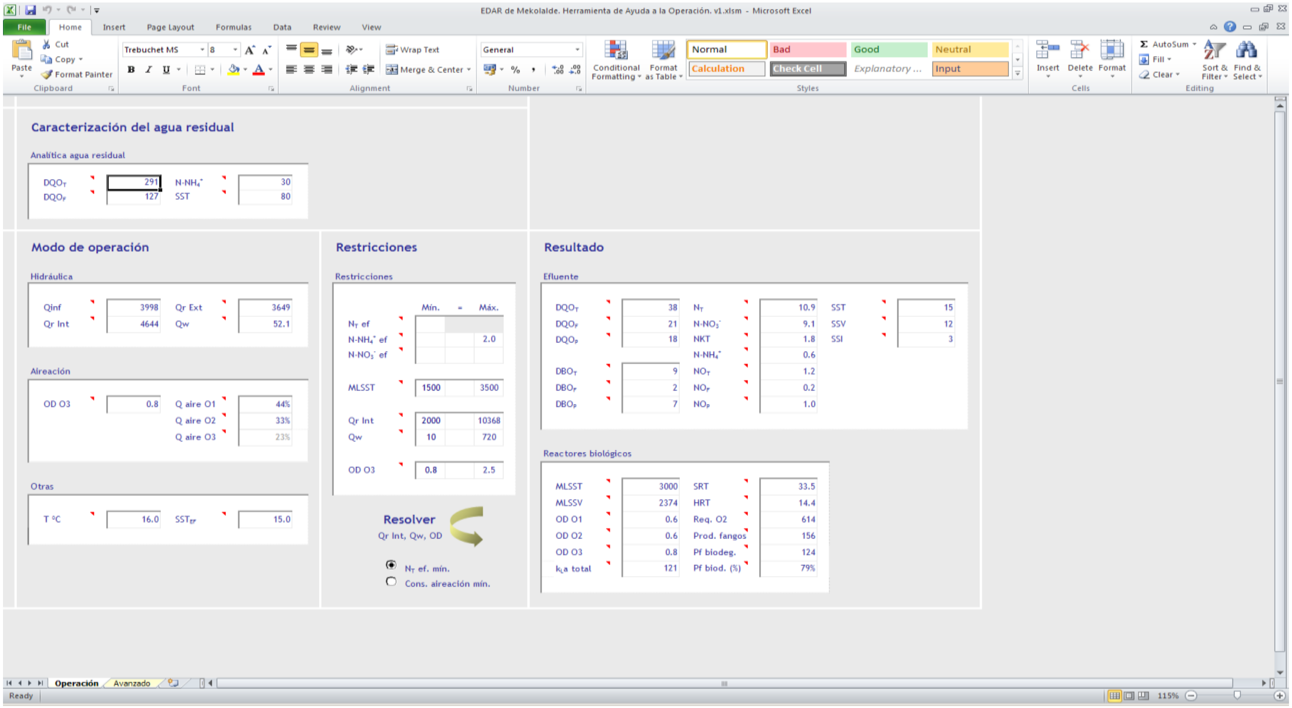


Figure 22 Visual appearance of the input and output interface of the of the plant performance prediction algorithm of the Mekolalde WWTP

Once these data are introduced, the algorithm automatically solves the set of ordinary differential equations and obtains the outputs at steady state conditions. These outputs are: the quality of the effluent in terms of chemical oxygen demand (dissolved and particulate), biological oxygen demand (dissolved and particulate), total organic nitrogen (dissolved and particulate), ammonia, TKN (dissolved and particulate), nitrates, total nitrogen (dissolved and particulate) and total suspended solids and volatile suspended solids. In addition, the total and volatile suspended solids of the reactors are also given, as well as the oxygen requirements, the sludge production, the hydraulic retention time and the sludge retention time.

To sum up, a plant performance prediction algorithm has been developed to predict the future values of several KPIs so that plant operators could know the future status of the Mekolalde WWTP. The future plant status is shown in terms of sludge production, suspended solids concentration in the biological reactors, nitrification rates, effluent quality (concentration of organic matter –COD, FCOD, BOD_5–, nitrogen –total nitrogen, ammonia, nitrates and organic nitrogen–, and solids –total, volatile and inorganic–) and energy consumption of the aeration system. Thus, and thanks to the plant performance prediction algorithm, the AMS can be also seen as a tool that helps plant operators with the decision making of the medium and long term operation of the WWTP.

As an example of the potential applicability of this tool, a graphical abacus for the operation of the dissolved oxygen concentration is shown in **Figure 23**. In this case, the tool is used to assess the plant operator in the selection of the set point of the dissolved oxygen in the aerobic reactors. Although the Mekolalde WWTP has an ammonia nitrogen control that continuously moves the set point of the dissolved oxygen, this control sometimes has to be switched off, for example when the ammonia analysers are been calibrated. For a maximum limit of 1 gN/m^3 for the effluent ammonia, the vertical axis represents the total suspended solids concentration, which varies between 1500 gSS/m^3 and 4500 gSS/m^3, while the horizontal axis represents the temperature of the water, which is in the 10 °C and 25 °C.

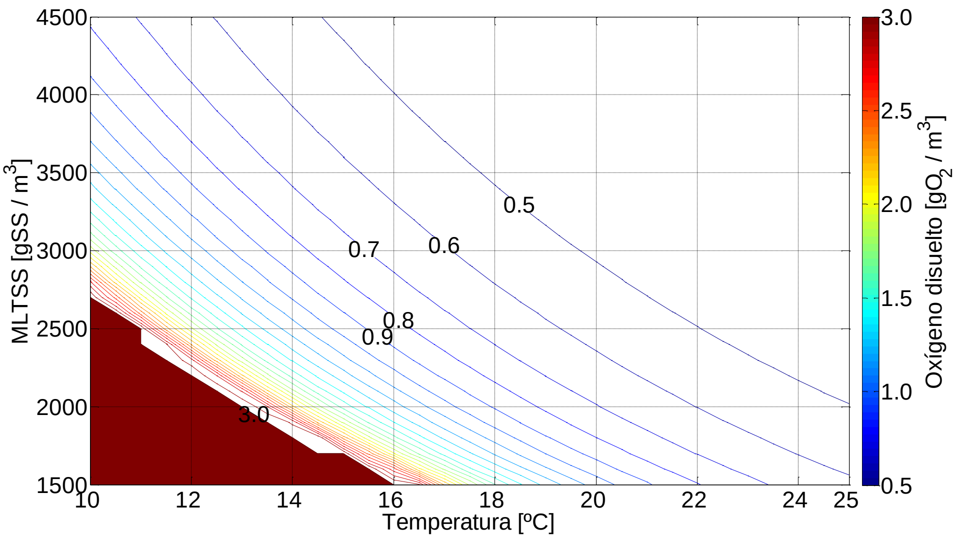


Figure 23 Graphical abacus of the operation of the dissolved oxygen for 1 gN / m^3 of ammonia in the effluent

General conclusions for the Advanced Monitoring System

In conclusion, it can be said that the Advanced Monitoring System proposes a solution for facilitating the monitoring of plant state by both synthesising all the available data and generating new useful key-process information for plant operators. As an example, the tool has helped the Mekolalde WWTP operators to provide and monitor overall plant status information, speeding up their decision-making processes. With this tool, a model of excellence in the management of all the available data has been achieved, supporting this way the optimisation of the global operation of WWTPs in terms of water quality, resources and energy.

The Advanced Control System (ACS)

The DIAMOND Advanced Control System (ACS) aims at the improvement of the operation of WWTPs in terms of process stability, effluent quality and operational costs. ACS consists of two modules (the Optimal Process Control Module –ACS-OPCM– and the Statistical Process Control Module –ACS-SPCM–) and a set of data transmission interfaces between the ACS applications (implemented in Otaniemi, at the Aalto University, Finland), Mipro’s data exchange center (at Mipro Oy precincts, in Mikkeli, Finaland) and the WWTP control system (in Mussalo, Finland).

Given the tight effluent requirements of the Mussalo WWTP, the significant disturbances of the influent, the low temperatures and the need of an essential reduction of the energy consumption, the testing and full-scale validation of ACS involved:

1. Offline simulation of ACS-OPCM in virtual Mussalo WWTP

To test the potentialities of the ACS-OPCM on the Mussalo WWTP, Line 2 is selected as representative of the activated sludge process and it is modelled in GPS-X (www.hydromantis.com/GPS-X.html), a commercial simulation software for simulating municipal and industrial WWTP. ACS-OPCM is developed in Matlab and linked with GPS-X model of the ASP line 2 in Mussalo that acts like the real-plant.

The model is calibrated using the available influent laboratory analysis and the online data, allowing a realistic representation of the dynamic behaviour of the ASP Line in Mussalo. As an example of the simulator performances, **Figure 24** reports a comparison of the simulation results for the nitrogen compounds. An overall good fit between the measurements (the dots) and their reconstruction (the solid lines) can be appreciated.

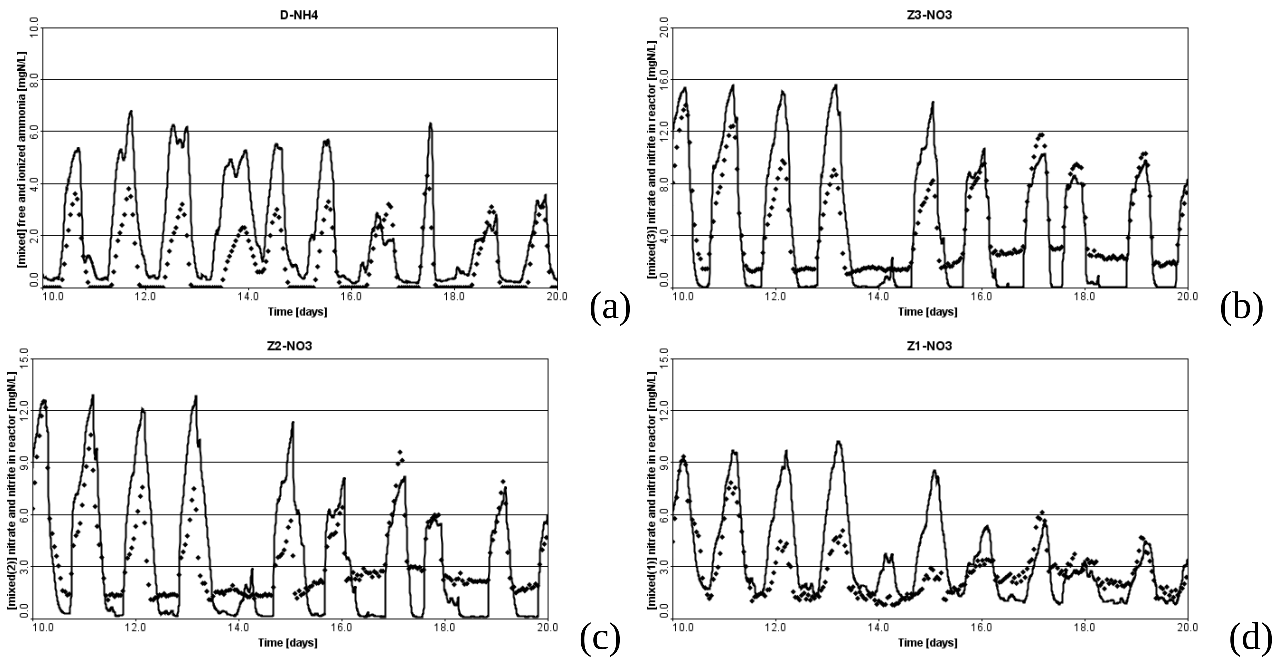


Figure 24 Simulation results: Ammonia at the end of the ASP (a), Nitrate in Zone 3 (b), Nitrate in Zone 2 (c), Nitrate in Zone 1 (d) comparison between simulation (solid line) and plant data (dots).

The simulation environment in GPS-X is used as testing platform for the design and selection of the configurations for the Mussalo WWTP as well. In particular, two simple structures are tested and validated on the GPS-X model of the Mussalo plant:

• ACS-OPCM with ammonia:
- One ACS Controlled Variable: Ammonia concentration at the end of the ASP.
- Three ACS Manipulated Variables: DO set-points in the anoxic zones and internal recirculation flow rate.

• ACS-OPCM with nitrates:
- Two ACS Controlled Variable: Ammonia concentration at the end of the ASP and the NO3 concentration in the last anoxic zone.
- Three ACS Manipulated Variables: DO set-points in the anoxic zone and the internal recirculation flow rate.

For both structures, the dissolved oxygen in the aerobic zones is strictly kept at the given set-points by the low-level DO-controllers. The DO set-points in those zones are set as 2 mg/l in the fourth aeration zone, 1 mg/l in the fifth and last zones of the bioreactor.

Simulations results show that the total energy consumption is reduced in a single line in Mussalo (Line 2) by 17 % and the effluent quality is improved by 20 % in terms of total nitrogen. As expected, ACS-OPMC marginally affects the organic compounds and only the effluent concentration of BOD5 is improved only by 8 %.

2. System identification

The simulation platform allows also the characterisation of a dynamic model between each pair of manipulated and controlled variables of the ACS-OPCM structures. This task is known as system identification task.

On the plant level, the identification tests are usually carried out manually: a series of independent steps to the manipulated variables in the ACS-OPCM structures should be given and the response on the controlled variable should be registered. This practice is however rather intrusive for the regular operation of the plant. For the Mussalo WWTP, the task is approached only by using the available data and the available measurements, avoiding any extra burden to the plant managers at the Mussalo WWTP. A specific period of about one week is selected such as during this time, the ACS-MVs are varying in a step-fashion way and the ACS-CV is responding accordingly.

Result of the identification procedure allowed the definition of simple first order plus time delay models (FOPTD) for every pair ACS-Controlled Variable/ACS-Manipulated Variable.

3. Pre-test activities

A tight adhesion of the low-level dissolved oxygen controllers to the set-point given by the ACS-OPCM controllers represents a fundamental requirement for the ACS success. The assessment of the status of the low-level DO-controllers at the Mussalo WWTP was already addressed in the first period. The tuning of low-level DO-controllers at the Mussalo WWTP towards tighter disturbance-rejection and set-point-tracking performances has started in October 2013 for Line 4. Results of the pre-test activity produced improvements in the low-level DO controllers: a) reducing overshoot, particularly in the anoxic zones of the bioreactor when the DO set-point is changed from 0 to 3 mg/l; b) slight stabilisation of the low-level DO controllers in the aerobic zones.

As an example **Figure 25** shows the DO measurement in ASP Line 4 (upper) and Line 3 (lower): On the left-hand side column a zoom on winter (19 – 23 February 2014), on the central column the considered period (16 February – 10 August 2014) and on the right-hand side a zoom on spring (10 – 14 May 2014).

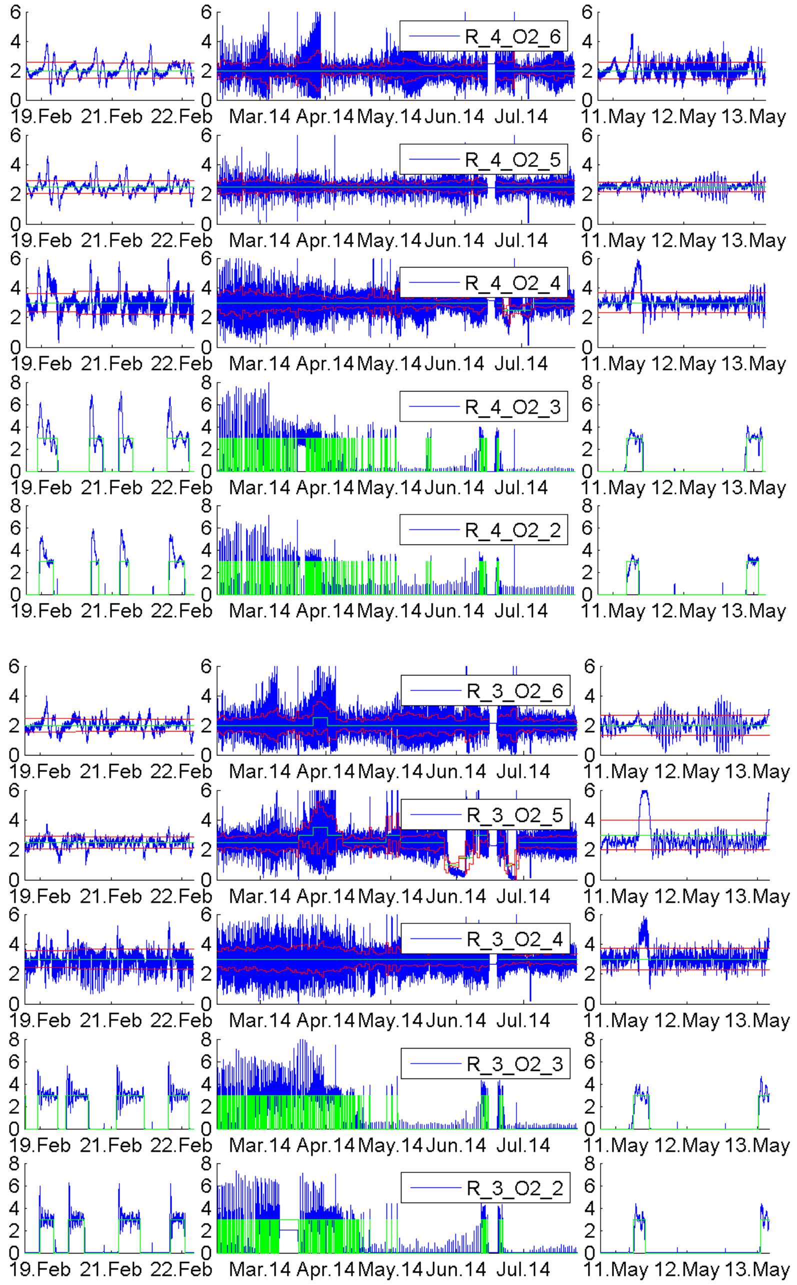


Figure 25 Example of the pre-test activities for Line 4 (upper) and Line 3 (lower)

4. Offline simulation of ACS using real data collected from the Mussalo WWTP

The real-time on-line data retrieved by the Miso System made as 15-minute averages are used for testing the offline testing of the ACS-OPCM structures. Two scenarios are investigated: (i) Winter scenario (26 February – 7 March, 2014); (ii) Spring scenario (12 May – 29 May, 2014). The ACS-OPCM offline results on the real Mussalo are reported for the two DIAMOND-ACS configurations.

Figure 26 and **Figure 27** give examples the two ACS structures for ASP Line 2. The comparison is done between the simulation results obtained with the ACS-OPCM (blue line) and the data collected from the Mussalo WWTP (red line) during the winter (left column) and spring (right column) scenarios, shows a consistent agreement. In **Figure 26** the controlled variable is the ammonia (first row) and the manipulated variables the DO set-point in zone 2 (second row), DO set-point in zone 3 (third row) and internal recirculation flow-rate (fourth row).

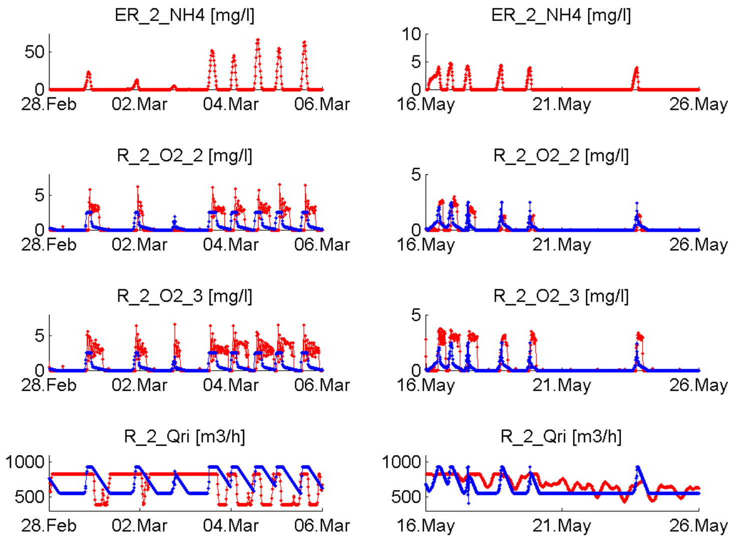


Figure 26 Example of the offline simulation results on Line 2 of the Mussalo WWTP for winter (left-hand side column) and spring (right-hand side column) scenarios with the ACS-OPCM with nitrates

Figure 27 shows the offline simulation results on real Mussalo WWTP for winter (left-hand side column) and spring (right-hand side column) scenarios. Here the controlled variables are the ammonia (first row) and nitrates (second row). The manipulated variables are the DO set-point in zone 2 (third row), DO set-point in zone 3 (fourth row) and internal recirculation flow-rate (fifth row). The Mussalo measurements are red and the offline simulations of ACS-OPCM are blue.

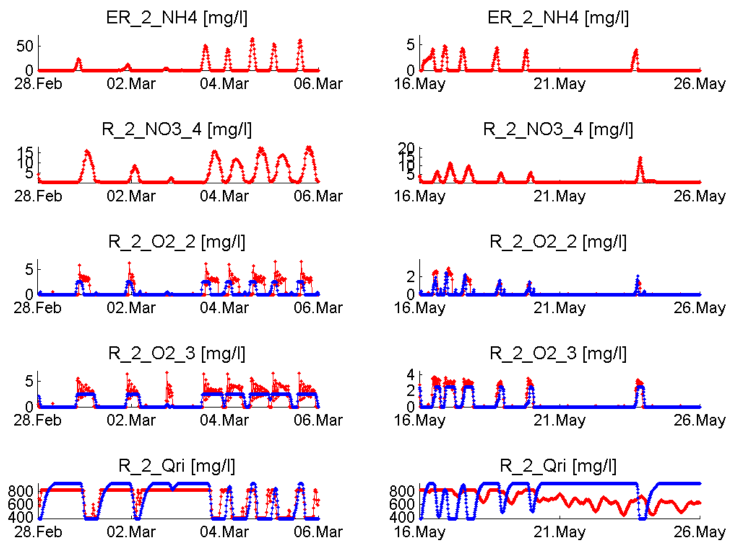


Figure 27 Example of the offline simulation results on Line 2 of the Mussalo WWTP for winter (left-hand side column) and spring (right-hand side column) scenarios with the ACS-OPCM with nitrates

Results of the offline simulation showed an average percentage reduction between the ACS-OPCM DO set-points and the real Mussalo data during the offline testing period (February 16 – August 10, 2014) of 38 % for every ASP line. This percentage is meant as approximated index of the aeration energy consumption in the second and third zone of every ASP line included in ACS-OPCM. The aerobic zones in Mussalo ASP lines are not included in the current ACS-OPCM configurations and not considered in the index.

5. Set-up of the communication interfaces

Some specific pieces of software have been developed for data exchange between the automation system of the Mussalo WWTP and the testing systems located at Aalto University. This will not be the procedure when the ACS is installed locally in the wastewater treatment plant under control, but this work has been very valuable to validate the technical feasibility of the ACS solution as Software as a Service (SaaS). In this regard, this software will be reused for the future in commercial installations. Examples of the communication interfaces are reported in **Figure 28** and **Figure 29**. In particular, **Figure 28** shows the selection menu of the ACS set-points, as “graphical switch board” developed for the project. The operators can switch between the “normal” and “ACS” modes line wise. **Figure 29** shows the ACS output of the status of the aeration lines is presented by “traffic lights”.

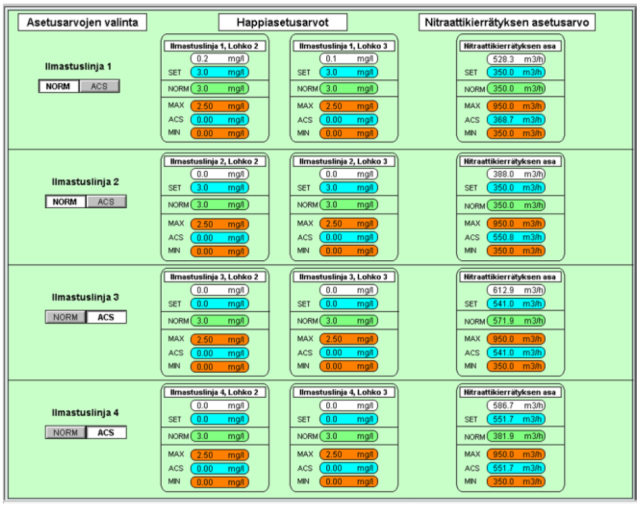


Figure 28 Selection menu for ACS set points ()

Figure 29 ACS status data for the plant and lines

6. Online closed-loop implementation and validation of ACS

The online closed-loop implementation and validation of DIAMOND-ACS began for Line 4 (Aug 7, 2014) and for Line 3 (Aug 21, 2014) and the other lines will follow. DIAMOND-ACS is ready to be implemented in every ASP line of the Mussalo WWTP.

DIAMOND-ACS is currently running on Mussalo WWTP. Being the first implementation of advanced controllers in the plant and given the short validation time in the DIAMOND project, the main goal in DIAMOND has been achieving primarily a robust ACS implementation in one line of the Mussalo WWTP. Based on the results obtained in Line 3 and Line 4, the DIAMOND-ACS will be implemented in the rest of the lines.
General conclusions of the ACS

General Conclusions for the Advanced Control System

The design and validation of the ACS on the Mussalo WWTP and its integration with the ADAM tool has been possible due to the generality of the ACS control technologies.

The results obtained on the Mussalo WWTP demonstrate the potentialities of multivariate predictive control and statistical process control for efficient and automated operation of wastewater treatment plants through the easy integration of the ACS with the existing plant instrumentation and control systems.

The generality of the technologies and its modularity of the proposed design will allow the inclusion of more complex control structures, the addition of more units in the plant, and higher levels of plant-wide optimization.

Conclusions for the Scientific & Technical results

With the results that have been obtained at three full-scale WWTP, it can be stated that the DIAMOND project proposes a solution for the advanced data management in WWTPs. The solution is based on an efficient and intelligent use of all the available data by a standard centralisation of the heterogeneous data acquired from the different sources, an effective data processing to detect faulty sensor data, an agile extraction of adequate.

The results show that the DIAMOND Systems facilitate the monitoring of plant state by both synthesising all the available data and generating new useful key-process information for plant operators; detect plant data faults and incoherencies, and reconstruct incomplete data so that data reliability and robustness is guaranteed; and optimise plant control based on complete, non-faulty and reliable sensor data. This way, is has been demonstrated that DIAMOND Systems allow plant operators to optimise the global operation of WWTPs in terms of water quality, resources and energy by achieving a model of excellence in the management of all the available data.

The lack of appropriate data management is no longer a bottleneck for a more efficient use of sensors, monitoring systems and process controllers in WWTPs.

1.4. Potential impact

Due to this project, four exploitable systems have been designed, developed, implemented, tested and validated at full-scale. These four products are now technically ready to be commercialized:

(1) The ADAM Tool and the (2) Advanced Monitoring System will benefit MSI through their commercialization.

(3) The Online Data Pre-processing Tool will be commercialized by CERLIC.

(4) The Advanced Control System will benefit MIPRO through its commercialization.

The potential customers of the results are environmental engineering companies, end-users such as public entities and WWTP exploitation companies and automation engineering and sensors manufacturing companies.

For the two End-Users of the project, CAG and SVVAAB, in addition to having special market conditions, energy consumption and treatment costs savings have been achieved from the implementation of the different project results in the Mekolalde and Bromma plants. On the other hand, the Kymen Vesi Oy Mussalo WWTP has also benefited of energy consumption and treatment cost savings.

And as for DIAMOND’s societal implications, they are in line with different socio-economic and strategic objectives defined by the European Union and associated European Water Directives, such as improving the competitive position of the SME project partners from the technological point of view, with more added-value in its products. In addition, they consider minimizing the waste from WWTPs by using advanced data management, monitoring and control tools and algorithms. Also, they take into account reducing the energy consumption at WWPTs via dedicated plant-operational optimization control algorithms, attaining sustainability by protecting water and the developed technology by itself. Finally, they are in line with introducing information and communications technology in all sectors, looking at the system as a whole instead of looking at and optimizing only part of it.

In summary, the gathered knowledge will put Europe in the pole position with respect to tackling the optimization of wastewater treatment processes in a sustainable way.

1.4.1. Main dissemination activities

Although DIAMOND is a research project with restricted access, significant efforts to identify and promote effective dissemination practices and specific actions for the DIAMOND project have been undertaken.

The progress of the project and its results have been disseminated through different channels: the DIAMOND web page, communication of results to broad public (printed and electronic press releases, publications, pieces of news), and participation in industrial and academic conferences related with the relevant subjects of DIAMOND. The effort in these activities has increased once the project results have been available. Among the dissemination activities, it is especially important to highlight the DIAMOND video clip and the Wikipedia webpage of the project and its results.

Since the development of the DIAMOND website (http://www.diamond-eu.org/) on November 2012, it has been the gateway to the public audience throughout the project. Highlights of the research and development results, announcement of events and public deliverables have been posted here. The number of sessions in the website from the 1st of November of 2012 to the 31st of August was 1036, being 776 the number of users, 722 the new visitors and 264 the returning visitors. The visit average duration has been 1 min and 24 s. Internet home users have been the 77 % of the visits, universities the 10 %, companies and others the 10% and research centers the 3%. The most active countries have been Spain (43.34 %), Finland (11.68 %) and Sweden (7.53 %). **Figure 30** shows a screenshot of the main page of the DIAMOND external website.

Figure 30 Diamond external website screenshot (upper part, left hand side) (below part, right hand side)

DIAMOND has been introduced and presented during 13 presentations, conferences and seminars, being the main goal to reach different audiences with various backgrounds and experiences. The main events were the DIAMOND project was presented are the following ones:

1) IWA World Water Conference, held in Busan (Korea) during the 16th to the 21st of September of 2012, title: "Advanced data management for optimising the operation of WWTPs" (oral presentation);

2) 11th IWA conference on instrumentation control and automation (ICA2013), held in Narbonne (France) during the 18th to the 20th of September of 2013, title: "DIAMOND: AdvanceD data management and InformAtics for the optimuM operatiON anD control of WWTPs" (poster presentation, **Figure 31** left);

3) IWA World Water Conference, held in Busan (Korea) during the 21st to the 26th of September of 2014, title: "DIAMOND: AdvanceD data management and InformAtics for the optimuM operatiON anD control of WWTPs" (poster presentation, **Figure 31** right).

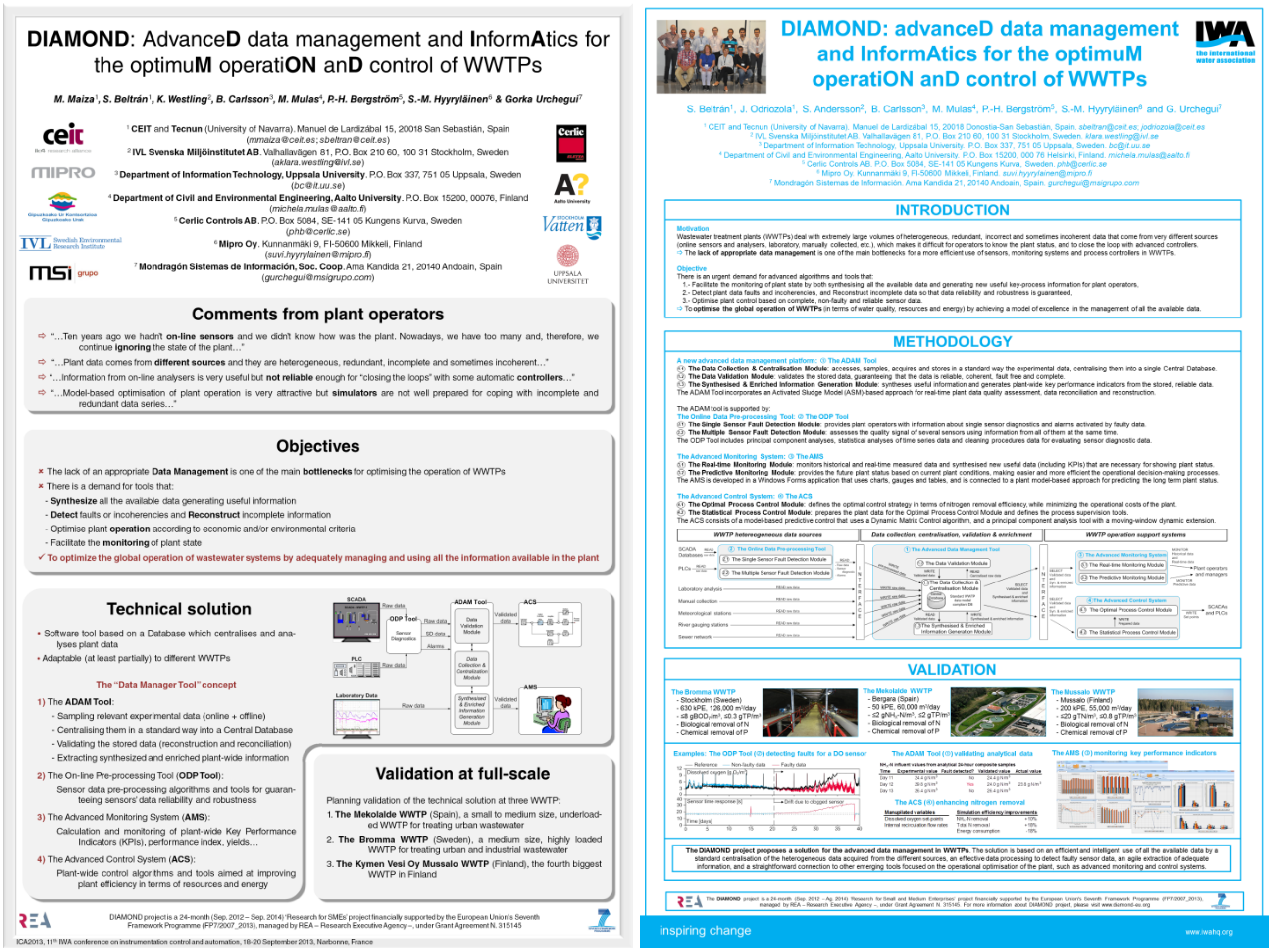


Figure 31 Poster presentations of the DIAMOND project

The scientific and the technical communities have had a key role for the dissemination of the DIAMOND project. Several technical magazines have been used for spreading the DIAMOND project, being the following two ones the most important ones: 1) Retema, Revista Técnica de Medio Ambiente, issue January-February 2013, title "Proyecto DIAMOND. Excelencia en la gestión de los sistemas de tratamiento de agua"; 2) Kuntatekniikka (2013), title "Jätevedenpuhdistamojen data tehokkaampaan käyttöön. DIAMOND-hankkeesta työkaluja puhdistamojen operointiin".

The DIAMOND Consortium has done a huge work regarding the presentation of the project and its final results on the mass media related to water, wastewater, engineering, automation and control. This media was very heterogeneous: webpages, newspapers, technical magazines, newsletters, presentations, exhibitions, thesis, interviews, video clips and posters. The targeted audience has been diverse: scientific community (higher education and research), industry, civil society and medias. This way, the DIAMOND Consortium has written almost 45 pieces of news and press releases of different languages.

Especially important among the dissemination activities is the video clip of the DIAMOND project that has been shot and edited by CEIT. During the 8 minutes and 29 seconds, DIAMOND project is introduced (type of project, consortium and objectives), all the nine partners are then represented and main conclusions are finally drawn. The 20th of September of 2014 the video clip was published into the video-sharing website YouTube under the following link: www.youtube.com/watch?v=QP2L6FTFw9Q. It can be watched under the following link with a quality of 1080p (High-Definition video, HD), 720p (HD), 480p, 360p, 240p or 144p (see **Figure 32**):

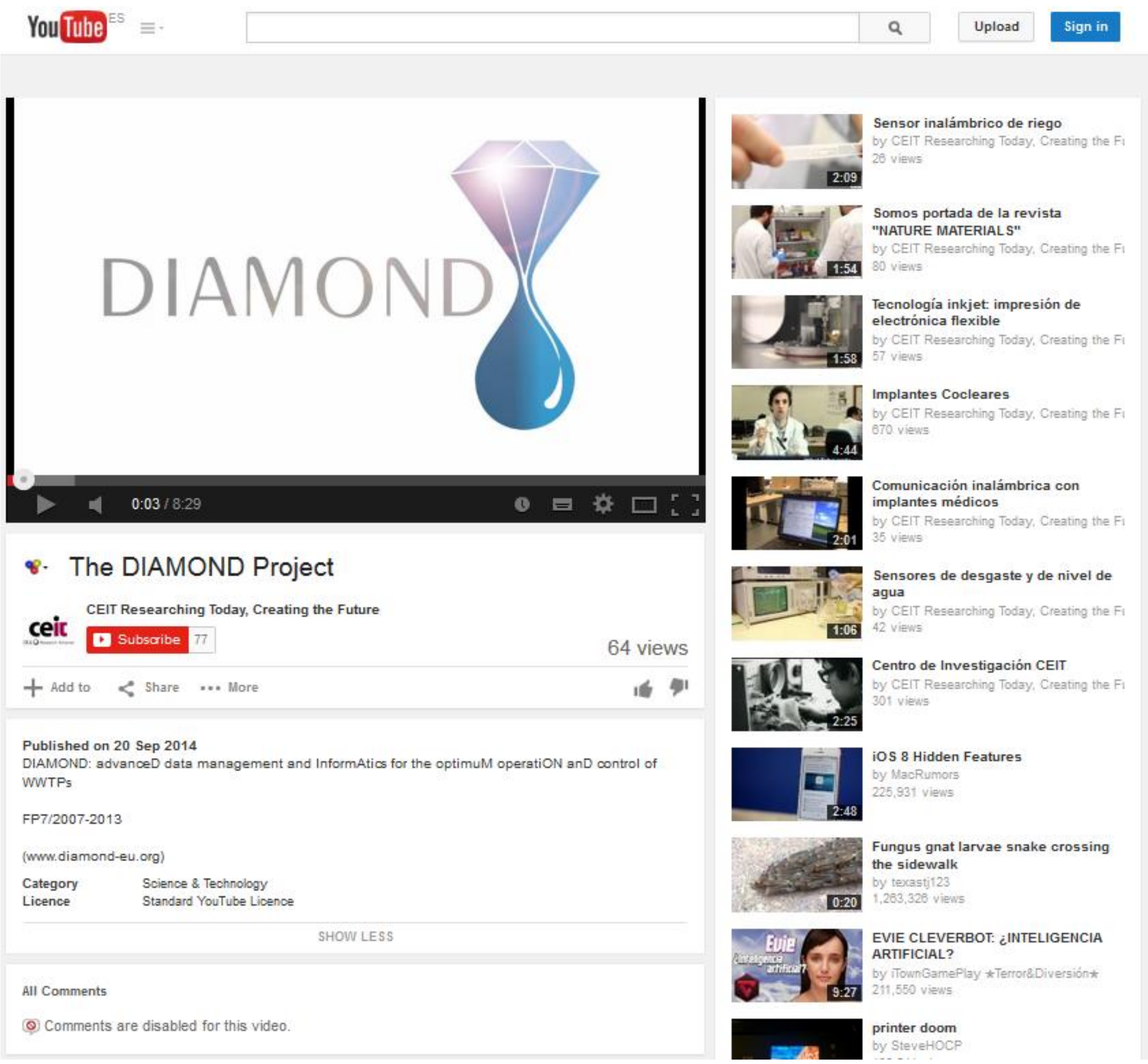


Figure 32 Diamond video clip on YouTube website

The following statistics are provided by YouTube Analytics regarding the DIAMOND video clip (from the 20th of September of 2014 until the 12th of October of 2014): 259 viewers; 639 minutes of visualization; main demographic countries: Spain (87 %), Colombia (2.7 %), México (1.5 %), Argentina (1.5 %), Peru (1.5 %), Finland (1.5 %) and others (4.3 %); viewers: 44 % on YouTube site, 55 % on websites where the video is embedded; main traffic sources: unknown embedded reproduction (55 %), unknown direct access (29 %), external website (5 %), others (11 %); type of device: PC (78 %), mobile phone (21 %), tablet (1.2 %). The DIAMOND video clip is embedded on seven webpages related to water and wastewater treatment.

In addition, it is also important among the dissemination activities the Wikipedia page on the project and its results. The internet encyclopaedia Wikipedia page of the Diamond project (https://en.wikipedia.org/wiki/DIAMOND_%28project%29) has been operational since the 16th of

July of 2014 (46 days before the project finished) to the 4th of September 2014 (4 days after the project finished). The 4th of September the Wikipedia editors closed nearly all 7th Framework Programme projects. In the period it was operational, the DIAMOND project page received 588 visits. **Figure 33** shows a screenshot of the Wikipedia page of the DIAMOND project.

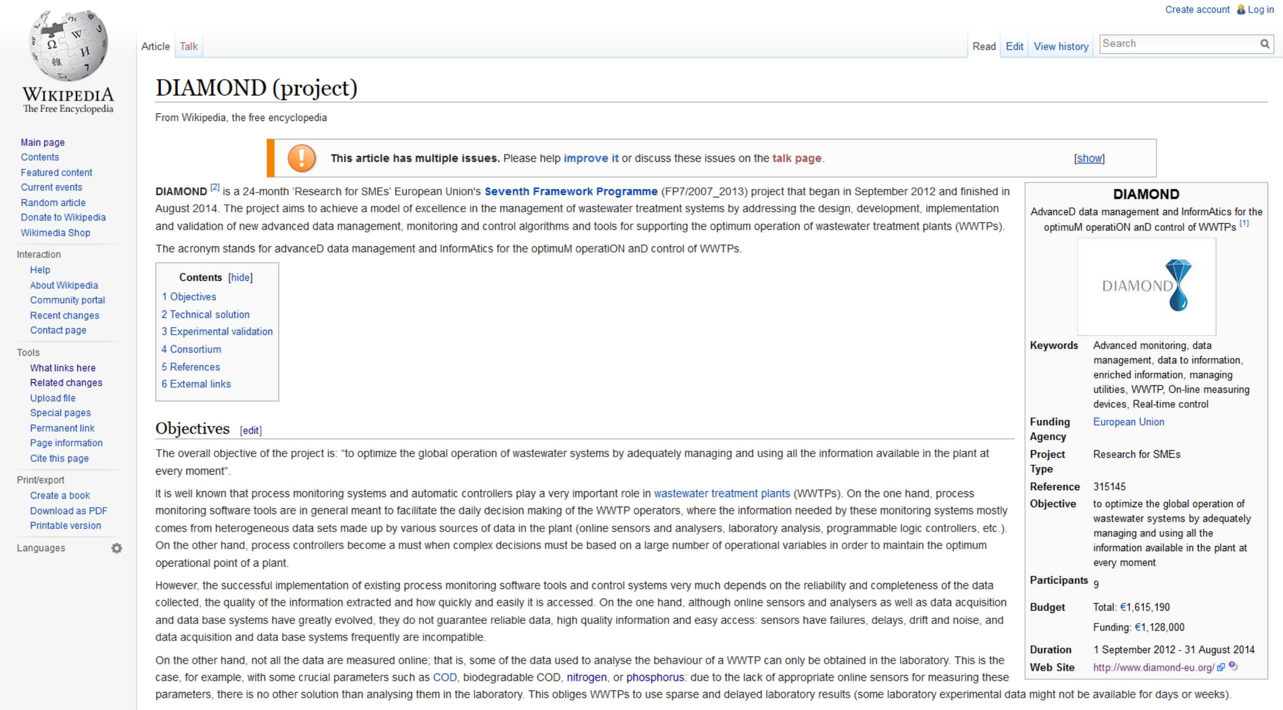


Figure 33 Visual appearance the Wikipedia page of the Diamond project

1.5. Project public website and logo

The DIAMOND logo is shown in
Figure 34, and the DIAMOND project website address is http://www.diamond-eu.org/.

Figure 34 Logo of the DIAMOND project

1.5.1. Consortium

The DIAMOND project has been carried out by nine partners –three small and medium-sized enterprises (SME), two research centres, two universities and two end-users– from three European Union countries (Spain, Finland and Sweden). **Table 2** shows a list of the DIAMOND consortium.

Table 2 List of the DIAMOND Consortium

No	Name	Short name	Beneficiary type	Country	Contact name	Contact email
1	Mondragón Sistemas de Información, Sociedad Cooperativa	MSI	SME participant	Spain	Mr. Gorka Urchegui	gurchegui@msigrupo.com
2	Aguas de Gipuzkoa, SA	CAG	End-user	Spain	Mr. Óscar Fernández	oscarf@gipuzkoakour.com
3	Cerlic Controls, AB	CERLIC	SME participant	Sweden	Mr. Pär-Håkan Bergström	phb@cerlic.se
4	Stockhom Vatten, AB	SVVAAB	End-user	Sweden	Mr. Lars-Gunnar Reinius	lars-gunnar.reinius@stockholmvatten.se
5	Mipro OY	MIPRO	SME participant	Finland	Mr. Seppo Hyvönen	Seppo.Hyvonen@mipro.fi
6	Centro de Estudios e Investigaciones Técnicas de Gipuzkoa	CEIT	RTD performer	Spain	Mr. Sergio Beltrán	sbeltran@ceit.es
7	Uppsala Universitet	UU	RTD performer	Sweden	Mr. Bengt Carlsson	bengt.carlsson@it.uu.se
8	IVL Svenska Miljoeinstitutet, AB	IVL	RTD performer	Sweden	Mr. Fredrik Hallgren	Fredrik.Hallgren@ivl.se
9	AALTO-Korkeakoulusaatio	AALTO	RTD performer	Finland	Ms. Michela Mulas	michela.mulas@aalto.fi

1.5.2. Contact details

Mr. Gorka Urchegui
Mondragón Sistemas de Información Sociedad Cooperativa (MSI)
Ama Kandida Etorbidea 21, 20140 Andoain (Spain)
Tel: +34 943 594 400
Fax: +34 943 590 536
E-mail: gurchegui@msigrupo.com