Sustainable water use in chemical, paper, textile and food industries

Evaluation of sensors and data processing tools

Marie Maurel, VEO

AquaFit4Use is co-financed by the European Union's 7th Framework Programme

Executive summary

This report is a result of the project AquaFit4Use, a large-scale European research project co-financed by the 7th framework programme of the European Union on water treatment technologies and processes.

This report highlights the results of two additional studies:

- The evaluation of water quality sensors and analyzers in order to measure the quality of water in industrial processes, particularly in case of reuse.
- The development of software in order to manage effectively on-line and off-line measurements and specifics modules in order to get high-content information concerning the state of the process.

The first task was to develop a test protocol in order to define criteria of comparison to evaluate sensors. This test protocol allows calculating comparison criteria to benchmark the instruments. In such a project as Aquafit4Use one, where the water quality is one of the main concerns, sensors can stand as a key for many studies (simulation, efficiency of new processes…). Four parameters were chosen to describe water quality among the list established in WP 1.1. Lab-scale tests on substitute waste waters (synthetic samples) are thus performed for:

- 5 Chemical Oxygen Demand (COD) sensors and analyzers,
- 1 Total Nitrogen (Total N) analyzer,
- 2 Orthophosphate (PO_4^{3-}) analyzers,
- 2 Calcium (Ca) sensors.

The tests were executed according to the protocol built in the WP 2.1 and the devices are evaluated and compared by the aid of common criteria about static (accuracies) and dynamic (response times) characteristics. At lab-scale, the instruments based on physical UV absorption are the only devices that meet all the requirements set in the preliminary study for COD measurement. The other devices use more complicated measurement principle based on full spectrum absorption (Spectro::Lyser from S::CAN and CarboVis from WTW) or chemical reactions (PeCOD P100 from aqua Diagnostic) that induce more important data treatments.
Pilot-scale tests are also executed on the COD and Ortho-phosphate sensors and analyzers in order to evaluate the devices in conditions closer to real operating conditions. The results remain poorer than the lab-scale tests because of the real and complex water matrix. Nevertheless, one analyzer gives quite good results given that its calibration overcomes the water matrix.
In a later stage of the project, industrial tests should be executed to complete the evaluation and comparison.

Regarding the data processing software, the interface is developed and tested. This decision support tool aims to improve the user information levels. The architecture used with high level of modularity allows further developments of modules. One module was implemented to make classification and data reconciliation. The method chosen allows making reliable the information brought by measurements and to determine the current state of the process. Moreover, four pre-treatments were implemented to detect sensors faults. A case study was achieved to illustrate the results. At last, a remote access tools was added to the software in order to provide remote web access to the high-level information module software.

Content

1 Introduction

This report is a result of the project AquaFit4Use, a large-scale European research project co-financed by the 7th framework programme of the European Union on water treatment technologies and processes.
The research objectives of AquaFit4Use are the development of new, reliable cost-effective technologies, tools and methods for sustainable water supply use and discharge in the main water using industries in Europe in order to reduce fresh water needs, mitigate environmental impact, produce and use water of a quality in accordance with the industries specifications (fit-for-use), leading to a further closure of water cycle.

For more information on AquaFit4Use, please visit the project website: www.aquafit4use.eu.

1.1 State of the art

1.1.1 Evaluation of sensors

To manage waste water treatment processes in the industry, the use of devices measuring water quality which has to be controlled is needed. The choice of an instrument for a given process is not easy and is often made according to the basic metrological characteristics given by the suppliers. However these characteristics are not always sufficient and may even lead to a wrong choice: most of them do not take into account real operating context with varying conditions. This is why a procedure for the evaluation of sensors and analyzers of water quality was built in WP2.1. This method is a means to obtain "true" data on sensors performances which could be compared to the datasheets from the suppliers. Several approaches already exist: National normalization organizations, like AFNOR in France, wrote document as norm ISO 15839 and ISO 8466 so as to define a test protocol for instrumentation. They well defined lab test protocols and methodologies for sensors evaluation. The first added value from these norms is the definition of common glossary in order to understand the different concepts and promote exchange on this topic.

1.1.1.1 Measurement chain

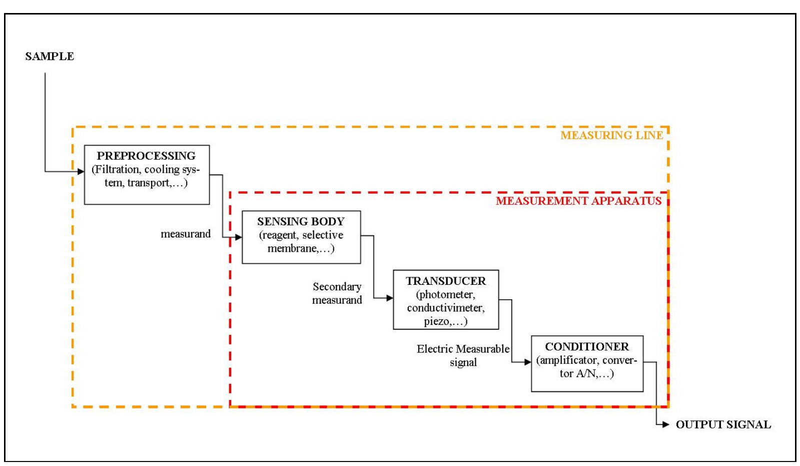


Figure 1: MEAURING LINE - Definition

The operating features are established for the whole of the measuring line: from the "packaging" of the sample to the output signal. Before beginning the study of the measuring chain, it is important to well define this chain and to identify all the elements precisely (see figure 1):

- Sample: Complex Matrix containing the characteristics to measure (or Main measurand).
- Preprocessing: All operations applied to the sample before its arrival to the analysis system.
- Sensing Body: Element which selectively reacts to the variations of the measurand, it has to transform this measurand into another physical parameter known as measurable: secondary measurand.
- Measurand: Parameter which is measured. The whole of the experimental operations which give the knowledge of the numerical value of the measurand is the measure.
- Transducer: It transforms the measurand into measurable electrical signal.
- Conditioner: It transforms the signal of the transducer into electrical or numerical signal exploitable by an operator or an automat.

The following paragraphs describe the main evaluation criteria used in the test protocol.

1.1.1.2 Response time

One of the first criteria useful to evaluate a sensor is the time needed to obtain the measurement value. Some definitions already exist to describe this criterion:

- Response time (RP): Time interval between the moment when the material of the instrument is exposed to an abrupt change of the value of the characteristic to determine and the moment when the reading passes the limits of a band ranging between 90% and 110% of the difference between the initial value and the end value of the abrupt change and remains inside this band.
- Delay time: Time interval between the moment when the instrument is exposed to an abrupt change of the value of the characteristic to determine and the moment when the reading indicates in a continuous way a value higher or equal to 10% of the difference between the initial value and the end value of the abrupt change.
- Rise time: Difference between the Response time and the Delay time when the abrupt change of the value of the characteristic to be determined is positive.
- Fall Time: Difference between the response time and the latency time when the abrupt change of the value of the characteristic to be determined is negative.

1.1.1.3 Linearity

Most of time, the linearity of a sensor is evaluated by the value of the correlation coefficient. Its definition is given below:

$$R^2 = \frac{\left(\sum_1^n (xi - xm)(yi - ym)\right)^2}{\sum_1^n (xi - xm)^2 \sum_1^n (yi - ym)^2}$$

Where xm and ym are respectively the average of xi and yi.

The main concept is that the response of the sensor is a linear function of the concentration of the pollutant, expressed like: Y=AX + B where A and B are:

$$A = \frac{\sum_1^n (xi - xm)(yi - ym)}{\sum_1^n (xi - xm)^2} \quad \text{and} \quad B = ym - A.xm$$

After calibration (automatic or manual), A and B are good indicators of the efficiency of this calibration. A must be closed to 1 and B must be closed to 0.
A and B can be interpreted as an expression of the trueness linked to the calibration. The interpretation of R² is not trivial, the perfect value is 1, and the difference between R² and 1 can be due to the non linearity of the measurement AND/OR to the dispersion of points. Those examples below have the same coefficient R² - one is linear with an important dispersion (left side) and not the other (right side).

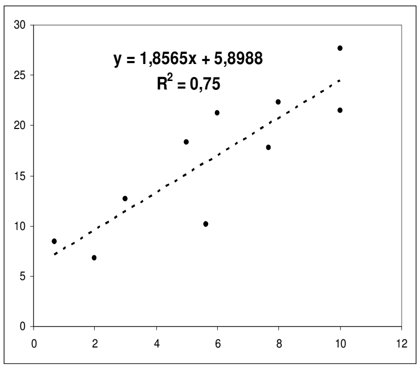


Figure 2 : R² interpretation - Linear dispersed

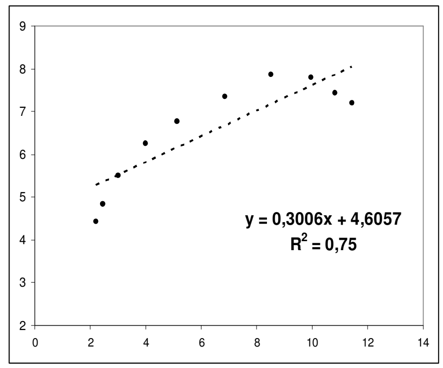


Figure 3 : R² interpretation - not Linear

The coefficient R² is very influenced by the presence of aberrant points. The practice is to eliminate 10% of the most aberrant points – but not those at the extremity of the series. Of course, the choice of these points is subjective. To illustrate this statement, let see the figure below where the 2 series have the same correlation coefficient.

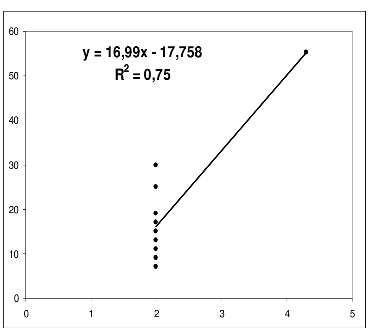


Figure 4 : R² interpretation - with Aberant Point - not Linear

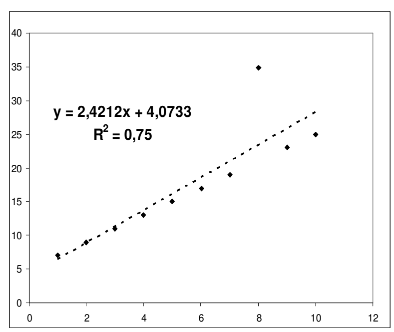


Figure 5 : R² interpretation - with Aberant point - Linear

To conclude about the interpretation of the R² coefficient, it is really needed to see the graph to avoid bad interpretations as illustrated by figures in this paragraph. All the figures have the same correlation coefficient but only 2 are linear relations. Coefficients A and B are linked to the trueness and so to the quality and the efficiency of the calibration process.

1.1.1.4 Accuracy

Accuracy may be split into trueness and precision, where trueness accounts for the closeness of agreement between the mean value and the true value, while precision accounts for the closeness of agreement between the individual values among themselves. The figure 6 below is a good and famous illustration of the difference between precision and trueness.

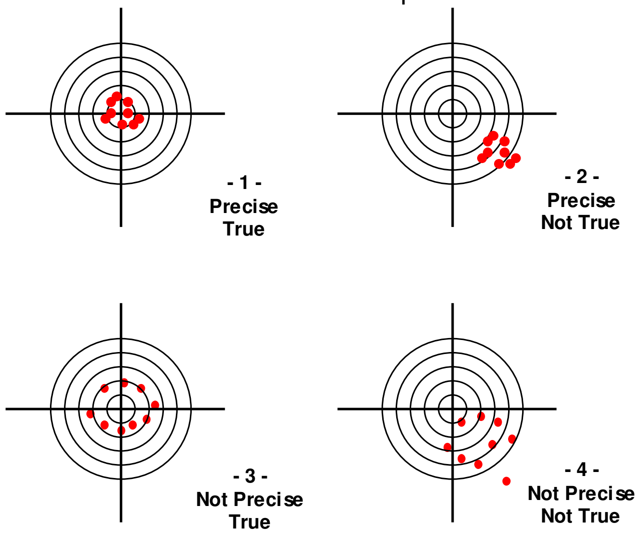


Figure 6 R² interpretation - Accuracy

Two simple mathematic tools are used to study the accuracy: the standard deviation and the average of the series.

The **standard deviation s** is the root of the average of the squared variations:

$$s^2 = \frac{1}{n}\sum_{1}^{n}(xi - xm)^2$$

Where xm is the average of xi

Based on the standard deviation, norm defined other standardize performance criteria:

- Limit of detection (LOD): The lowest value, considerably higher than zero, of the characteristic to be determined which can be detected. To calculate the limit of detection while multiplying by three the standard deviation of $y_{1,j}$ measurements, for J = 1 to 6. This factor 3 is corresponded to 0.13% of risk to conclude that the substance is there then it is not.

  LOD = 3.S$_{5\%}$

- Limit of quantification (LOQ): The lowest value of a characteristic to be determined which can be measured with an acceptable level of exactitude and fidelity. To calculate the limit of quantification while multiplying by ten the standard deviation of $y_{1,j}$ measurements, for J = 1 to 6. This factor 10 is corresponded to 0.5‰ of risk to conclude that the substance is there then it is not.

  LOQ = 10.S$_{5\%}$

These 2 definitions can be discussed. Indeed, they don't take into account the systematic error on zero level. The instrument is considered perfectly linear and calibrated… and that the average of blank measurement is Zero.

- Repeatability: Fidelity under repeatability conditions. In these conditions, independent test results are obtained with short time intervals (for example one day) with the same method, in the same laboratory, by the same operator and using the same material and the same reagents. The repeatability must be studied as well for the high values of the characteristic to determine as for the low values and to record in the form of two different results (index 20 and index 80 respectively) calculated as standard deviation of $y_{2,j}$ measurements, J for J = 1 to 6 and of $y_{3,j}$ measurements, J for J = 1 to 6.

$$\mathbf{R_{20\%} = S_{20\%}}$$
$$\mathbf{R_{80\%} = S_{20\%}}$$

- <u>Smaller Detectable Change (SDC):</u> Smallest difference being able to be measured significantly between two measurements. To determine the smallest detectable change as well for the high values of the characteristic to determine as for the low values and to record in the form of two different results (index 20 and index 80 respectively) calculated while multiplying by three the standard deviation of $y_{2,j}$ measurements, for J = 1 to 6 and of $y_{3,j}$ measurements, for J = 1 to 6.

$$\mathbf{SDC_{20\%} = 3.S_{20\%}}$$
$$\mathbf{SDC_{80\%} = 3.S_{20\%}}$$

- <u>Bias:</u> Regular variation of the value measured compared to the reference value. To determine bias as well for the high values of the characteristic to determine as for the low values and to record in the form of two different results (index 20 and index 80 respectively) obtained by calculation of the difference between the average value of $y_{2,j}$ measurements, for J = 1 to 6 and the x_2 value and of the difference between the average value of $y_{3,j}$ measurements, for J = 1 to 6 and the value of x_6.

$$\mathbf{B_{20\%} = M_{20\%} - X_{20\%}}$$
$$\mathbf{B_{80\%} = M_{80\%} - X_{80\%}}$$

Where M is the average of the measurement and X the reference measurement.

The calculations given by the norm ISO 15839 seems to be redundant and could be simplified in order to be applied and easily used for industrial applications. They are not fitted for real onsite application especially for industrial water/wastewater plant. For instance, it is not possible to screen easily the whole range of measurement needed for the application because this value depends only on the process state. Moreover, the measured value could be different in case of a process drift or failure. In this last case, we should provoke a failure to apply the evaluation procedure and it is impossible to access to all these previously defined criteria. To conclude, link to this background, Veolia Environnement worked to define a sensors test methodology taking into account the operational constraints.

1.1.2 Data processing software

There is a lot of commercial software for data acquisition. These software are powerful but not optimized for “process” specialised data acquisition and are not flexible to integrate the know-how for real-time data analysis, the most important issue for the task of the WP2.1. The following state of the art could not be exhaustive; nevertheless, it describes the most relevant existing software with their main functions.

1.1.2.1 IHM complementary tools

<u>PcVue Solutions from ARC Informatique</u>

PcVue Solutions provides a suite of software and hardware products for displaying, controlling, managing and analyzing the information from installations. PcVue is SCADA software for monitoring and control. There are features, add-ins and tools to handle communications, networking and database management.

<u>DataVue: analysis of historical data:</u> DataVue runs queries on PcVue's historical data held in proprietary format: trends, logs and log reports. It displays the results in a Microsoft Excel spreadsheet, where they can be formatted for presentation as a report, graphic or pie chart. Results can also be exported in ASCII or Windows Clipboard formats. DataVue caters for predefined queries with user defined criteria (data, event type, value, variable properties etc.). Statistical calculations can be performed as the data are extracted (Fig. 7).

<u>Dream Report: real-time report generator:</u> Dream Report collects and exploits all data accessible from any data source to generate reports dynamically. User can access and interact with information coming from industrial systems such as Supervisors (SCADA), real-time drivers, OPC servers and all other data sources. This offers dynamic, real-time generation of reports, automatic distribution and a management engine with which users can protect and administer access to each of the modules (Fig. 7).

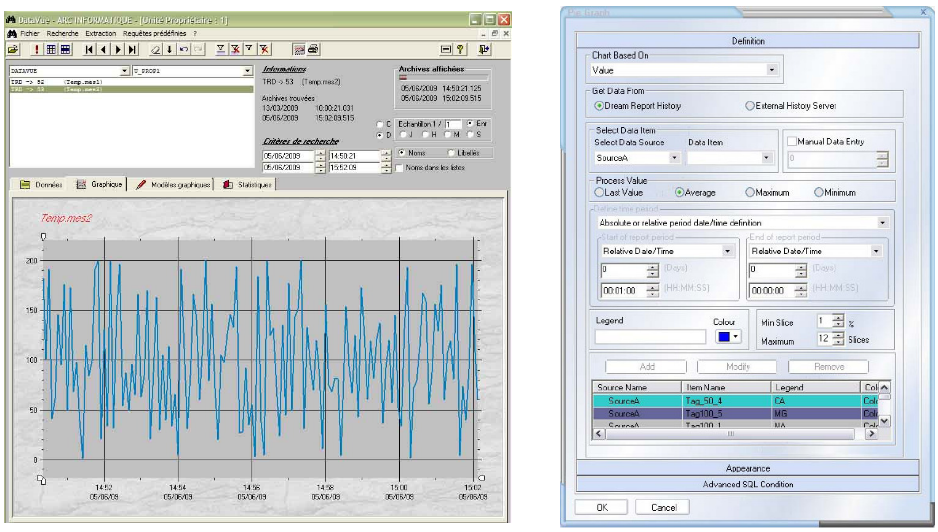


Figure 7 DataVue and Dream Report

<u>Wizcon from Elution</u>
Wizcon provides tools to build operator interfaces and supervisory control applications.

<u>Multi-variable trend charts track real-time and historical data:</u> Users can view and analyze both real-time and historical data on Wizcon's trend charts. Wizcon's charts provide graphical views of process behaviour and operational trends over a period of time. A single chart window can display up to 16 process parameters. The user can freely scroll the charts over the time axis or zoom in on the time interval to see rapid changes in the process behaviour. Wizcon's charts provide linear and logarithmic scales (Fig. 8). The user can change the chart from online to historical mode to present historical information for a set of process parameters.

<u>X-Y charts:</u> For in-depth analysis of dependencies among process parameters, the user can configure the chart to an X-Y chart. The X-Y chart presents up to 15 process variables as the functions of another process parameter.

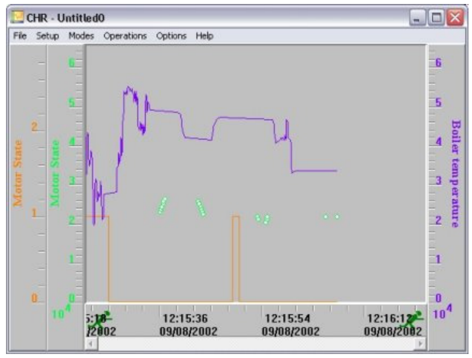


Figure 8 Wizcon chart

Intouch from Wonderware

Wonderware HMI Reports: Wonderware HMI Reports is a reporting solution. It can create reports from Wonderware InTouch HMI, and many other data sources. Wonderware HMI Reports requires no IT or programming skills to design, schedules and produces informative dynamic reports (Fig. 9).

Wonderware HMI Reports allow users to transform raw data into great looking, information-filled reports. Wonderware HMI Reports works with Wonderware Information server or an included web portal to extend the availability of reports throughout enterprise network.

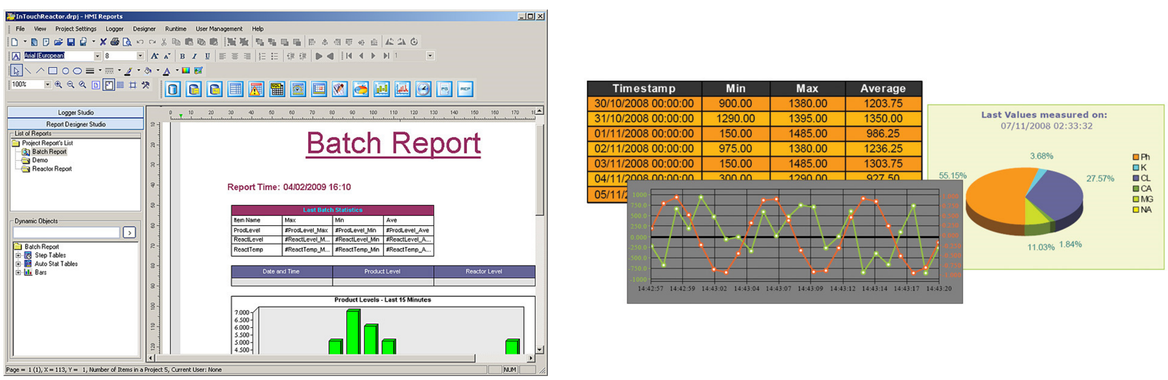


Figure 9 Wonderware HMI Reports

Wonderware Historian: The Wonderware Historian component is a real-time database for historical information. Wonderware Historian is designed to collect a wide variety of plant data. Advanced data retrieval modes enable to generate the detailed, focused information needed to accelerate the decision-making process and provide access to the right information when a problem is identified or an opportunity uncovered (Fig. 10).

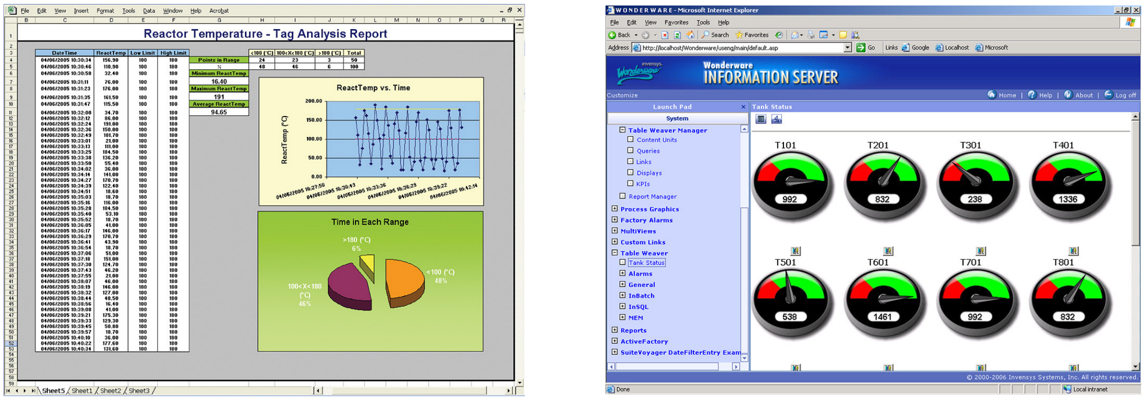


Figure 10 Wonderware Historian

1.1.2.2 Other standalone software

<u>Matlab from the MathWorks</u>

<u>Data Acquisition Toolbox:</u> Data Acquisition Toolbox software provides a set of tools for analog input, analog output, and digital I/O from a variety of PC-compatible data acquisition hardware. The toolbox lets user configure external hardware devices, read data into MATLAB and Simulink environments for immediate analysis, and send out data.

Data Acquisition Toolbox enables to customize acquisitions, access the built-in features of hardware devices, and incorporate the analysis and visualization features of MATLAB and related toolboxes into design. User can analyze or visualize data, save it for post-processing, and make iterative updates to test setup based on analysis results (Fig. 11).

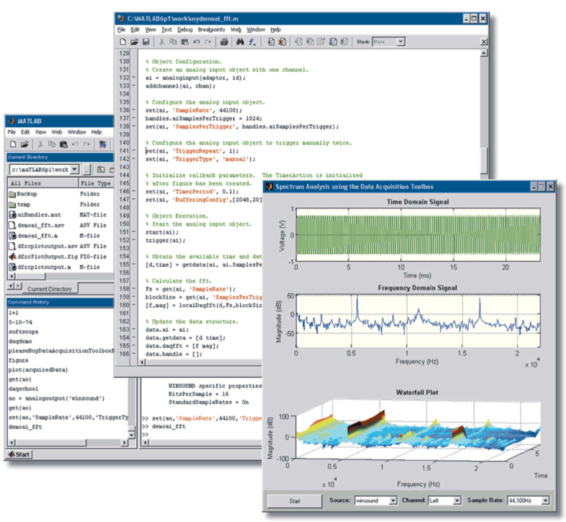


Figure 11 Matlab Data acquisition toolbox

Figure 12 LabView

<u>LabView and NI-DAQ from National Instruments</u>

National Instruments offers a complete family of data acquisition hardware devices and software that extends too many languages, buses, and operating systems (Fig. 12).

1.1.2.3 Evaluation

Tab.1 presents an evaluation of theses software.

Table 1 Software evaluation

	PcVue	Wizcon	InTouch	Matlab	LabVue
Data format	Proprietary	Proprietary	Proprietary	Open	Proprietary
Acquisition flexibility	No	No	No	Yes	Yes
Data export	Yes	No	Yes	Yes	Yes
Data treatment (pre / post)	No	No	No	Yes	Yes
Developer competence	-	-	-	Yes	Yes

All these software are not able to provide all required qualities needed:

- An open data format or a database in order to export data to other software (Excel, Matlab, etc.).
- Acquisition configuration flexibility, to allow all users to configure themselves data sources, recording frequencies, etc.
- An ability of data pre-treatment, to reduce data quantity, and post-treatment, to produce high-content information to user
- No developer competence needed for software installation and configuration
- Modularity, in order to add process know-how as “plug-in”.

This is why it was decided to develop a solution for data acquisition, using existing toolboxes for low-level communication, in order to be able to implement, in a modular and evolutionary way, our know-how in term of real-time and specific data treatment from processes.

1.1.3 Methods for data reconciliation

To achieve data reconciliation, the proposed strategy is based on analysis of available information which comes from sensors of the process. Methods using historical data were studied. The strategy takes place in two stages:

- Off-line step: historical data are analysed to characterize known comportment of the system. A learning method has to be chosen.
- Both process behaviour, obtained during the first step, and on-line data are used to determine the current state of the process. This is called the recognition step. Situation can be known (normal functioning or faults identified during the learning stage), or unknown (new process state). In the case of an unknown situation, the recognition method must be able to update the model behaviour of the process.

The principal features of the different supervised classification methods are presented in the table 2.

Table 2 Comparison of supervised classification methods

Classification method	**K nearest neighbour**	**Classification tree (Decision tree)**	**Neural networks**	**SVM (Support Machine Vector)**	**LAMDA**
Data type	Numerical	Numerical	Numerical & Qualitative	Numerical	Numerical & Qualitative
Need of a known set of learning	Yes	Yes	Yes	Yes	Yes/No
A priori allocated classes	Yes	Yes	Yes	Yes	Yes/No
Fixed number of initial classes	Yes	Yes	No	Yes	No
Parameters to be regulated	K		activation function, number of hidden layers of the network and, number of neurones by layer	hyper-parameters C and ε and the kernel parameters	MAD function, connective and α
Classification type	strict	strict	strict	strict	strict & fuzzy
Classes update	No	No	No	No	Yes

By looking at this table of comparison, the LAMDA method was chosen. This classification method is presented in the following chapter.

1.1.4 Remote access tool

There are many ways of remotely accessing to IT equipments. It will be necessary Internet connection in both sides, apart from a middleware (middleman software or service), among other staff in some cases. But as the range of IT solutions is wide, the most important point here is to decide which the main access purpose is: control the device, just act in a read-only way, change its state, file transferring, data access, etc. In the following section, a quick review (tab. 3) of the most popular remote access commercial software is presented, paying attention to the possibility that any of them can fit to the required solution for the data processing tool.

Table 3 Remote supervision tools evaluation

Remote access tool	**Basic**	**Subscription / cloud-based**	**VPN**	**Server-based bundle**
Remote access	Yes	Yes	Yes	Yes
Web access	Not all	Yes	Yes	Yes
Security (encryption)	Yes	Yes	Yes	Yes
Remote screen blanking / Keyboard locking	No	Yes	No	Yes
Chart oriented features (historical data checking)	No	No	No	No
Multiple concurrent user access / actions	No	No	No	No

All these remote access type of tools are able of giving access to the target computer. However, this access should not be as meddling as they end up being. To illustrate this, suffice it to say that if one user connects remotely to a PC, any operator / developer / technical staff who was using this equipment will not be able to work or do anything more in this computer but looking at the screen (or disturb one each other). Another not covered situation is the multiple remote accesses; if more than one user need to check some information in the remote equipment, a satisfactory action is not possible. In this way, the establishment of the minimum required qualities for the appropriate remote access tool should include:

- Remote access to the data / information being served.
- Multi-user (more accesses than one at the same time).
- Transparent access, not monopolizing the use of the remote computer against a local user
- Web based, what ensures it to be done from any platform, any device …
- Chart / Graphic features to, for example, draw historical data (no need of having any additional staff in the supervised PC).

Considering all the above information, the remote WWTPs supervision tool, AqquaScan, developed by ATM was the tool chosen as the Web access system for the software modules to be developed within the WP 2.1.

1.2 Objectives

1.2.1 Evaluation of sensors

The general objective for this study is to execute test studies for sensors evaluation on pilot plant or labscale. In detail, the sub-objectives set by Veolia Environnement for this study are the following:

- Define sensors benchmark and criteria of comparison
- Select relevant and representative sensors and analyzers of water quality
- Apply the protocol of evaluation on these devices at lab-scale
- Benchmark these devices according to the results
- Execute complementary tests at pilot-scale

1.2.2 Data processing

Data acquisition and data treatment are key issues for process modeling and survey. Firstly, it is necessary to record measurement and operational process condition to be able to calibrate and validate modeling stage. Secondly, effective on-line data treatment and acquisition software development will allow the integration of faults detection procedures or algorithms for process survey. In case of recycling water, the efficiency of water treatment processes should interfere on industrial production line and become a constraint: for instance in pulp paper industries, if the water quality decrease, the paper quality could be affected. The aim of this work is to provide standard software architecture to manage effectively on-line and off-line measurements and to develop specifics modules in order to get high-content information concerning the state of the process. At last, a web remote access to collect and process data has to be added in order to enable remote data processing.

2 Methods

2.1 Methods

2.1.1 Evaluation of sensors: test protocol

The experimental approach is adapted to the time, the needs and tools available. A standardized report of results should allow for a large number of readers to understand and exploit the results of the tests. The test protocol is divided into 2 parts: laboratory tests and on-site tests. Laboratory test is a way to evaluate the performances of the instrument on synthetic samples. The industrial test is the way to evaluate the behavior of the measuring line in real conditions.

2.1.1.1 Pre-required information

The test protocol must be fitted with the targeted application: what is the mission of the instrument? The application is defined by the description of the process and the location where the measuring line will be installed. The necessary operating range must be settled for the application. Especially for industrial application, the environmental constraints must be considered and will be taken into account in the second part of the test: the on-site tests.
It is important to consider the different roles that an instrument can have. There are 3 classes of role. This classification isn't exhaustive but allows keeping in mind that the applications for an instrument can be very different:

- **Auto-control**: the most important quality is the precision of the measurement. Indeed, it can be used to meet the standards (at the end of a waste water treatment plant for instance) and the measured value must be the reference.
- **Process control:** the most important quality is the availability. Indeed, this one is used for process control and stop the instrument may have bad consequences on the process. The accuracy of the measurement is less important than the one before but the dynamic of the measure must be correctly detected.
- **Early warning:** the implementation of sensors must be dense (for instance, for a potable water network: every 10 km) and no maintenance is possible. The accuracy can be neglected because the aim is to detect a crossing of a limit and not the exact value of the measure.

In order to materialize the needed reflection before starting the test, it's important to complete as precisely as possible the questionnaire. Examples of questionnaire will be presented in the results.

2.1.1.2 Laboratory tests

Based on the state of the art, the laboratory tests is an application of the main concepts describe in the different ISO norms. It could be compare as a pre-study in order to check the theoretical capabilities of the instrument. The retain procedure try to simplify the actual norm.

i. Dynamic characteristics

The values of the characteristic to measure solutions allowing the abrupt change are 20% and 80% of the operating range. The preliminary response time is given by doing the abrupt change of the solutions (from 20% to 80%). This time is used to set the time delay for the next sequence. Then, the abrupt change between 20% and 80% of the measuring range is done 6 times (6 rises & 6 falls). The settling time before each change is 3 times of the preliminary response time. Be careful to not pollute the standard with the precedent one. Dynamic characteristic times are then the average of the times determined during these 6 rises and 6 descents (see figure 13).

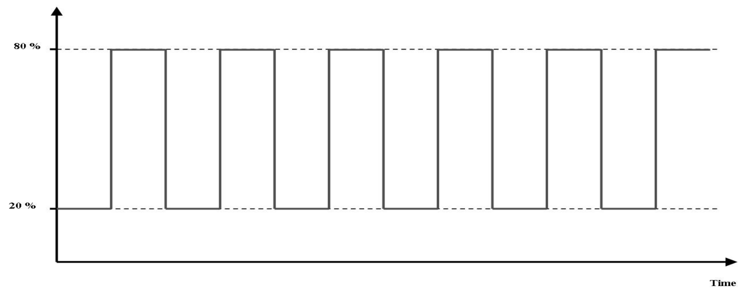


Figure 13 : Protocol for Dynamic Study

ii. Study of linearity

The analysis of the linearity is done by comparing the value of the characteristic to be measured reading on the apparatus and the value of a standard solution. These couples (standard / Reading) are taken in range 20%-80% of the range of operation with a constant step on this range (see figure 14). There must be at least 5 points measuring range (advised: 10 points).

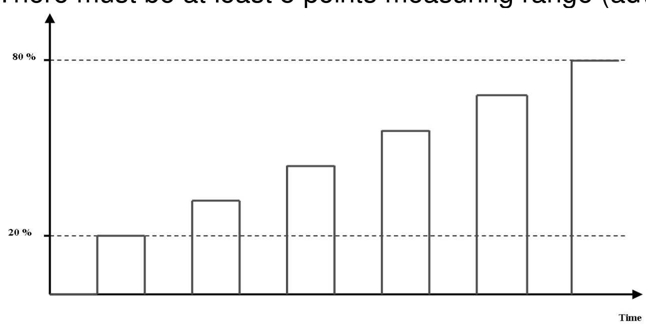


Figure 14 : Protocol for Linearity Study

Do a linear regression with the couples of results (References in Y-axis and reading in X-axis) without forcing the intercept to the origin. The correlation coefficient gives an account of the linear dependence between the read values and the reference values of the standards.
If it is possible to calibrate the instrument, this study of linearity must be done just after this calibration. The calibration must be done by following the recommendations of the supplier.
2 criteria can be calculated with the linear test:

- **Offset:** Average bias between the measured value and the reference one in the low values of the studying range. The offset is the constant coefficient (Y-intercept) of the linear regression calculated upon the reference measure expressed in function of the device measure and obtained from the linearity tests.
- **Error in slope ((Slope-1)*SR):** Average bias between the measured value and the reference one in the high values of the studying range (SR). The error in slope is calculated by multiplying the studying range and the deviation of the slope (slope – 1) of the linear regression calculated upon the reference measure expressed in function of the device measure and obtained from the linearity tests.

iii. Study of accuracy

Two notions, precision and trueness, are studied by doing repeated tests on the same sample. The table below synthesizes the necessary results for the calculation of the "static performances". The values (5%, 20%, and 80%) are expressed in % of the operating range and results must be expressed in the unit of the measurement. <u>The instrument must see a blank between each measurement.</u>

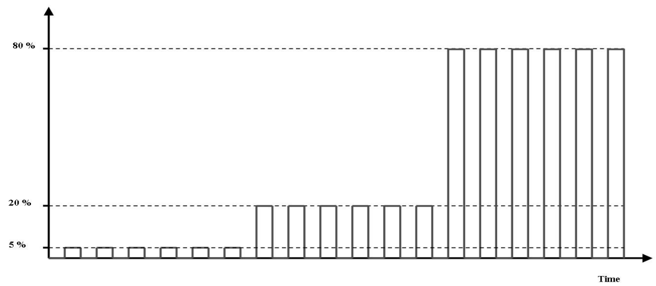


Figure 15 : Protocol for Accuracy Study

<u>Reminder and choice of the most important value:</u>

Table 4   Static performances Results

NAME	SYMBOL	CALCUL
Standard deviation at 5%	S5	-
Standard deviation at 20%	S20	-
Standard deviation at 80%	S80	-
Limit of Detection	LOD	3 x S5
Limit of Quantification	LOQ	10 x S5
Repeatability at 20%	Rp20	S20
Repeatability at 80%	Rp80	S80
Smallest detectable change at 20%	SDC20	3 x S20
Smallest Detectable Change at 80%	SDC80	3 x S80
Bias at 20%	B20	average
Bias at 80%	B80	average

Table 4 reminds all the calculation given by the norm ISO 15839. There are some redundancies as said on the state of the art. To not perturb the reader and to simplify the reading of the results, we decided to choose the most significant values (highlighted in grey on the table). The 3 values calculated with the standard deviation (LOD, SDC20 and SDC80) can be seen as the thickness of the line of the graphic; under these values, the change is not significant. Static performances will be summarized in a graphic as figure 16 below. For representation issue, the absolute values were considered for the bias. This kind of graphic is very useful to compare several instruments.

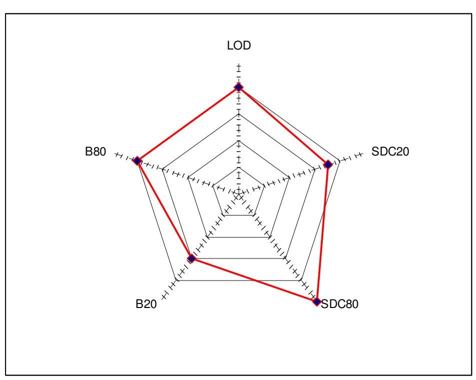


Figure 16: static performances presentation

On-site Study

The on-field tests complement of the laboratory tests. Two families of information are accessible by the on-site study: the accuracy of the instrument in real conditions and the necessary maintenance under real exploitation conditions. If several apparatus are to be tested for the same application, it's better to do it at the same time to compare the behavior of all the instruments. Another detail for the environment of the on-field test is the competence of the staff which takes care of the instrument. For this kind of study, the danger is to not have a well-adapted staff in term of education. It's possible to obtain very different conclusion for on-site experiment depending on whether the technician is experienced or not. In this case, even if the measurement could be very good, the conclusion of the test could be wrong.

There are several ways to evaluate the behavior and the accuracy of an instrument in real conditions. In most cases, it is very difficult to cause variations and/or control the characteristic to be measured in the plant in order to reproduce laboratory tests. Thus, the on-field tests will be done in comparison with a reference measurement associated to a sampling. From here, two points have to be study with attention: the accuracy of the reference measurement and the response time of measuring chain.

i. The reference measurement

The protocol for the reference measurement must be defined precisely. This protocol – defined by the Norm associated to this measurement or by the manufacturer – is the way to have a right measurement and a common method for all the operators. In the report, the accuracy of the reference measurement must be written, both from the supplier datasheet and from the analysis of a repetitive test on the same sample by several operators: two measurements on a synthetic sample (concentration known at 20% and 80% of the measuring range) and one on the field. These last tests give 3 confidence intervals objectives: 2 linked with the measurement principle or the apparatus and the third linked to the perturbation of the real environment. A good example is the DPD measurement of free chlorine in a cooling tower with often a high value of combined chlorine: the error due to the presence of combined chlorine is linked to the rapidity of the operator to do the measurement. Of course, it's important to be sure that all the operators have exactly the same way to measure.

<u>Confidence interval</u> is calculated as follows: X is the average and s the standard deviation of a set of measurement (synthetic sample at 20% and 80% and real sample).

The 99.7% confident interval is: I = [X-3.s ; X+3.s].

It means that the probability to have the measurement X in the interval I is 99.7%. The average due to the sample method must be taken into consideration.

ii. Response time of the whole measuring chain

Lab tests allow calculating exactly the response time of the instrument. For the industrial study, it's important to know exactly the response time of the whole measuring chain. In addition to the response time of the instrument, we have to consider the delay due to sample preparation and transport. Mostly, the sampling delay is mainly due to the sample transport. It's often easy to evaluate this delay with a calculation like Delay=Tube_volume x flow. For more complex situation, it's possible to use tracer like salt with a conductimeter to evaluate sample preparation delay.

It's important to exactly describe where is the sample point. The best is to sample closest as possible to the instrument. The delay or the average due to the sample method must be taken into consideration.

iii. On Field Error study

During the on field test, comparisons to the reference measurements have to be done. For the previous laboratory tests, the systematic errors (trueness) and the random errors (precision) were considered. This split of the accuracy was possible because of the repeated tests on the same

sample. With an on-field study, it's impossible to do this kind of analysis. What is considered constant in the laboratory test becomes dynamic in the long term study. The error to consider is the deviation of the measurement. Indeed, during the laboratory test, the systematic error (or trueness) is constant and in a long term study it is linked to the time: this is the long term drift. Long-term drift is the slope of the straight regression line obtained starting from a series of differences between the values of reference and the measured ones during the on-field test, and expressed in the form of percentage of the operating range over one period of 24h.
With this definition, the period for the study of the long-term drift has to be set. The Maintenance period (MP) is defined as follows: time between 2 successive sequences of maintenance requiring the intervention of an operator (change of reagent is not necessary the limit of a maintenance period). It is clear that during on-site test, the availability and the attention of the operator managing the measuring chain must be higher than those during a real operation. It is important to not distort the result of the test by a surplus of maintenance. Operator's actions must be only the essential ones. One way to solve this problem is to define the Maximum Allowed Difference (MAD) for the application (example: ± 0.1 ppm for free chlorine in potable water). Thus, the maintenance must be done as soon as the error is higher than the MAD. Another way is to respect the recommendation of the supplier and to do not take care of the apparatus more than the recommendation.
In case of a cyclic behavior of the characteristic to be determined - daily for NH_4 in Aeration tank inlet, shorter for sequenced Aeration Tank… - attention should be paid not to make correspond the frequency of comparison to this possible cycle and thus always compare the same value. When a cycle is well identified, a follow-up of this cycle has to be carried out (10 measurements on this cycle) with 3 recoveries at the time of Maintenance Period: just after maintenance, in the middle of MP and right before maintenance.
In case of not cyclic behavior, it is recommended to have at least 20 points of comparison during the period of maintenance.

2.1.2 Evaluation of sensors: choice of parameters of interest

Water quality involves many parameters which have to be controlled. A list of relevant parameters for industrial waste water from textile, paper, chemical and food industries was set in SP1.
This study is focused on four particular parameters that are chosen in this list as representative of water quality in order to test various sensors and analyzers:

- Chemical Oxygen Demand (COD)
- Total Nitrogen (Total N)
- Ortho-phosphate (PO43-)
- Calcium (Ca)

Only the results on COD sensors will be presented in this report as a case study of the test protocol.

2.1.2.1 Chemical Oxygen Demand (COD)

The Chemical Oxygen Demand (COD) is the measure of the mass of oxygen required to oxidize the organic and inorganic matter present in one liter of solution. COD is a measure of the pollution load of wastewater. Moreover, COD is a way to evaluate the Total Organic Carbon (TOC) of an effluent which is a factor of pollution as a source of eutrophication in case of discharge and which interferes with the cleaning function of water. COD is expressed in mgO_2/L.
Total COD is divided in two categories:

- Soluble COD: related to the part of carbon dissolved in the water.
- Particulate COD: related to the particulate carbon in suspension in the water.

For this study, total COD is considered.

COD is the main surveillance parameter used by the European legislation. For instance it is used for the design and the evaluation of treatment capabilities of wastewater treatment plants (WWTP). In the industry, waste waters always contain organic carbon which has to be removed. Therefore COD is interesting at all stages of many processes as an indicator of efficiency and potentially as a control parameter of these processes, whether water is to be discharged or re-used.

2.1.2.2 Choice of substitute wastewaters for COD

Lab-scale tests require substitute wastewater to evaluate the comparison criteria. For every parameter chosen, the corresponding substitute wastewater has to be a standard solution:

- reproducible in any laboratory,
- usable by everybody aiming at testing their own sensors,
- as representative as possible for industrial waste water.

The difficulty lies in the fact that wastewaters are extremely variable depending on the nature and quantity of the compounds.
Several synthetic solutions already exist to represent wastewater with a certain level of COD. Three of them are listed in the table 5 below:

Table 5 Main existing synthetic samples for COD

Designation	Composition	Aspect	Initial purpose
BrenntaPlus	Acetic acid (25-50%), alcohols, sugars, proteins	Dark brown lightly viscous solution	Doping of biological processes
Canadian Recipe	Potassium biphtalate, phosphate (KH_2PO_4, Na_2HPO_4)		Calibration of UV-visible spectrophotometer at 254 nm
OCDE	Soy peptone, viandox (salty sauce with a little meat extract and flavoured), sodium propionate, ethanol, sodium chloride	Turbid dark brown solution	Respirometry tests

OCDE solution is chosen for this study as substitute wastewater for two main reasons:

- Contrary to BrenntaPlus and the Canadian recipe, the composition of OCDE is not only an association of chemical compounds but aims at simulating a waste water sample with several ingredients which are source of COD.
- Although the absorption spectrum of OCDE solution is slightly different from the one of raw waste water sampled in Maisons-Laffitte (France) as shown below in Figure 17, OCDE solution shows an absorption level closer to the one of the waste water in the range 200 – 300nm which corresponds to organic matter.

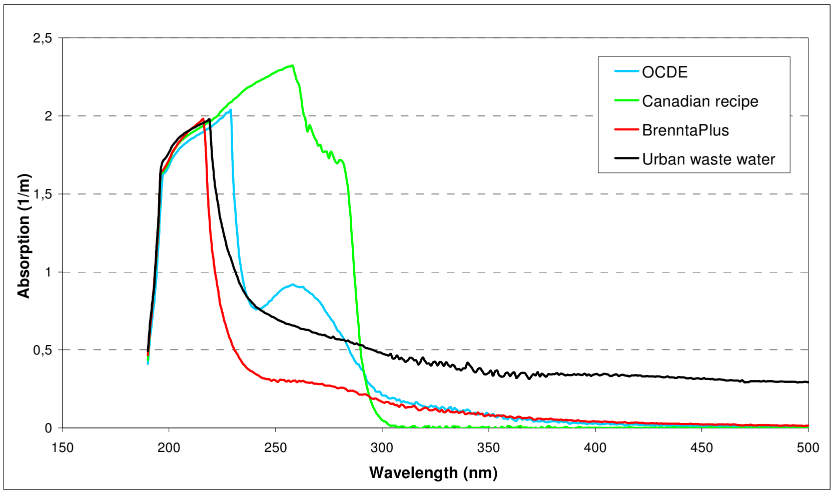


Figure 17: Absorption spectra of three substitute wastewaters

The formula of OCDE is given is the **Error! Reference source not found.**Table 6 below:

Table 6 Ingredients of OCDE solution

Ingredient	Quantity (g for 1L)
Ethanol	9,9
Sodium propionate	13,8
Soy peptone	48,0
Viandox (water, sugar, salt, dye, sodium glutamate, meat extracts, aroma, yeast extracts, milk proteins, lactose, citric acid, lactic acid)	33,0
Sodium chloride	2,1
Ultra pure water	Top up to 1L

The concentration of a few chemical compounds and parameters of OCDE is given in the table 7 below:

Table 7: Chemical composition of OCDE solution

Parameter or compound	Concentration (mg/L)
COD	84 500
Ammonium	163,5
Nitrate	58,7
Nitrite	4,55
Phosphate	354,5
Sulphate	32,7

The targeted COD concentration in then obtained by dilution of the initial OCDE solution.

2.1.2.3 Reference measurement

According to ISO 15705, COD is defined as the volume of oxygen equivalent to the mass of potassium dichromate ($Cr_2K_2O_7$) reacting with the oxidable materials in the water. Mercury, silver sulfate and sulfuric acid are indicated as auxiliary reagents. The reaction time is 2 hours at 148℃.
The standard method is the following:

- the water sample is oxidized at 148℃ with potassium dichromate in excess, silver sulfate, mercury (II) sulfate and sulfuric acid.
- the excess of potassium dichromate is then measured by titration with ammonium iron (II) sulfate (Mohr's salt) in presence of ferroin ($C_{36}H_{24}FeN_6O_4$).

The reference method chosen for the measurement of COD is the Hach Lange micro-method:

- Kit LCK 114 for the range 150 – 1000 mgO_2/L
- Kit LCK 414 for the range 5 – 60 mgO_2/L

Figure 18: Heating block, spectrophometer and reagent kits for the Hach-Lange COD micro-method

This method is based is the same principle as the standard method:

- a 2 mL water sample of the water to be analysed is oxidized at 170℃ during 15 minutes with potassium dichromate in excess in an acid medium: the ion Cr^{6+} is reduced to ion Cr^{3+} which gives a green color to the solution;
- the green coloring is measured with the aid of a Hach Lange DR 5000 UV-visible spectrophotometer: the absorption at a specific wavelength is proportional to the COD.

2.1.3 Data processing

The data acquisition software was developed using very high modularity method, in order to be able to plug different modules for data pretreatment and reconciliation.
The main software (core module) is able to access to

- Common module for low-level communication with equipments.
- Common module for database access (read and write).
- Specifics modules for data acquisition.
- Specifics modules for process control and operator information.
- Specifics modules for data pretreatment and reconciliation.
- Reusable mathematical libraries for all data processing.

2.1.3.1 Core software

The core software is the main HMI (Human Machine Interface) and is in charge of user interface and information, software configuration management, data structures management, accesses to

common modules for low-level communication and database management, dynamic accesses (using programming models) to specifics modules and mathematical libraries.
Data structures:

- Data object : Low-level source, Value, Timestamp and Quality (Good / Bad)
- Data set: Collection of Data objects, with Timestamp and methods for collection management (adding, removing, clearing, counting, etc.)
- Data table: Collection of Date sets, with methods for collection management (adding, removing, clearing, counting, etc.) and for mathematical functions (average, sum, min, max, etc.)

2.1.3.2 Low-level Communication

The communication protocol used is OPC (Ole for Process Control) which is a standard communication protocol now used by automation industries. This protocol can be used in development environments like Microsoft Visual Studio. This protocol is developed independently, in order to be adapted for other communication protocol if necessary.
The communication protocol used between the data processing software and AqquaScan is OPC as well since AqquaScan had relied on an OPC XML DA client driver.

2.1.3.3 AqquaScan architecture

The modularity of AqquaScan software architecture assumes perfectly the supervision of facilities using the data processing tool as a new WWTP:

- The controller module is the piece of software that collects data in a customized way for anything to be read, or modifies any operational value.
- The engineering module manages and serves not only the information in the database, but also the whole events to be warned.
- The Web page is one of the client modules available. This is the friendly Web interface for the data processing tool.

All the modules interact among the rest ones using Web Services, a set of protocols and standards oriented to exchange data between applications (figure 19).

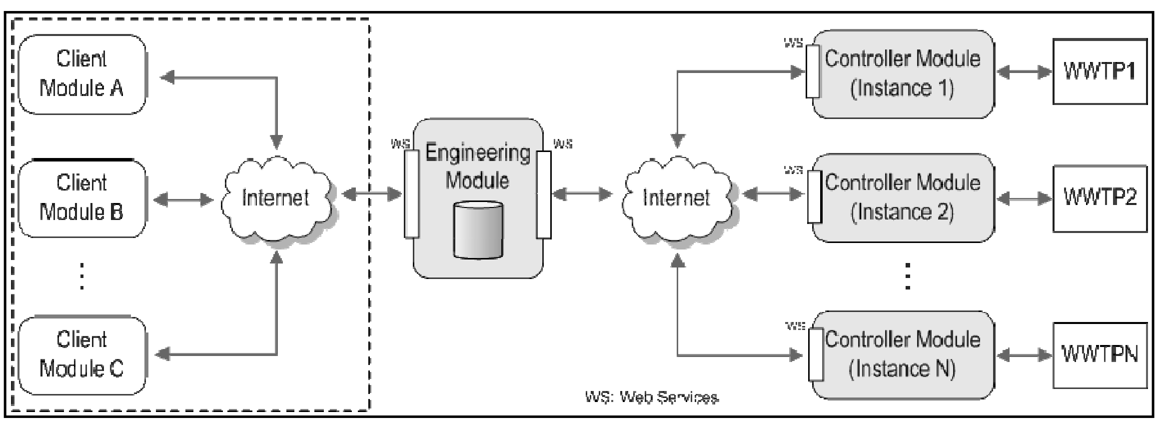


Figure 19 AqquaScan software architecture

2.1.3.4 Database management

Data from processes are stored in a database for local and distant access and post-treatment. An independent data management module is developed in order to switch easily between different

database formats (Microsoft SQL server, Microsoft Access, etc.). In a first time, the Microsoft Access database format has been chosen.
For AqquaScan, data is stored in the object-relational database management system PostgreSQL. Although it is free and open source software, it is supposed to be one of the most robust DBMS and it is ready for managing large amounts of data.

2.1.3.5 Specifics modules and mathematical libraries

All specifics modules and mathematical libraries can be developed independently, and recognized by the core software without having to recompile it. The core software is giving heritable models for such developments. All mathematical libraries must be developed with standard methods and objects, in order to be reused without the core software, for off-line data processing.

2.1.3.6 LAMDA Classification

The LAMDA (Learning Algorithm for Multivariate Data Analysis) methodology is a classification technique introduced by Aguilar-Martin in the early 1980s, developed by Piera Carreté and al. and studied within LAAS (Laboratory for Systems Analysis and Architecture) of CNRS, Toulouse.
LAMDA is a fuzzy methodology of conceptual clustering and classification. It is based on the adequacy concept. The idea is to determine the adequacy degree of an individual to each of the existing classes. This adequacy is obtained from the analysis of a relation between each characteristic of a given object and the respective parameters of each class. This contribution is called the marginal adequacy degree (MAD). MADs are combined using "fuzzy mixed connectives" as aggregation operators in order to obtain the global adequacy degree (GAD) of an individual to a class.

LAMDA is a classification method for multivariate data, using learning and pattern recognition. It addresses to elements sequences, characterized by a fixed number of quantitative or qualitative descriptors.
The LAMDA classification technique can be used according to three different strategies:

- <u>Unsupervised learning (self-learning):</u> there is no previous knowledge of any class, there are no pre-defined classes, and the only existing class is the NIC (Non Informative Class); the first learning is made by calculation of the MADs of descriptors to obtain the best partition of these objects. There is creation of classes and storage of their parameters. This partition depends on the valuation method of the MADs and of the way of aggregating them (through connectives). The relevance of the classification is judged by the expert.
- <u>Supervised learning:</u> the user has previously set some classes and decided to which class each element must be assigned to; from this set, he can create other classes in using those already created. Indeed, it’s possible to form new classes and to evaluate the parameters of the old classes.
- <u>Pattern recognition:</u> Classes are known and recovered well the universe of elements. In this case, a set of objects is allocated in these classes. During the allocation, the classes’ parameters are not modified. Not classified elements are allocated to a class named “residual class”, which has the same characteristics that the NIC class in the learning case.

These three strategies can be complementary. Usually, the user begins by an unsupervised learning and then in using created classes he can get the best partition. All these classes are used to do pattern recognition.

2.1.3.7 Pre-processing

This phase is primarily aimed at preparing the data in a suitable and useable format, so that a knowledge extraction process can be applied.

Four pre-treatments have been chosen to detect sensors faults in data:

- Null variance error: gradient considered as null on a determined number of successive points.
- Temporal error: the step of time between two measures is not respected.
- Gradient error: too high gradient (in absolute value) between two measures
- Min-Max error: value inferior to the minimum or superior to the maximum

On the graph below, an example of each pre-treatment is given: one can see the variables and for each error, a stair when a fault is detected.

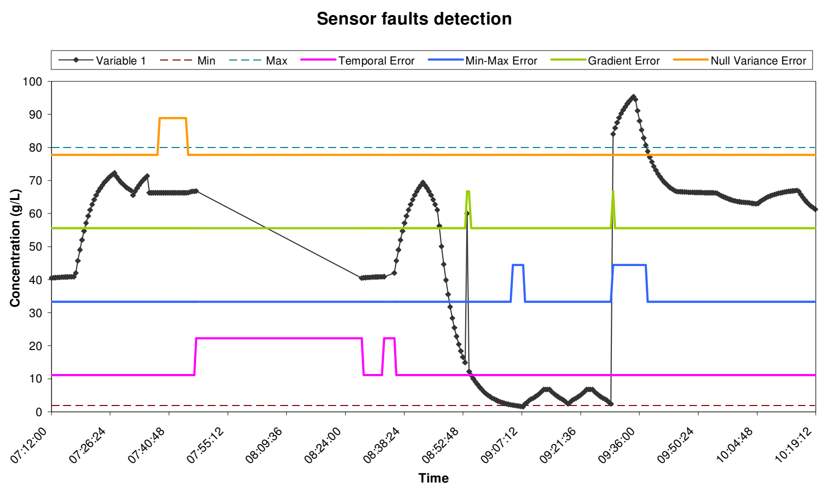


Figure 20 Example of sensor faults detection

Each pre-treatment is customizable with one or two parameters:

- Null variance error: gradient value which will do to consider two consecutive values as no different and the maximum number of successive null gradient beyond which we consider that the variance of the variable is null.
- Temporal error: the time step.
- Gradient error: gradient value beyond which it is considered as too high.
- Min-Max error: the minimum and the maximum of the variable.

Only if no fault is detected, the variable is sent to the classification.

2.2 Materials and equipment

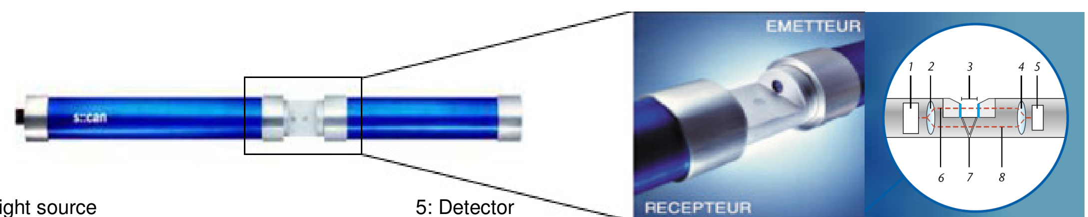
2.2.1 Evaluation of sensors

The added value of this work is that it is not only an office and bibliographic work but it has been done with technicians doing the tests. The main default of the norm ISO15839 is that any example of application is given even if the approach is good. This part details the main characteristics of the COD devices chosen to be evaluated in this study.
Three in-situ COD sensors and two COD analyzers are selected to be tested according to the lab-scale protocol defined above.

2.2.1.1 Spectro::Lyser from S::CAN

Table 8 Description of Spectro::Lyser from S::CAN

Designation	Spectro::Lyser
Supplier	S::CAN
Type of instrument	In-situ probe
Type of measurement	Continuous
Dimensions	650 x 44mm (length x max. diameter)
Measurand(s)	COD [mgO_2/L] Nitrates Nitrogen dioxide Turbidity Suspended matters
Measuring range for COD	0,1 – 3 600 mgO_2/L
Measurement principle for COD	UV-visible spectrometry (220 – 720 nm)
Working for COD	The sensor contains a broadband light source that emits a beam through a slit to a photometric detector. This beam is crossed by the water to analyse: the absorption by the organic carbon at specific wavelengths is then measured and analysed with turbidity and suspended matters compensation.
Calibration	Device calibration with three user solutions and a blank.
Others	The slit is automatically cleaned with compressed-air injection.
Price per unit	23 000 €
Comments	Contrary to an electrochemical sensor, this device does not contain any membrane of mobile part.



1: Light source
2: Sender of the optical system
3: Measuring gap between the measurement windows
4: Receiver of the optical system
5: Detector
6: Measuring beam
7: Cleaning nozzles of the compressed-air cleaning system
8: Reference beam

Figure 21: View of Spectro::Lyser from S::CAN

2.2.1.2 CarboVis from WTW

Table 9 Description of CarboVis from WTW

Designation	CarboVis
Supplier	WTW
Type of instrument	In-situ probe
Type of measurement	Continuous
Dimensions	650 x 44 mm (length x max. diameter)
Measurand(s)	COD [mgO_2/L] Biological Oxygen Demand (BOD) [mgO_2/L] Spectral Absorption Coefficient (SAC) [m^{-1}]

Measuring range for COD	0,1 – 800 mgO_2/L
Measurement principle for COD	UV-visible spectrometry (200 – 750 nm)
Working	The sensor contains a broadband light source that emits a beam through a slit to a photometric detector. This beam is crossed by the water to analyse: the absorption by the organic carbon at specific wavelengths is then measured and analysed with turbidity and suspended matters compensation.
Calibration	Device calibration with one user solution and a blank
Others	The slit is automatically cleaned by compressed-air injection.
Price per unit	13 000 €
Comments	The Carbovis is made by S::CAN and commercialised by WTW with its own treatment algorithm.

2.2.1.3 Uvas from Hach Lange

Table 10 Description of Uvas from Hach Lange

Designation	Uvas
Supplier	Hach Lange
Type of instrument	In-situ probe
Type of measurement	Continuous
Dimensions	70 x 329 x 333 mm
Measurand(s)	Spectral Absorption Coefficient at 254 nm (SAC 254) [m^{-1}]
Measuring range for COD	0,1 – 600 m^{-1} (~ 0 – 7000 mgO_2/L)
Measurement principle for COD	UV spectrometry at 254 nm
Working	The sensor contains a light source that emits a beam at 254 nm through a slit to a photometric detector. This beam is crossed by the water to analyse: the absorption by the organic carbon is then measured and analysed with turbidity compensation (absorption measurement at 550 nm).
Calibration	Manual calibration to convert m^{-1} into mgO_2/L
Others	The slit is automatically cleaned by a mechanical lens wiper. Uvas does not directly give a level of COD but UV absorption at a wavelength of 254 nm with turbidity compensation, expressed in m^{-1}. The signal given by the sensor must be converted into mgO_2/L.
Price per unit	6 000 €
Comments	

Figure 22: View of Uvas from Hach Lange

2.2.1.4 TresCon SAC OS210 from WTW

Table 11 Description of TresCon from WTW

Designation	TresCon, SAC OS210 module
Supplier	WTW
Type of instrument	On-line analyzer
Type of measurement	Batch, quasi-continuous (up to two measures per minute)
Dimensions	612 x 775 x 329 mm
Measurand(s)	Spectral Absorption Coefficient at 254 nm (SAC 254) [m^{-1}] Nitrate [mg/L] (others possible)
Measuring range for COD	0,1 – 200 m^{-1} (~ 0 – 2000 mgO_2/L)
Measurement principle for COD	UV spectrometry at 254 nm
Working	The analyzer samples the water to analyse with a peristaltic pump. The sample is filtrated to remove the suspended matter then it is sent to the measurement cell where UV absorption at 254 nm is measured.
Calibration	Daily automatic zero-calibration on a blank (demineralised water) and manual calibration to convert m^{-1} into mg/L
Others	
Price per unit	7 000 €
Comments	The Trescon is not dedicated to the measurement of COD: the measuring cells can be chosen by the user depending on the parameter to be measured.

Figure 23 View of whole TresCon analyzer (left) and SAC OS 210 module (right) from WTW

2.2.1.5 PeCOD 100 from Aqua Diagnostic

Table 12 Description of PeCOD P100 from Aqua Diagnostic

Designation	PeCOD P100
Supplier	Aqua Diagnostic
Type of instrument	On-line analyzer
Type of measurement	Batch
Dimensions	215 x 340 x 225 mm
Measurand(s)	COD [mgO_2/L]
Measuring range for COD	Several ranges are possible depending on configuration: 0 – 20 mgO_2/L 20 – 150 mgO_2/L 150 – 1500 mgO_2/L 1500 – 15000 mgO_2/L
Measurement principle for COD	Photocatalyst
Working	The analyzer samples the water to analyse with a peristaltic pump. The organic species of the sample are oxidized with titanium dioxide (TiO_2) and UV activation: the photocurrent charge originating from the chemical reaction is then measured by an auxiliary electrode (see Figure 24)
Calibration	Automatic calibration with benchmark samples.
Others	
Price per unit	25 000 €
Comments	The PeCOD P100 is an innovative analyzer that still requires development.

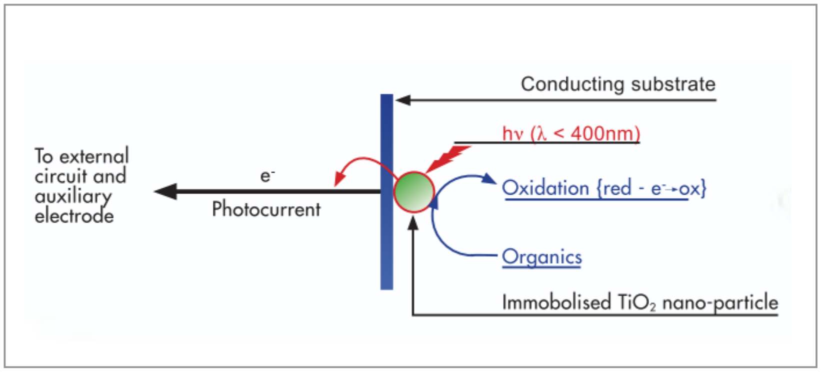


Figure 24: View and measurement principle of PeCOD P100 from Aqua Diagnostic

2.2.2 Data processing software

2.2.2.1 OPC Data Control

For the development of the low-level communication module, the OPC communication protocol has been chose. The commercial product used is the Software Toolbox OPC Data ActiveX Control, which is usable with Microsoft Visual Studio (Fig. 25). The ActiveX Control is included in the software with an ActiveX container and can access to OPC servers (like RSLinx for Allen-Bradley PLCs, or OFS for Schneider PLCs) via an Ethernet TCP/IP network. It can also use proprietary protocol to access directly to PLC, with an OPC server component or a multi-protocol gate.

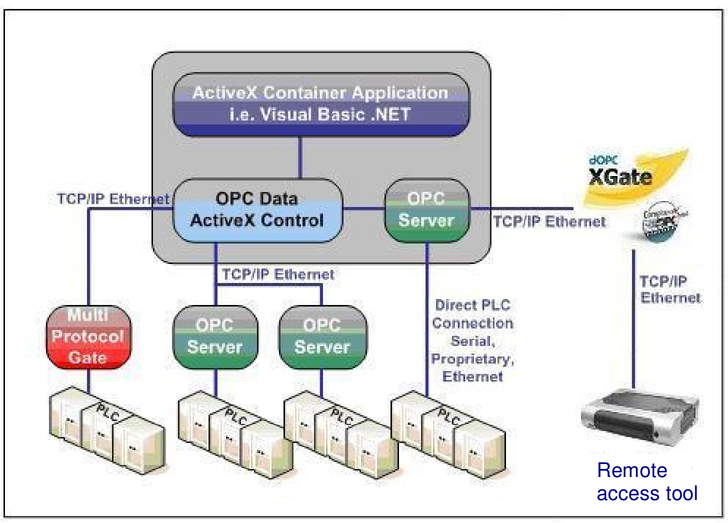


Figure 25 Low-level communication with OPC Data ActiveX Control

As the data processing software implements an OPC Server, based on OPC COM DA, in its low-level communication module, the remote access tool is able to connect to it using its OPC XML DA driver. But commercial middleware is needed, since this OPC COM DA specification is tied to

Microsoft platforms and a free Linux distribution is the Operative System running on remote access software. This middleware is Kassl dOPC Xgate and it acts as a proxy.

2.2.2.2 Programming language and Framework

The development language used is Microsoft Visual Basic .NET 2008, part of the latest Microsoft Visual Studio .NET Integrated Development Environment (IDE), and using Microsoft .NET framework 3.5.

Visual Studio is the Integrated Development Environment in which developers work when creating programs in one of many languages, including Visual Basic, for the .NET Framework. The .NET Framework is a development and execution environment that allows different programming languages and libraries to work together seamlessly to create Windows, Web, Mobile, and Office applications. Visual Basic 2008 is an evolution of the Visual Basic language that is engineered for productively building type-safe and object-oriented applications. Visual Basic enables developers to target Windows, Web, and mobile devices.
Existing common and specifics modules are developed with this language, but future developments can be realized with any other language supported by this IDE (like C# or C++).

2.2.2.3 Testing environment

The software testing environment is:

- Pilot plant for wastewater treatment, with online sensors,
- Allen-Bradley PLC for automation (sensors and actuators),
- ARC Informatique PCVue SCADA system,
- RSLinx professional for OPC communication with PLC
- dOPC Xgate server gateway

The specifications for the PC used for the software installation are:

- Standard PC with Intel processor and 2Gb RAM
- Microsoft Windows XP SP2

The specifications for the miniPC used for the remote supervision are:

- On-board VIA C3/Eden, VIA Eden Processor 533MHz, 512Mb RAM
- Linux Debian Etch.

3 Results and achievements

3.1 Major results and achievements

The major results and achievements for this study are:

- The evaluation of sensors:
  - 5 COD sensors were tested at lab scale and pilot scale. The lab scale tests were done thanks to the test protocol: comparison criteria were calculated for all the sensors and analyzers tested so as to benchmark the devices. At this stage of the test, 2 devices met the requirements for the chosen application.
  - The results of the complementary pilot-scale tests enable to compare the behaviour of the sensors on more complex water matrix.
- Data processing tool: The software is developed and is able to plug to any system with OPC protocol and optional gateway. The high level of modularity allowed developing modules for on-line and off-line data treatment. A remote web access to high-level information module has been successfully provided.

3.1.1 Evaluation of sensors

3.1.1.1 Lab-scale results for COD measurement

i. Preliminary study for COD

Application	
Name of the application:	COD regulation or control
Type of industry:	Paper, textile, food and chemical industries
Name of the process:	-
Type of Application:	Process Control
Measurand	
Measurand:	COD
Unit:	mgO_2/L
Range:	0 – 1000 mgO_2/L
Maximum Allowed Difference:	20 mgO_2/L
Time	
Smallest Characteristic Time of the Process:	A few hours (depending on the process)
Maximum time needed for change detection:	10 min
Situation	
Quality of the sample:	Raw industrial waste water after pre-filtration
Place in the process:	Outlet (or inlet) of a process
Preprocessing:	Transportation, filtration (for the inlet application)
Reference	
Reference Measurement:	Hach Lange micro-method
Norm:	ISO 15705
Apparatus:	Kits LCK 114 and LCK 414
Reference Measurand 1:	COD
Unit:	mgO_2/L

Accuracy:	~ 5 – 10 mgO_2/L

ii. Comparison of performances of COD sensors and analyzers

Table 13 Performances coefficients of COD sensors and analyzers

Criterion (perf)	Spectro::Lyser (S::CAN)	CarboVis (WTW)	Uvas (Hach Lange)	TresCon (WTW)	PeCOD100 (Aqua Diagnostic)
RT	-6.7	-1.6	-5.6	-1.4	2.0
LOD	1.6	*10 (-)*	*-5.1*	-5.8	-2.74
B20	-3.1	4.1	-2.1	-1.7	2.3
B80	4.7	4.4	-4.7	-1.1	-1.1
SDC20	2.4	4.9	-2.1	*-10 (-14.9)*	1.9
SDC80	2.5	5.4	-1.3	-6.8	3.3
Offset	1.8	4.6	-2.7	-1.9	1.3
(Slope-1)xSR	*10 (15.6)*	5.9	-10.0	-2.0	-1.3

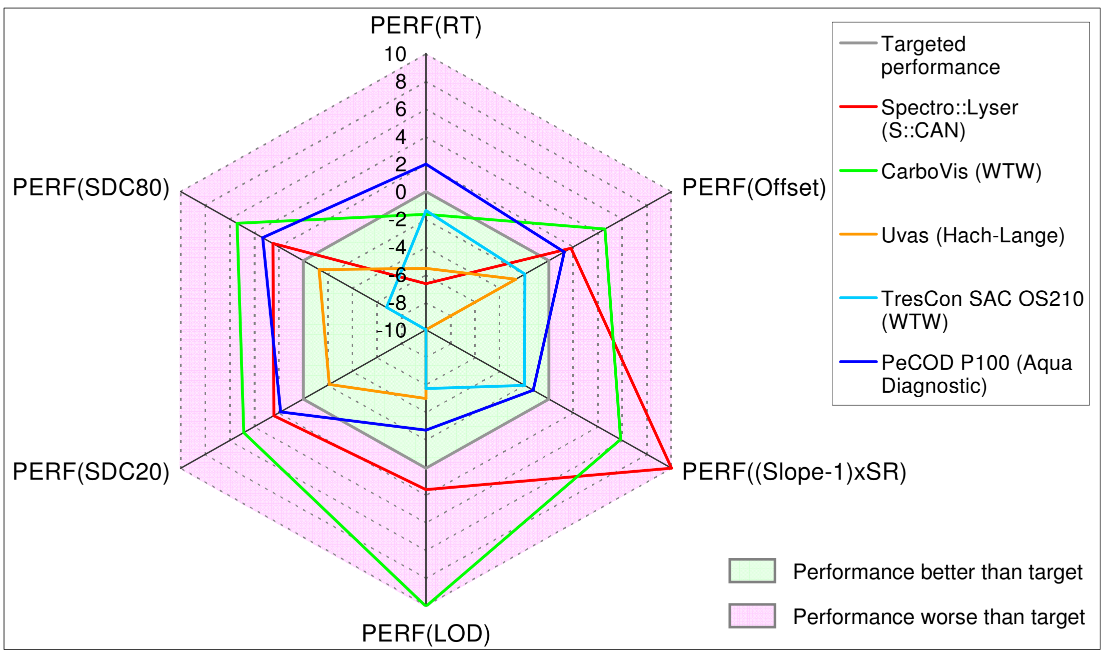


Figure 26: Comparison of performances of COD sensors and analyzers

This diagram is a way to compare the sensors and analyzers in a metrological point of view only at lab-scale.
Uvas from Hach Lange and TresCon from WTW are the only devices that meet all the requirements set in the preliminary study. Both apparatus are based on the same measurement principle which seems to be – at least at lab-scale – the most effective: indeed it is physical (UV absorption) and uses a very limited number of wavelengths to work, with a simple compensation for turbidity that does no induce any delay time.
The three other devices use more complicated measurement principle based on full spectrum absorption (Spectro::Lyser from S::CAN and CarboVis from WTW) or chemical reactions (PeCOD P100 from aqua Diagnostic) that induce more important data treatments.

3.1.1.2 Complementary results at pilot-scale

Complementary tests at pilot-scale are performed for the 5 COD sensors and analyzers
Two series of pilot-scale are run:

- online tests on a real urban waste water effluent,
- offline tests on real industrial waste water samples.

i. Online tests on urban waste waters

For this series of pilot-scale tests, the sensors and analyzers to be tested are exposed to real waste waters at different stages of the pilot treatment line in the water research centre of Veolia Environment (Maisons-Laffitte, France), depending on the technical possibilities:

- Decanted waste water: waste water after lamellar settling
- Filtrated waste water: waste water after simple filtration
- Treated waste water: waste water after biofiltration (Biostyr®)

Different calibration methods are tested:

- Waste water calibration (pilot calibration): the device is calibrated with samples of waste water from the effluent to be measured.
- Supplier calibration: the device is calibrated with a specific reagent provided by the supplier of the apparatus.

Measurement with urban waste water calibration

In these tests, the devices are calibrated with samples of the urban waste water to analyze then are exposed to the urban waste water for a few hours without recalibration. The Figure 27 and Figure 28 present the comparison of the measures from the devices with waste water calibration and the reference measures.

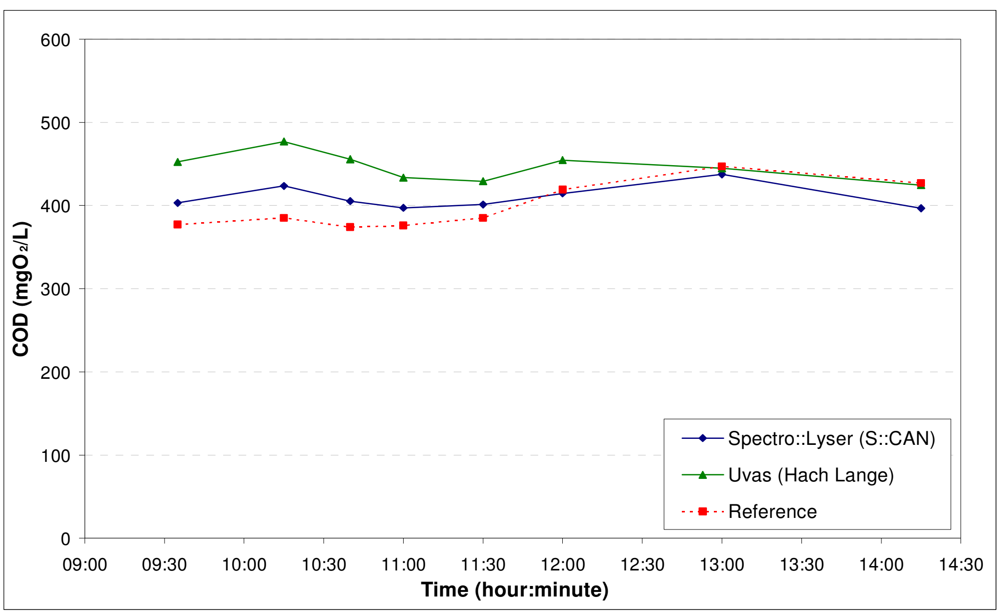


Figure 27: COD measurement on clarified water. Spectro::Lyser (S::CAN) and Uvas (Hach Lange. Waste water calibration

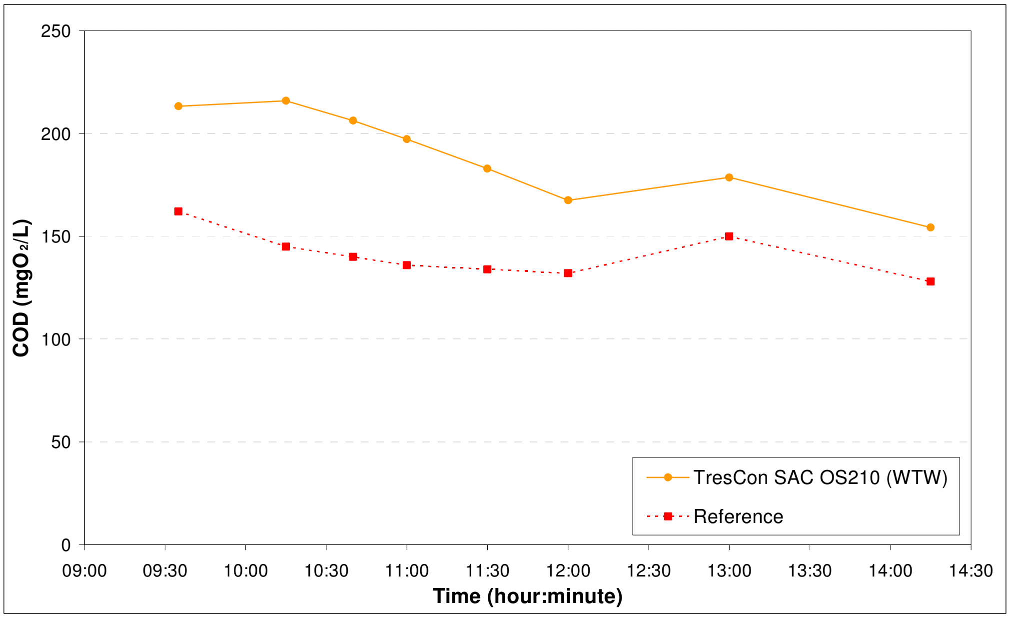


Figure 28: COD measurement on filtrated water. TresCon SAC OS210 (WTW). Waste water calibration

These graphs show that the waste water calibration gives the following results:

- Uvas from Hach Lange: the maximum error observed on the period is ~20% of the measure (455 mgO2/L for Reference COD = 380 mgO2/L)
- Spectro::Lyser from S::CAN: the maximum error observed on the period is ~8% of the measure (420 mgO2/L for Reference COD = 390 mgO2/L)
- TresCon SAC OS210 from WTW: the maximum error observed on the period is ~60% of the measure (230 mgO2/L for Reference COD = 145 mgO2/L)
- Carbovis from WTW: the maximum error observed on the period is ~15% of the measure (39 mgO2/L for Reference COD = 34 mgO2/L)

Still these results remain poorer than the lab-scale tests.

Measurement with supplier calibration

The only analyzer concerned by supplier calibration is PeCOD P100: Aqua Diagnostic provides a specific reagent that must be used for calibration. Thus in these tests, the device is automatically calibrated with the supplier reagent then is exposed to the waste water for a few hours without recalibration. Neither OCDE nor waste water is used for calibration.

The Figure 29 presents the comparison of the measures from PeCOD P100 from aqua Diagnostic with supplier calibration and the reference measures for two and a half days.

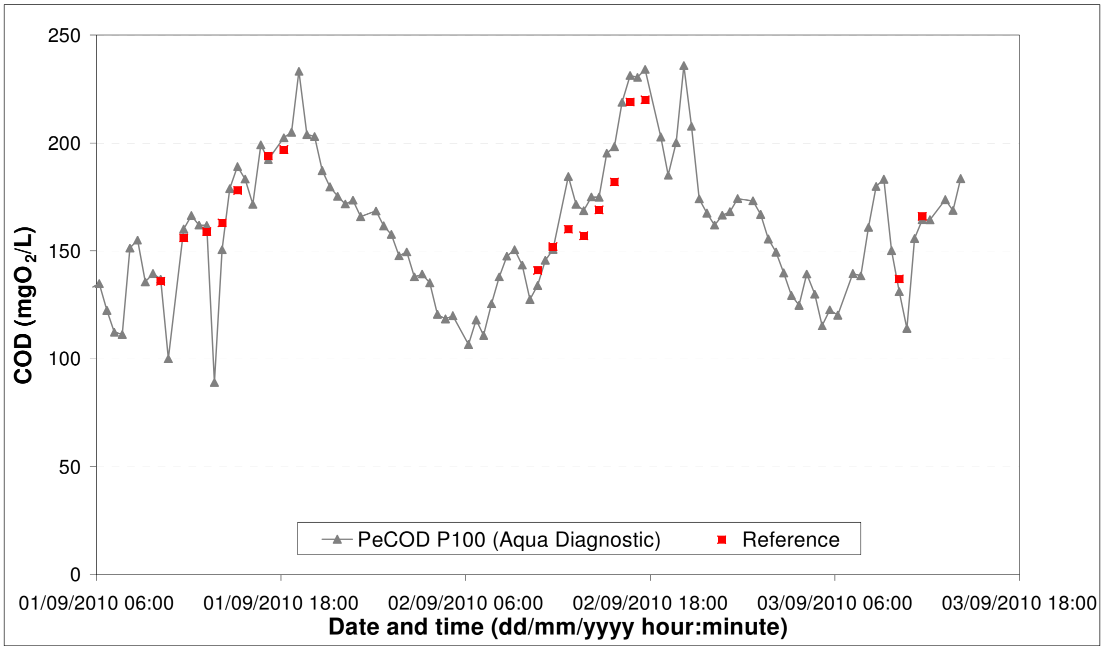


Figure 29: COD measurement on filtrated water. PeCOD P100 (Aqua Diagnostic). Supplier calibration

These graphs show that PeCOD P100 from Aqua Diagnostic and its particular calibration method give quite good results: the maximum error observed on the period is ~15% of the measure (160mgO_2/L for Reference COD = 185mgO_2/L).

These results are in the same order of magnitude than the lab-scale tests.

ii. Offline tests on industrial waste waters samples

For this second series of pilot-scale tests, the devices are exposed to industrial waste water samples from the outlet of the textile factory of Tekstina (Slovenia). Nine different solutions sampled on different days of a week at the outlet of the factory are used. They are numbered from 1 to 9. The measurements were done with industrial waste water sample calibration. The devices are calibrated with two samples of industrial waste water: samples 2 and 8. They are then exposed to the 9 samples.

> *Comment*: PeCOD P100 from Aqua Diagnostic is tested but it still remains calibrated with the reagent supplied by Aqua Diagnostic.

The Figure 30 presents the comparison of the measures from the devices with industrial waste water calibration and the reference measures on the 9 samples.

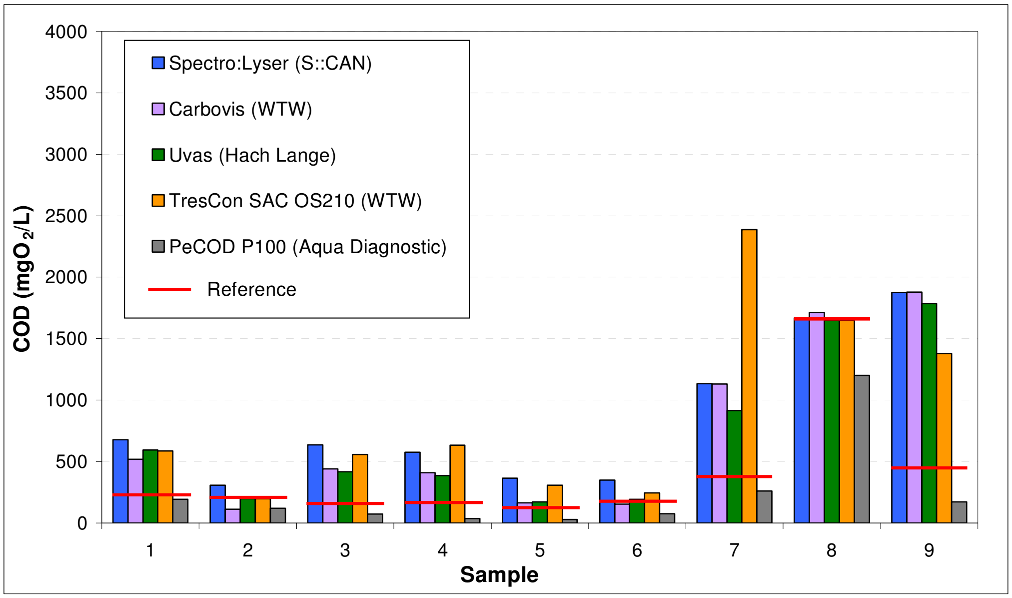


Figure 30: COD measurement on industrial waste water samples from Tekstina. Sample calibration

This diagram shows that the calibration with industrial waste water samples is a way to improve the accuracy of the measures: relative errors are globally divided by two. Still the results remain rather unsatisfactory in comparison to the exigencies:

- errors up to 550% of the measure are reached,
- for a same sample, measures can prove to be highly different from one device to another,
- for a same device, accuracy can be very different from one sample to another.

iii. Conclusion

The complementary tests prove that the matrix of real waste water is much more complex than OCDE which contains a very limited number of compounds sources of COD. Thus the accuracy results on an urban waste water effluent are much poorer than the lab-scale tests. An appropriated calibration (using samples of the water to analyze) is a way to improve the measures but is not sufficient regarding the needs.

At this stage of the study, it is difficult to recommend a particular sensor or analyzer only with the aid of lab-scale and rough pilot-scale results: maintenance needs, long-term drift and long-term calibration needs must be evaluated with complete on-field tests in real operating conditions.

3.1.2 Data processing

3.1.2.1 Software integration

The software integration is shown on figure 31.

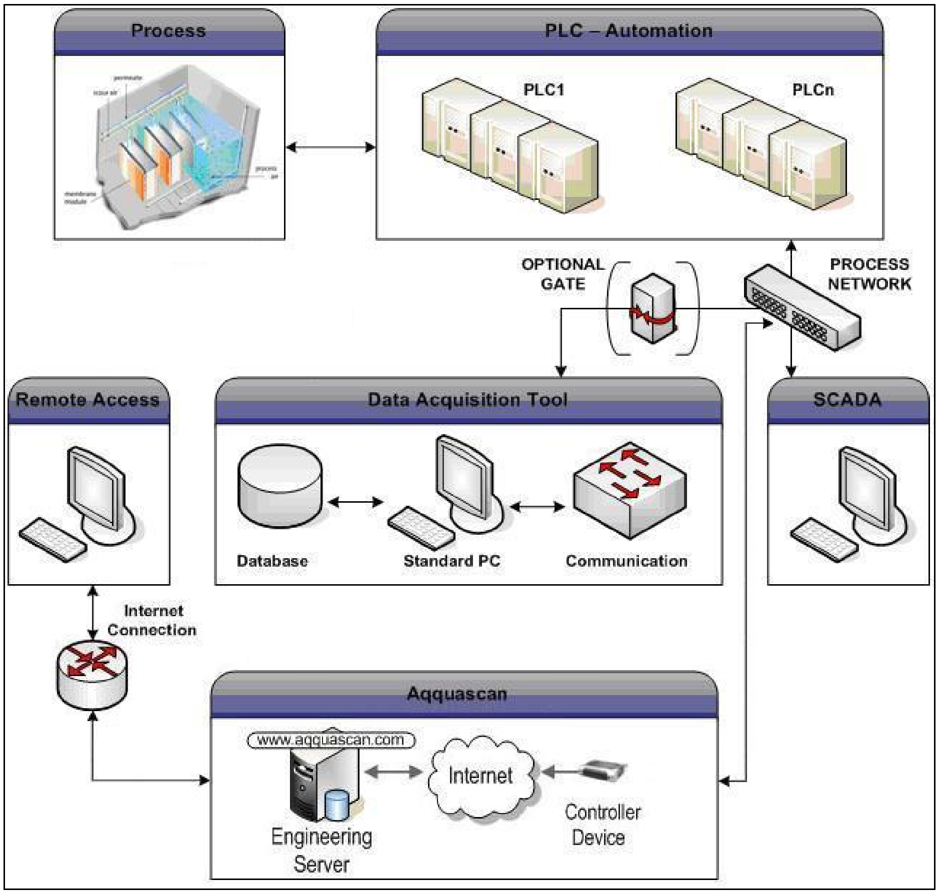


Figure 31 Software global architecture

Classical architecture of automation is composed of one or more PLC connected to process actuators and sensors in order to control and follow the process. A SCADA system allows the Human Machine Interface.

The software developed is able to connect independently to the Automation system, with direct OPC communication protocol, or via an optional gate. In such configuration, the data acquisition tool installed on a standard PC is totally independent, and there is no possible interference between existing architecture and data acquisition. The remote access is done using the AqquaScan Webpage. Through it, the information served by the data acquisition tool can be checked, both offline (historical data) and online (SCADA view).

3.1.2.2 Software main functionalities

Figure 32 shows the software architecture, organized with high level of modularity.

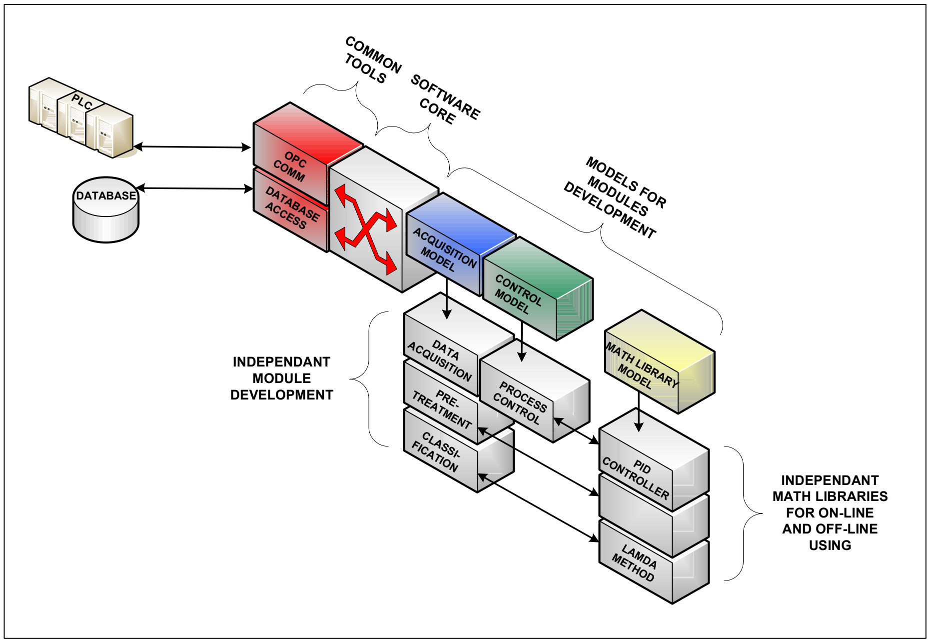


Figure 32 Software Modularity

The software is structured in many libraries, to allow high level of modularity. The main module (software core) is linking all theses libraries. Main libraries are:

- An OPC client library, for communication with PLC's.
- An OPC server library, to redistribute results with OPC protocol.
- A Database library, for database access (reading and writing).
- Models for modules development: One model for acquisition modules, one model for controller modules development, one model for mathematical modules development.
- Standard modules (developed with models libraries) for acquisition, control and mathematical treatment.

3.1.2.3 On-line and Off-line developing concept

Each module is developed with on-line and off-line concept: each mathematical processes used by on-line modules are developed in an external library, in order to reuse it with any other software (for off-line using). This concept is presented in figure 33.

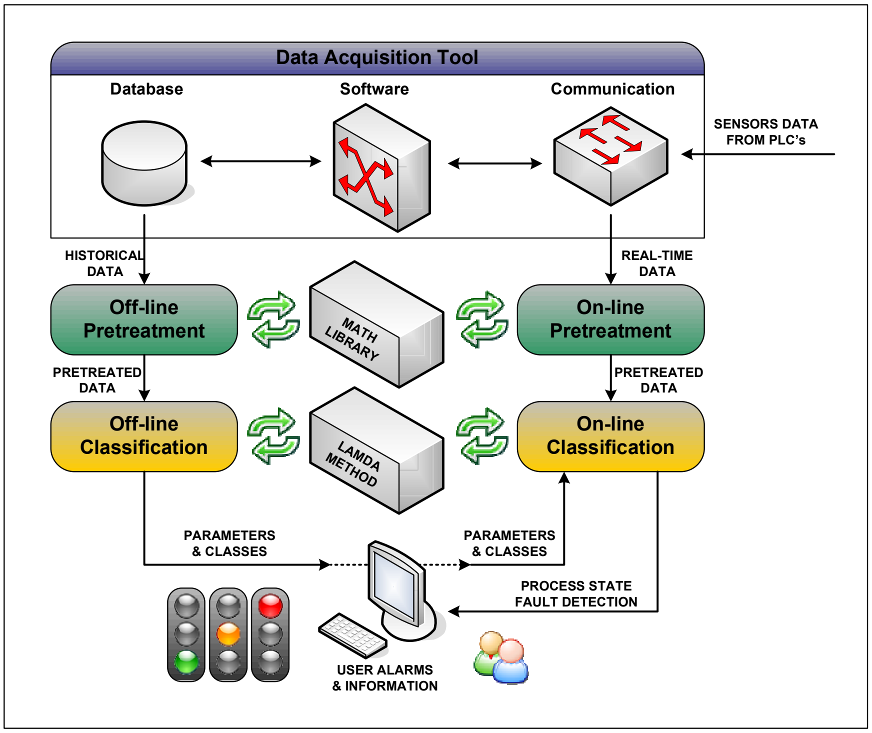


Figure 33 On-line and Off-line concept example

In this example for on-line and off-line classification, two libraries are presented: one for data pretreatment and one for classification. The off-line and the on-line software uses theses same libraries.

3.2 Technical progress of the work

3.2.1 Evaluation of sensors: Test protocol

In this part, the results of the test protocol applied to each COD sensor tested are presented in detail.

<u>Spectro::Lyser from S::CAN</u>
The calibration results of Spectro::Lyser are presented in the table 15 below.

Table 14 Calibration results of Spectro::Lyser from S::CAN

Reference COD (mgO_2/L)	COD measure before device calibration (mgO_2/L)
0	1.3
50	36.0
200	78.8
774	221.6

The final results of Spectro::Lyser are presented in the table 16.

Table 15 Criteria calculated for Spectro::Lyser from S::CAN

Criterion	Unit	Raw value	Performance coefficient
RT	minutes	1.5	-6.7
LOD	mgO_2/L	31.4	1.6
B20		6.4	-3.1
B80		93.8	4.7
SDC20		48.4	2.4
SDC80		49.8	2.5
Offset		-36.2	1.81
(Slope-1)xSR		312.2	15.6

These results show that Spectro::Lyser from S::CAN is an in-situ probe with a very fast response time compared to the exigency of 10 minutes: because the measurement principle (UV-vis absorption) is physical, it is only limited by the number of measures executed per minute (here: 1). It has a very good accuracy at 20% of the studying range but a poor one in higher values in spite of the calibration. The probe delivers quite a noisy, unsteady signal that makes its limit of detection and its smallest detectable changes higher than the exigencies. This light instability of its signal delivered may be due to the algorithm of compensation that takes into account the turbidity and the suspended matters and requires complicated calculation.

The study of the suspended matter (SM) measurement – that is used to compensate for the COD signal – shows that the sensor is very sensitive to the SM measure: when magnetic agitation is off, COD level decreases in the same way as SM level.

<u>CarboVis from WTW</u>
The calibration results of CarboVis are presented in the table 17 below:

Table 16 Calibration results of CarboVis from WTW

Reference COD (mgO_2/L)	COD measure before device calibration (mgO_2/L)
0	142
1000	361

The final results of CarboVis are presented in the Table 18:

Table 17 Criteria calculated for CarboVis from WTW

Criterion	Unit	Raw value	Performance coefficient
RT	minutes	6.1	-1.6
LOD	mgO_2/L	0.0*	*10*
B20		82.6	4.3
B80		87.3	4.4
SDC20		97.4	4.9
SDC80		108.2	5.4
Offset		91.6	4.6
(Slope-1)xSR		-116.9	5.9

* Comment: in this particular case, this zero value is misleading since the sensor actually measures zero when exposed to COD 50 mgO_2/L, hence a performance of 10 for LOD.

These results show that CarboVis from WTW is an in-situ probe with a slow response time compared to the exigency of 10 minutes: the signal is not even always stabilized after a 30min exposure to a sample (see Figure 47 in annexe) which means the result of 6.1 min is actually underestimated. This may be due to the calculation algorithm that must not be optimized and generates a quite long rise and fall times in case of abrupt changes. Its performances in matter of accuracy and repeatability are very poor in the entire studying range and it evens measure 0 mgO_2/L when it is exposed to a sample of COD ~ 50 mgO_2/L.

The probe itself is the same as Spectro::Lyser of which performances are much better. This is why the limits of this sensor must due to the algorithm that treats the signal from the probe.

<u>Uvas from Hach Lange</u>
The calibration results of Uvas are presented in the Table 19 and Figure 34 below.

Table 18 Calibration results of Uvas from Hach Lange

Reference COD (mgO_2/L)	Absorption measure before manual calibration (m^{-1})
0.0	0.0
53.7	3.6
192	15.6
397	32.0
625	50.3
789	65.5
970	79.5

Comment: the absorption measure is reported from a steady value given by the device after stabilisation.

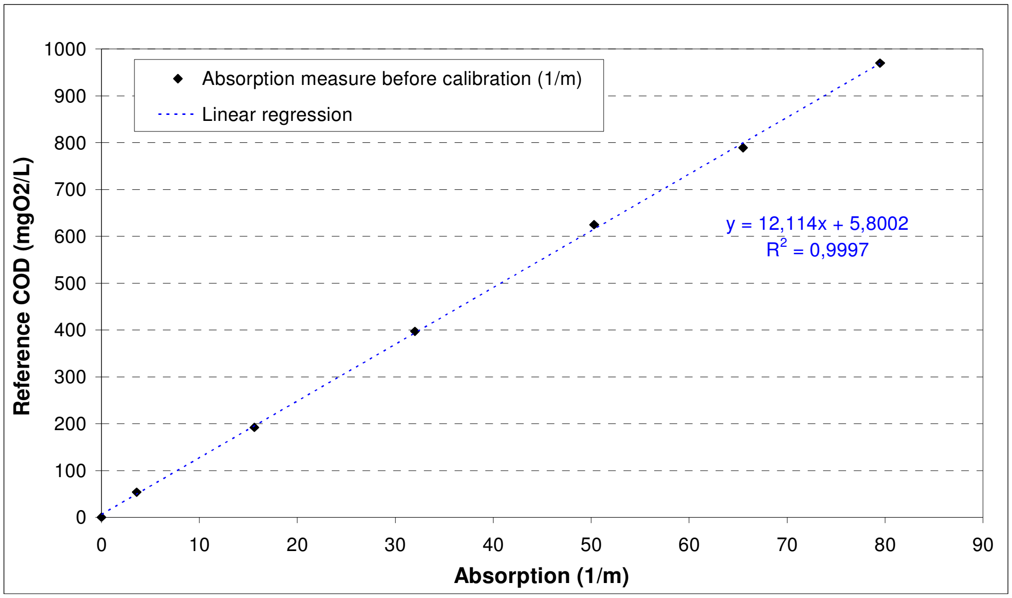


Figure 34: Linear regression for the calibration of Uvas from Hach Lange

The equation y = 12.11 * x + 5.80 is used offline on a computer to convert the absorption measure (in m^{-1}) into a COD measure (in mgO_2/L).This equation is considered as characteristic of Uvas and used only for this sensor.
The final results of Uvas are presented in the Table 20 below:

Table 19 Criteria calculated for Uvas from Hach Lange

Criterion	Unit	Raw value	Performance coefficient
RT	minutes	1.7	-5.6
LOD	mgO_2/L	4.0	-5.1
B20		9.5	-2.1
B80		4.2	-4.7
SDC20		9.5	-2.1
SDC80		15.6	-1.3
Offset		7.4	-2.7
(Slope-1)xSR		2.0	-10.0

These results show that Uvas from Hach Lange is an in-situ probe with all the lab-scale qualities required: very fast response time, excellent limit of detection and very good accuracy and repeatability in the whole studying range. The measure given by the sensor is very steady due to its simple physical measurement principle and an effective calculation algorithm.

<u>TresCon from WTW</u>
The calibration results of TresCon are presented in the Table 21 and Figure 35.

Table 20 Calibration results of TresCon from WTW

Reference COD (mgO_2/L)	Absorption measure before manual calibration (m^{-1})
0.0	0.1
53.7	5.2
192	23.4
397	40.4
625	62.4
789	80.7
970	102.5

Comment: the absorption measure is reported from a steady value given by the device after stabilisation.

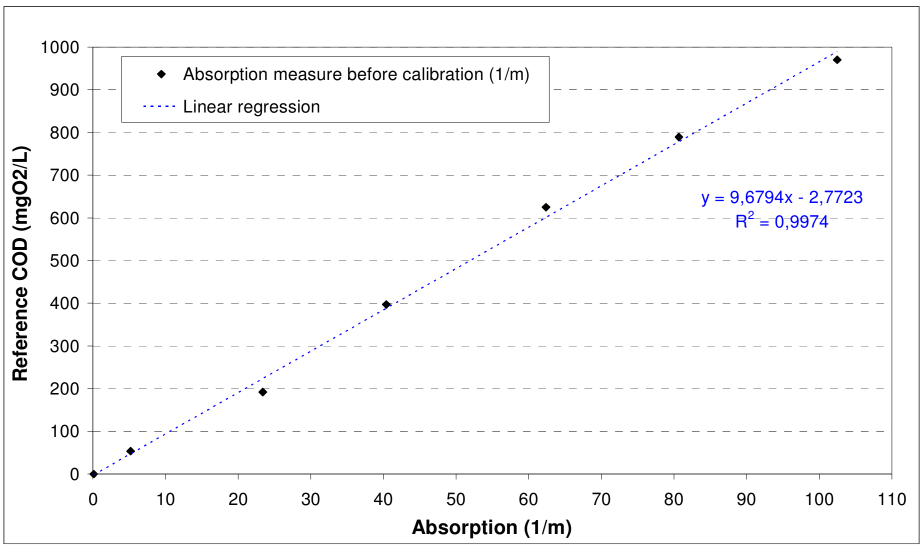


Figure 35: Linear regression for the calibration of TresCon from WTW

The equation y = 9.68 * x - 2.77 is used offline on a computer to convert the absorption measure (in m^{-1}) into a COD measure (in mgO_2/L). This equation is considered as characteristic of TresCon (SAC OS 210 module) and used only for this analyzer.

The final results of TresCon are presented in the Table 22 below.

Table 21 Criteria calculated for TresCon from WTW

Criterion	Unit	Raw value	Performance coefficient
RT	minutes	7.3	-1.4
LOD	mgO_2/L	3.5	-5.75
B20		12.1	-1.7
B80		18.2	-1.1
SDC20		1.3	-14.9

SDC80		3.0	-6.8
Offset		10.3	-1.9
(Slope-1)xSR		10.1	-2.0

These results show that TresCon from WTW is an online analyzer with a good response time compared to the exigency of 10 minutes because of its continuous-batch working (2 batches per minute). It also has an excellent accuracy and repeatability at both 20% and 80% thanks to very steady measures.

PeCOD P100 from Aqua Diagnostic
PeCOD P100 is the only device tested that is automatically calibrated without the aid of OCDE solution: it is designed to work with its own reagent supplied by Aqua Diagnostic.

Table 22 Criteria calculated for PeCOD 100 from Aqua Diagnostic

Criterion	Unit	Raw value	Performance coefficient
RT	minutes	20	2.0
LOD		7.3	-2.74
B20		-45.2	2.3
B80		-18.1	-1.1
SDC20	mgO_2/L	37.5	1.9
SDC80		66.12	3.3
Offset		-25.5	1.3
(Slope-1)xSR		15.8	-1.3

These results show that PeCOD P100 from Aqua Diagnostic is an online analyzer with a slow response time because of its batch working with only 1 batch every 20 minutes. Its accuracy is good in the high values of the studying range but lightly insufficient in the low ones. The repeatability of the measure is also insufficient.

3.2.2 Data processing tool

3.2.2.1 Data Structures

Figure 36 shows how data structures are used in the software.

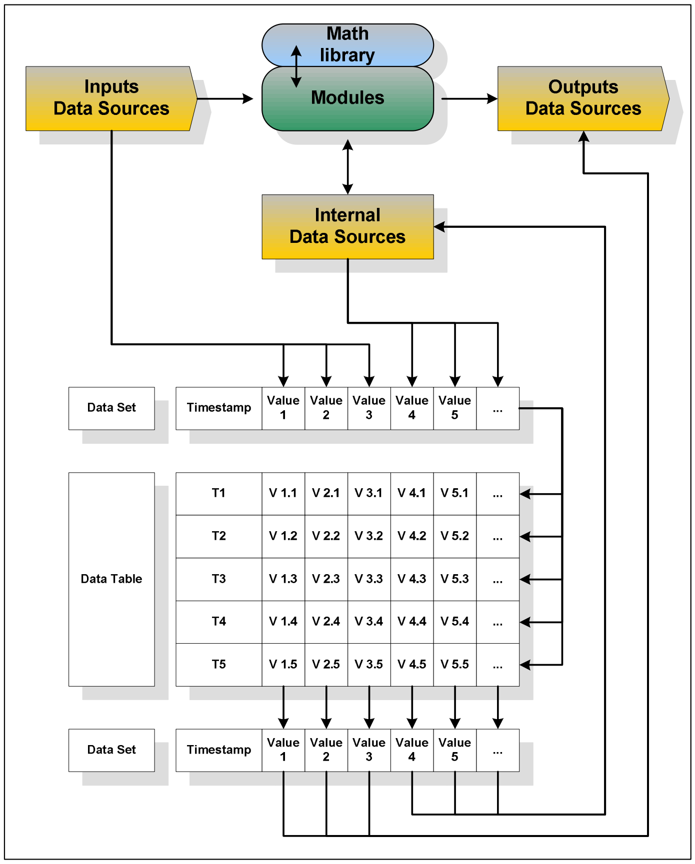


Figure 36 Data structures

i. Data sources

Data sources are either the input and output values of the software, and its internal values. Each value used by the modules are described by the properties of this object witch can be of three types: numerical, Boolean or text.

All data sources are described by:

- Name;
- Type of value;
- Value;
- Quality (Good or bad).

Input communication data sources are described by the way they are evaluated; a formula witch can be composed of elements presented in table 24. Their values are evaluated at a frequency defined in the software general configuration.

Table 23 Elements used for data evaluation

Elements	Examples
OPC Tags (communication with PLC's)	{Node|ServerName|TagName} Selected with Tag Browser
System tags	Now, Today, Now.Minute

Constants	10 ; 2,05 ;10^2 ; True ; False
Mathematical constants	E, PI
Mathematical operators	+, -, *, /, Mod, ^
Logical operators	<, >, <=, >=, <>, =, Not, Or, And, Xor
Mathematical functions	Abs(x), Acos(x), Asin(x), Atan(x), Atan2(x,y), Cos(x), Cosh(x), Exp(x), Log(x), Log(x,y), Log10(x), Max(x,y), Min(x,y), Pow(x,y), Round(x), Sin(x), Sinh(x), Sqrt(x), Tan(x), Tanh(x)
Priority characters	(and)

Outputs formulas can only be composed by a unique OPC tag witch can be written in the OPC server.
Numerical data sources are described by:

- Unit and decimals
- Limits (minimum and maximum values : under minimum value, and over maximum value, the data source quality is bad)
- Levels (Very low, low, high and very high levels witch can be used for operator information)

ii. Data sets

Data sets are groups of values with timestamp. A data set can be symbolized by a table line, in witch the first element is the timestamp, and others elements are values linked to their data source. Data sets are used either to store values of input data sources (acquisition), intermediate or final values (from calculation).

iii. Data Table

Data tables are groups of datasets. A data table can be symbolized by a table, in witch the first column is the timestamps, and others columns are values linked to their data sources. Data tables are used to store values for intermediate or final calculations (for example to calculate averages, slopes, etc.).

3.2.2.2 Mathematical libraries

Mathematical libraries are available in the software. They can be used by modules to transform raw values. For examples:

- Using one or more value from OPC to calculate another input;
- Calculate average from a data table column;
- Calculate an internal or an output value with one or more calculated value;
- Etc.

For calculation on data tables, available calculations are at the moment: sum, average, minimum, maximum, slope and duration.
Others mathematical libraries can be developed and used easily in the software, for example for data pre-treatment, statistical analysis, classification, etc.

3.2.2.3 Acquisition modules

Figure 37 shows an example of an acquisition module configuration.

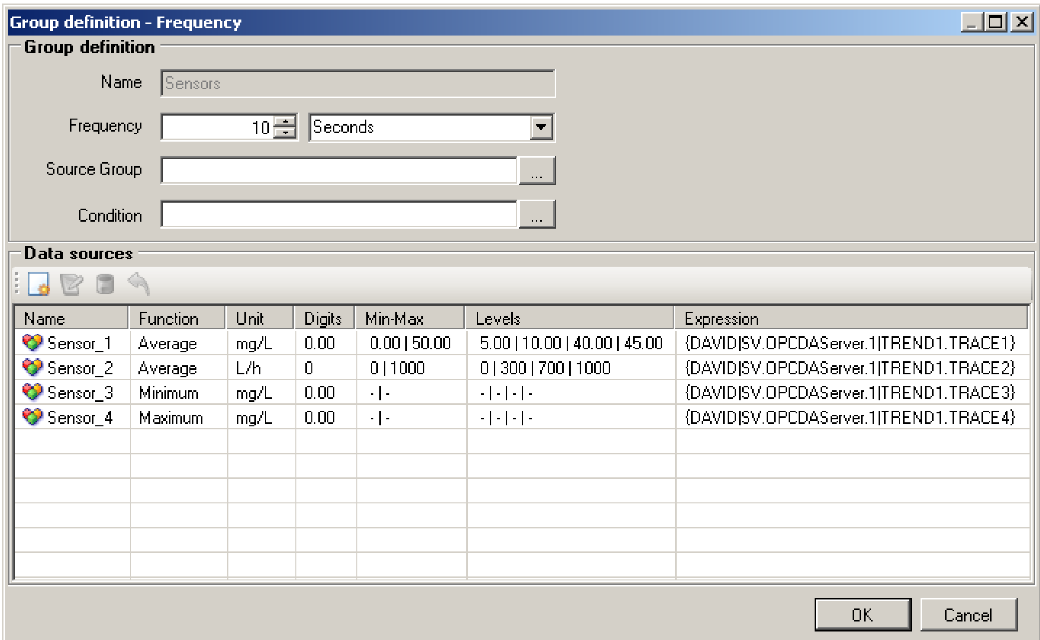


Figure 37 Example of standard module configuration form

i. Standard acquisition (frequency)

Operator defines a group of data sources. Each value will be stored simultaneously in a data set at the main software frequency, and each data set will be stored in a data table. At a frequency defined by operator between 1 second and 24 hours, a calculation is realized on each column of the data table (sum, average, minimum, maximum, slope, duration or last value) and each result will be stored in a result data set and written in the database. Possible data type for inputs is only numerical values. Recording can be conditioned by a Boolean formula.

ii. Statement acquisition

Operator defines a group of data sources. Each value will be stored simultaneously in a data set and written in the database at a frequency defined by operator. The frequency can be:

- Daily at a defined time.
- Weekly at a day and a defined time.
- Monthly at a day and a defined time.

Possible data type for inputs is only numerical values.

iii. Event acquisition

Operator defines a group of data sources. When the value of one data source changes, the new value is written in the database. Possible data type for inputs is only numerical values.

iv. Text acquisition

Like event acquisition, but for text data type.

3.2.2.4 Process Modules

Process modules are specialized acquisition module for one process. For example, a module can:

- Define witch data are needed, acquisition frequency, data types, etc.
- Make process calculations with data, record results in database

- Follow process events and make calculations between these events
- Give operator specific boards and tools to follow process
- Etc.

Process modules can be developed easily and quickly for each process.

3.2.2.5 Data format and extraction

The structure of the Microsoft Access database depends of the modules used. For standard modules, tables' structures are presented in table 25.

Table 24 Database structure for standard acquisition modules

Module	Table's structure
Standard acquisition	1 table by acquisition group First column : Timestamp (date format) Others columns : Values for each data source (real format)
Statement acquisition	1 table by acquisition group First column : Timestamp (date format) Others columns : Values for each data source (real format)
Event acquisition	1 table by acquisition group First column : Timestamp (date format) Second column : Name of the data source (string format) Third column : Value of the data source (real format)
Text acquisition	1 table by acquisition group First column : Timestamp (date format) Second column : Name of the data source (string format) Third column : Value of the data source (text format)

For specific modules, the database structure is free for the developer: number of tables, structure of tables, data formats, etc…

Data extraction can be done in different ways:

- Directly with Microsoft Access (2000 or newest versions) and SQL Queries
- With Microsoft Excel and its data extraction tool Microsoft Query
- With every data extraction tool using SQL language

3.2.2.6 Operator information and help

The main operator form is presented on figure 38.

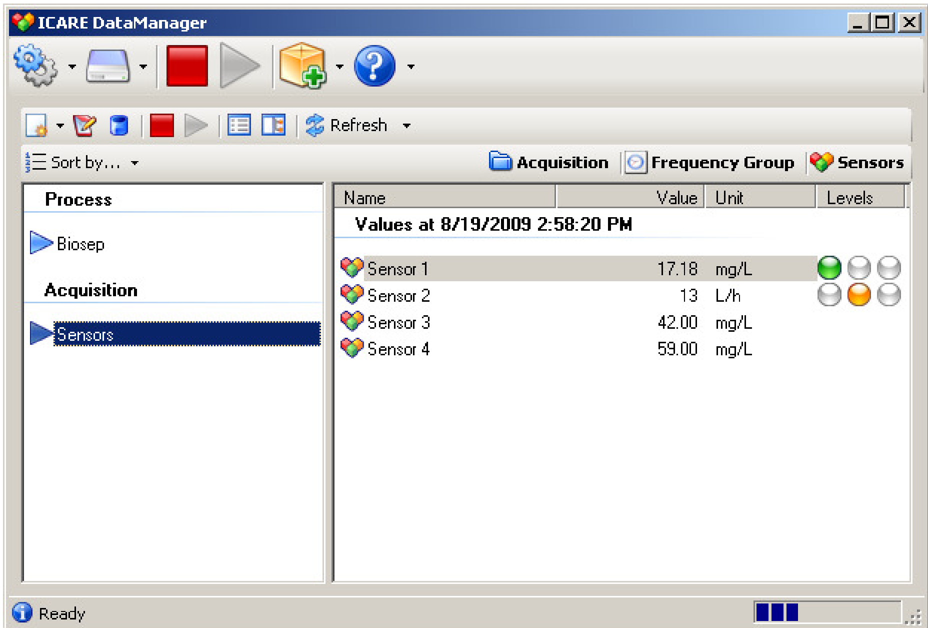


Figure 38 Main form

The left part of the window shows existing groups, sorted by type, category or user-defined sorting. The right part shows values of data sources for the selected group, with its units and levels (like traffic lights: green, orange or red). Each group can be started and stopped independently and there is a main command for acquisition start and stop. Groups can be added, edited or deleted with command buttons in toolbar.

For data sources levels, a main window can be enabled and is shown on figure 39.

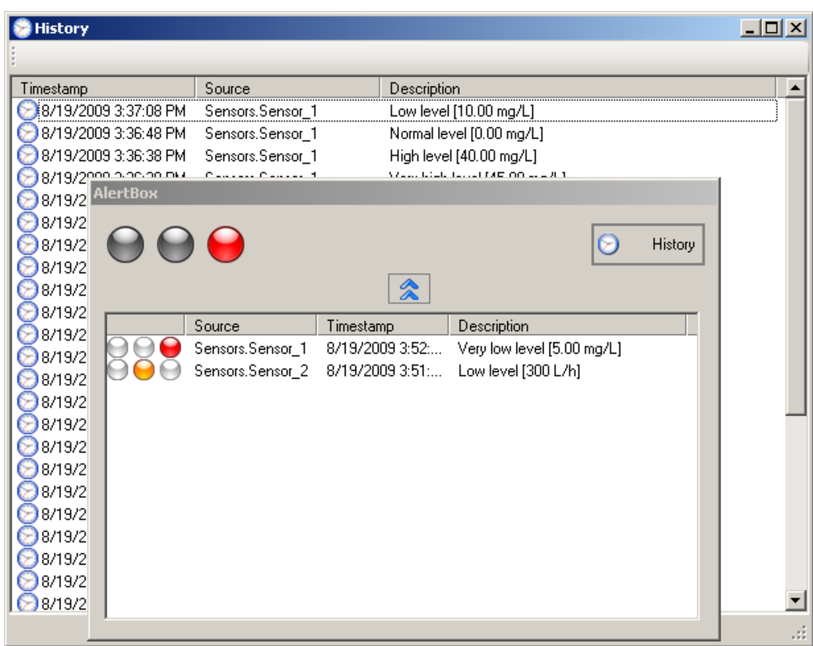
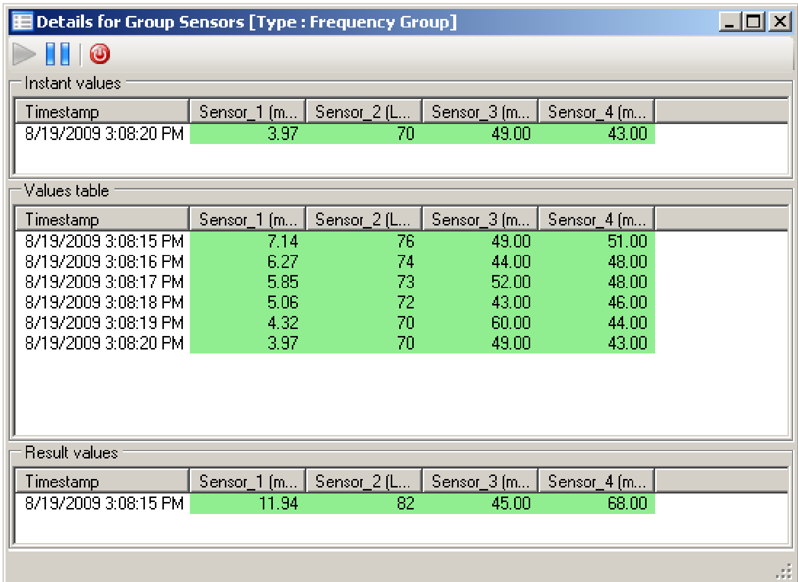


Figure 39 Operator warning forms

The main level represents the worst situation for every data source, and the detailed view describes the situation for each one. A history window can be opened and shows every level changing for each data source.
A detailed view is available for each group (figure 40)



Instant values

Timestamp	Sensor_1 (m...	Sensor_2 (L...	Sensor_3 (m...	Sensor_4 (m...
8/19/2009 3:08:20 PM	3.97	70	49.00	43.00

Values table

Timestamp	Sensor_1 (m...	Sensor_2 (L...	Sensor_3 (m...	Sensor_4 (m...
8/19/2009 3:08:15 PM	7.14	76	49.00	51.00
8/19/2009 3:08:16 PM	6.27	74	44.00	48.00
8/19/2009 3:08:17 PM	5.85	73	52.00	48.00
8/19/2009 3:08:18 PM	5.06	72	43.00	46.00
8/19/2009 3:08:19 PM	4.32	70	60.00	44.00
8/19/2009 3:08:20 PM	3.97	70	49.00	43.00

Result values

Timestamp	Sensor_1 (m...	Sensor_2 (L...	Sensor_3 (m...	Sensor_4 (m...
8/19/2009 3:08:15 PM	11.94	82	45.00	68.00

Figure 40 Acquisition details

In this example, the detailed view shows last values from data sources (first line), data table (center window) and last value calculated for each data source (last line).
A board can be developed for each module, for operator information and help. Figure 41 shows an example of board for standard acquisition.

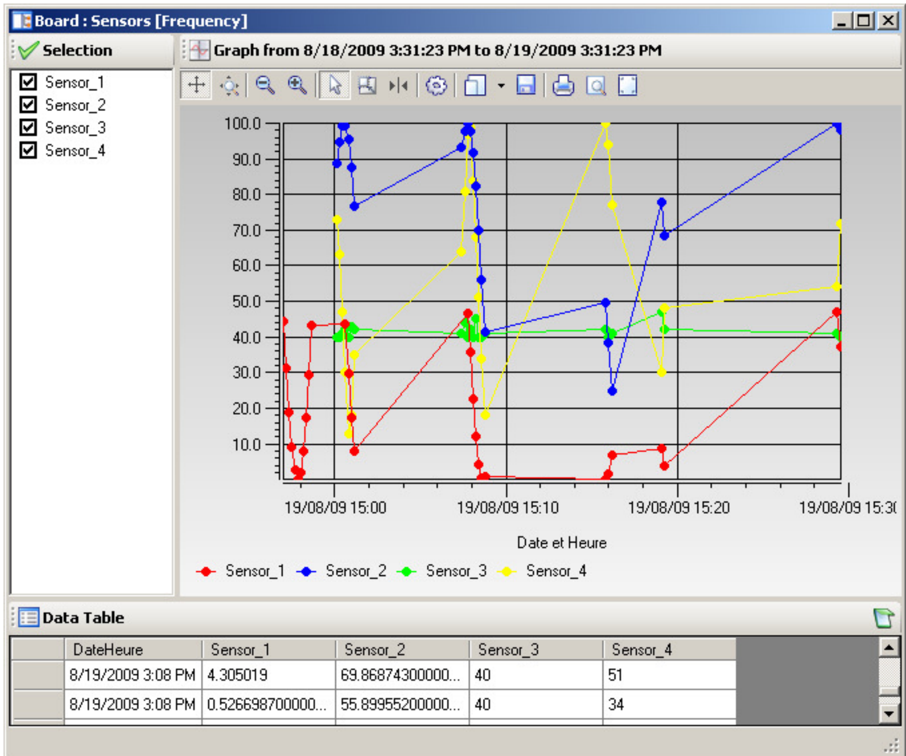


Figure 41 Operator Board

In this example, the board shows last values recorded in database for each data source, in a data table and in a graph. Theses information can be exported in different ways (graph picture, text file with data).

4 Conclusions

4.1 Major achievements

Concerning the evaluation of sensors, some points are worth pointing out:

- The methodology to compare different analyzers and probes is defined. This test protocol allows calculating comparison criteria to benchmark the instruments. In such a project as AQUAFIT4Use one, where the water quality is one of the main concerns, sensors can stand as a key for many studies (simulation, efficiency of new processes…). Thus, it was essential to develop a method to have an objective assessment on real capacities of a sensor. The first part of the method defines a common vocabulary and explains concretely the several concepts. Afterward, the methodology to evaluate sensors was define, inspired from the norm ISO 15839. The protocol has been checked and validated by several case studies on different kind of instruments. As said in this report, an accurate definition of the application is an essential step in this protocol. Indeed, the goal and the role of the instrument have to be clearly defined to set criteria requirement. Finally, the protocol development won't exempt the reader to have a critical mind and feeling to interpret the results. The interpretation of R^2 coefficient is a good example.

- Lab-scale and first pilot-scale tests were executed on COD, Total Nitrogen, Ortho-phosphate and Calcium sensors and analyzers. Results were analyzed to benchmark them. This work will help to measure the quality of the water reuse inside a process line.

    - At lab-scale, Uvas from Hach Lange and TresCon from WTW are the only devices that meet all the requirements set in the preliminary study for COD measurement. Both apparatus are based on the same measurement principle which seems to be – at least at lab-scale – the most effective: indeed it is physical (UV absorption) and uses a very limited number of wavelengths to work, with a simple compensation for turbidity that does no induce any delay time. The three other devices use more complicated measurement principle based on full spectrum absorption (Spectro::Lyser from S::CAN and CarboVis from WTW) or chemical reactions (PeCOD P100 from aqua Diagnostic) that induce more important data treatments.

    - At pilot-scale, the results remain poorer than the lab-scale tests, even for the Uvas frol Hach lange and TresCon from WTW. The maximum error observed is between 8 % and 60 %, far from the requirements. Nevertheless, the PeCOD P100 analyzer from Aqua Diagnostic and its particular calibration method give quite good results compared to those obtained at lab-scale : the maximum error observed on the period is ~15% of the measure (160 mgO_2/L for Reference COD = 185 mgO_2/L).

- The results from the tests on urban and industrial waste water samples prove to be very different from lab-scale tests – except for PeCOD P100 from Aqua Diagnostic –accuracies are poorer even with an adequate calibration. These bad results are due to the local characteristics of the water matrix. Indeed, the calibration of the instruments on a real and complex matrix is a tricky step. The measurement of PeCOD P100 from Aqua Diagnostic is independent of the water matrix, this is a real advantage given the problems encountered in calibrating all the sensors.

About the data processing tool, the following results were obtained:

- The software is developed and is able to plug to any system with OPC protocol and optional gateway. This decision support tool aims to improve the user information levels thanks to alarm windows with traffic light, historical values, graphical board, database contents…
- The software is made of some useful modules to improve the process management: an acquisition module allows to record data, a controller one to develop a control strategy on a process, *etc*.
- The high level of modularity allows development of modules for on-line and off-line data treatment:
    - Four pre-treatments are chosen, tested and validated for data reconciliation and fault detection
    - The classification method chosen is the LAMDA method.
    - Other pre-treatments methods can be added easily to the decision support tool thanks to the modularity of the architecture
- A remote web access tool was added to the software.

4.2 Future work

4.2.1 Within AquaFit4Use

For the evaluation of sensors, full on-field tests have to be executed to complete the evaluation and lead to solid recommendations for the choice of equipment. These tests will carry on next year.

The data processing software and the remote access are not limited to the WWTP monitored data, but can provide online data from any supervised variable in the whole industry. If it is of interest of any partner, VEOLIA and ATM could test the possibility of giving remote access to other type of industrial monitored data.

4.2.2 General recommendations

At the present time, the remote web access tool AqquaScan is only available in a Spanish version, so an English translation would be appreciated.

5 Literature

[1] ISO 15839_2003:
Water quality — On-line sensors/analysing equipment for water — Specifications and performance tests

[2] ISO 8466-1_1990:
Water quality – Calibration and evaluation of analytical methods and estimation of performance characteristics.

[3] OPC Foundation – OPC Data Access Automation Specification – Version 2.02 – February 1999

[4] Software Toolbox – OPC Data ActiveX Control Specifications

[5] DUDA R., HART P., STORK D. Pattern Classification, Second Edition, Whiley-Interscience - 2001.

[6] COMBACAU M., BERRUER P., ZAMAI E.; CHARBONNAUD P., KHATAB A., Supervision and monitoring of production systems. 2nd management and control of production and logistics, MCPL, Grenoble, France, 2000.

[7] FAYYAD U., PIATETSKY-SHAPIRO G., SMYTH P., From Data Mining to Knowledge Discovery: An overview. Advances in knowledge discovery and data mining, AAAI Press/MIT Press, California, USA, 1996.

[8] MICHIE D., SPIEGELHALTER D.J., TAYLOR C.C., Machine Learning, Neural and Statistical Classification, Ellis Horwood series in Artificial Intelligence, February, 1994.

[9] KEMPOWSKY Tatiana. Surveillance de procédés à base de méthodes de classification : Conception d'un outil d'aide pour la détection et le diagnostic des défaillances. Thèse de doctorat, Institut National des Sciences Appliquées, Toulouse, France, décembre 2004.

[10] LAMRINI Bouchra et al. Conduite supervisée d'une station de traitement d'eau potable par la méthode de classification LAMDA. Récents Progrès en Génie des procédés, Numéro 92 – 2005.

[11] SALSA – User's Manual 2009 v1.0. LAAS-CNRS

[12] Castro, A., Sanz, J. M., Ayala, I., Ayesa, E., Alferes, J. & Irizar, I. (2008) AqquaScan: design and implementation of an internet-based service for the remote monitoring and management of decentralised WWTPs. Water Sci. Technol. 57(10), 1495–1502.

[13] I. Irizar, J. A. Zambrano, D. Montoya, M. De Gracia and R. Garcia (2009).Online monitoring of OUR, KLa and OTE indicators: practical implementation in full-scale industrial WWTPs. Water Science and Technology, 60(2), 459,466.

[14] Irizar I., Sanz J.M., Ayala I., Pérez C., Castro A., Arrizabalaga S. and Salterain A. (2006). Integrated web-based supervision service for teleoperation of distributed WWTPs. Proceedings of the 7th IWA Specialised Conference on Small Water and Wastewater Systems, Mérida, Mexico, 7-10 March 2006.

[15] Suescun J., Irizar I., Ostolaza X. and Ayesa E. (1998). Dissolved oxygen control and simultaneous estimation of oxygen uptake rate in activated-sludge plants. Water Environment Research, 70(3), 316-322.

6 Annex

Here are presented the raw results of all lab-scale tests.

COD results

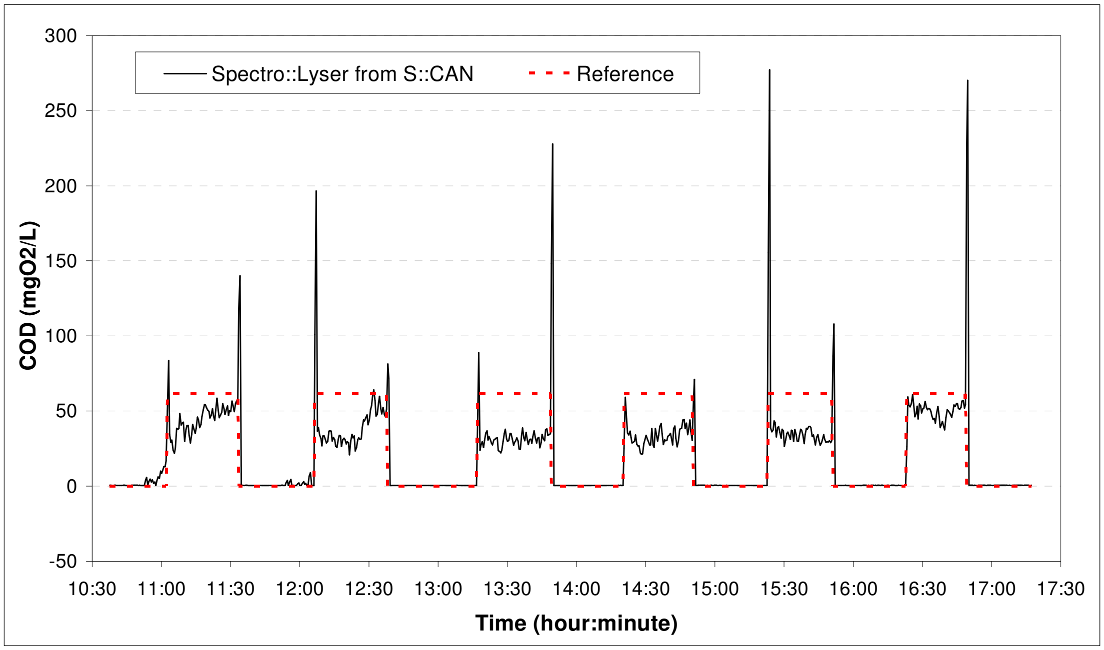


Figure 42: Accuracy test of Spectro::Lyser from S::CAN

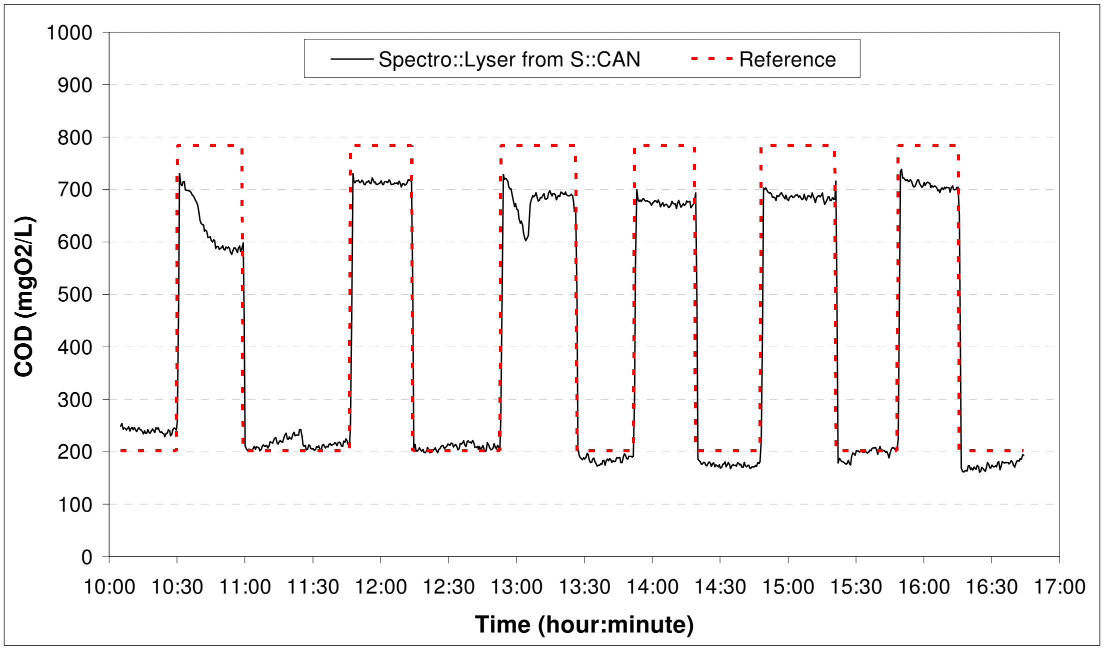


Figure 43: Dynamic test of Spectro::Lyser from S::CAN

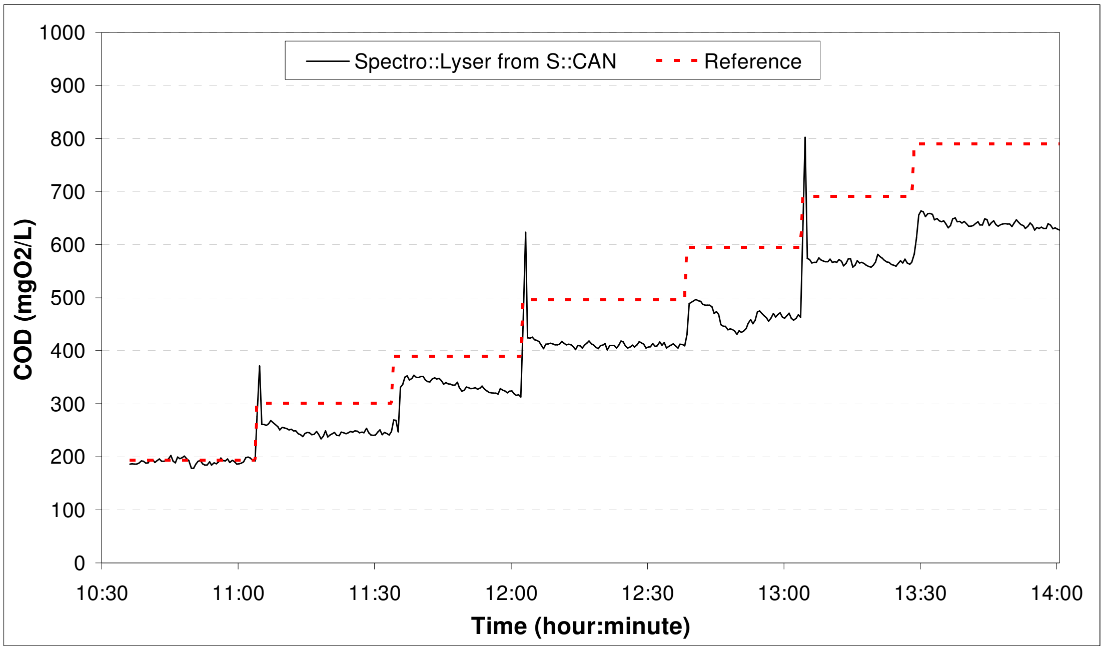


Figure 44: Linearity test of Spectro::Lyser from S::CAN (1/2)

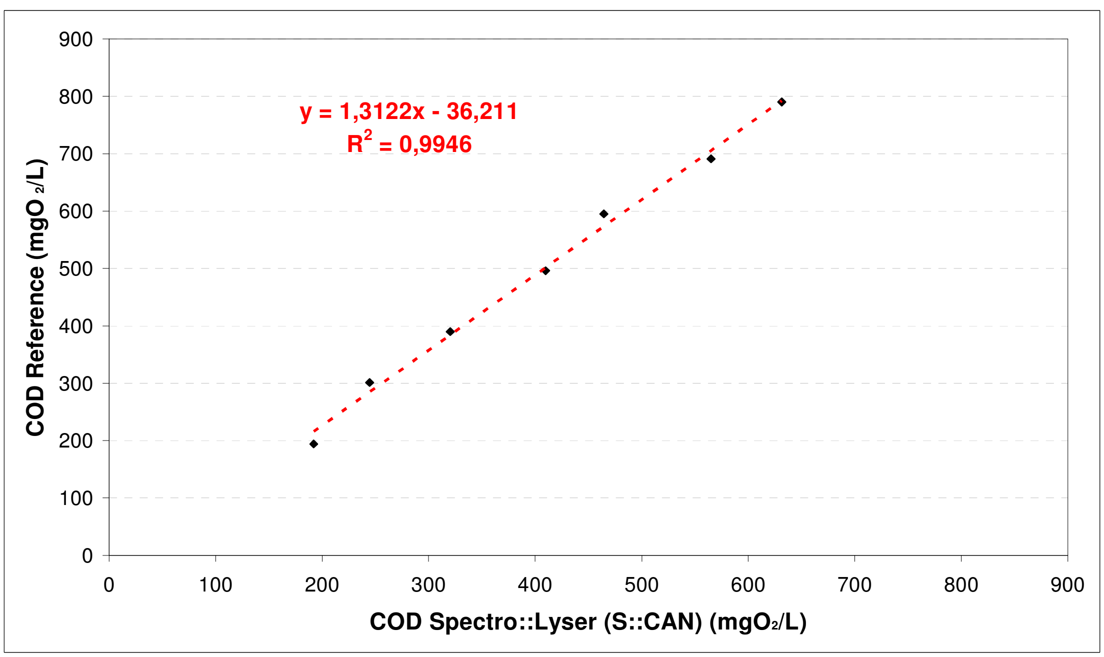


Figure 45: Linearity test of Spectro::Lyser from S::CAN (2/2)

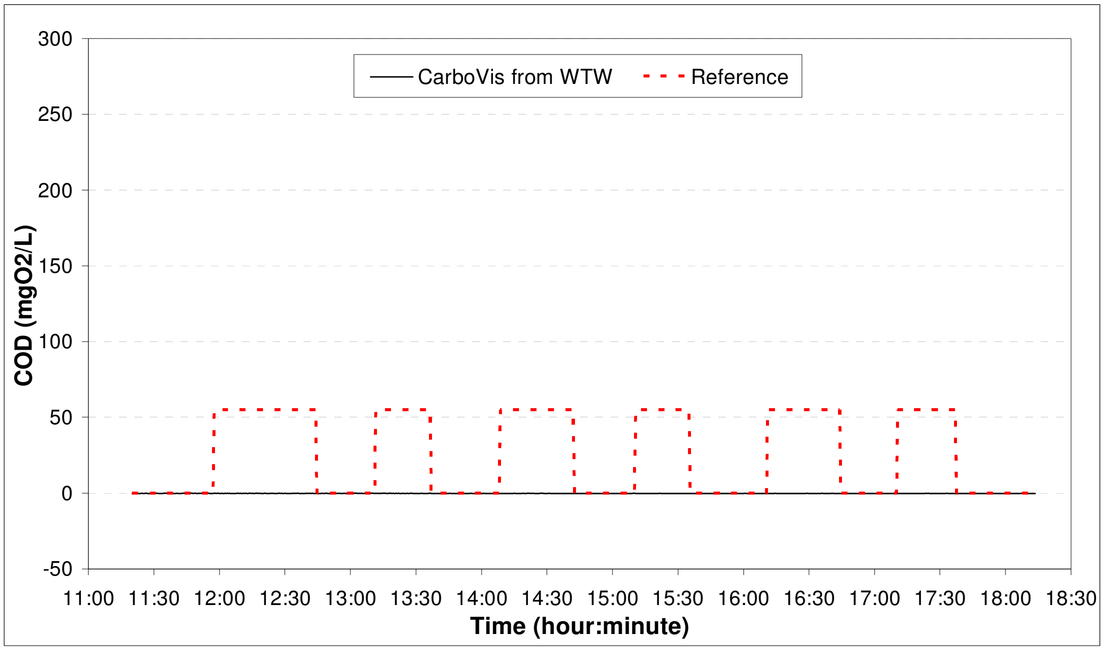


Figure 46: Accuracy test of Carbovis from WTW

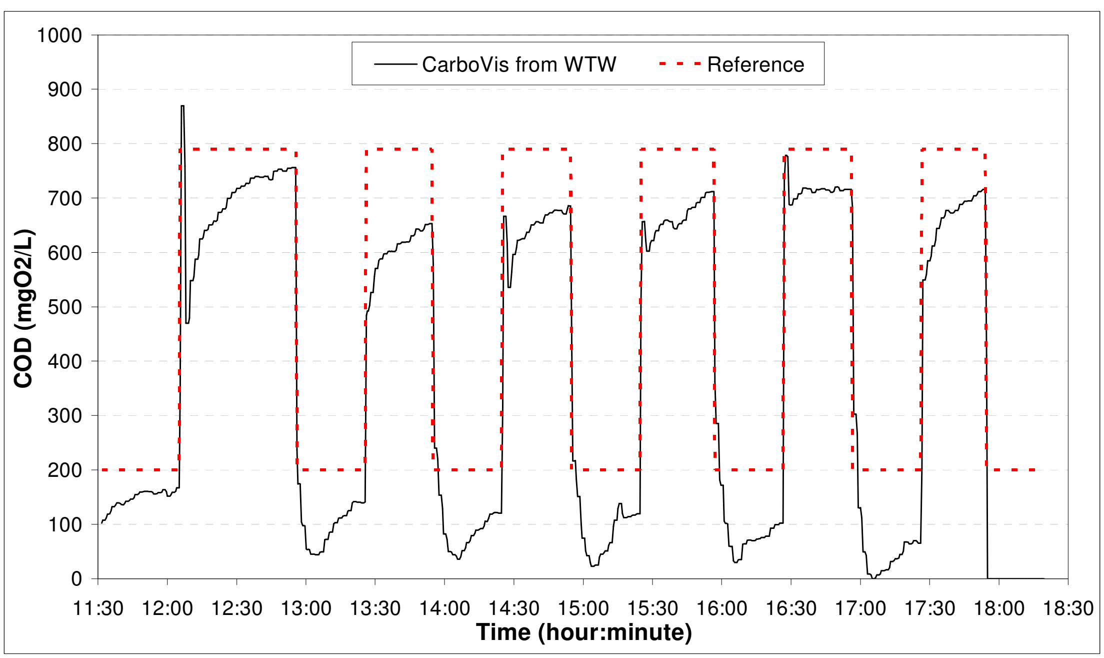


Figure 47: Dynamic test of Carbovis from WTW

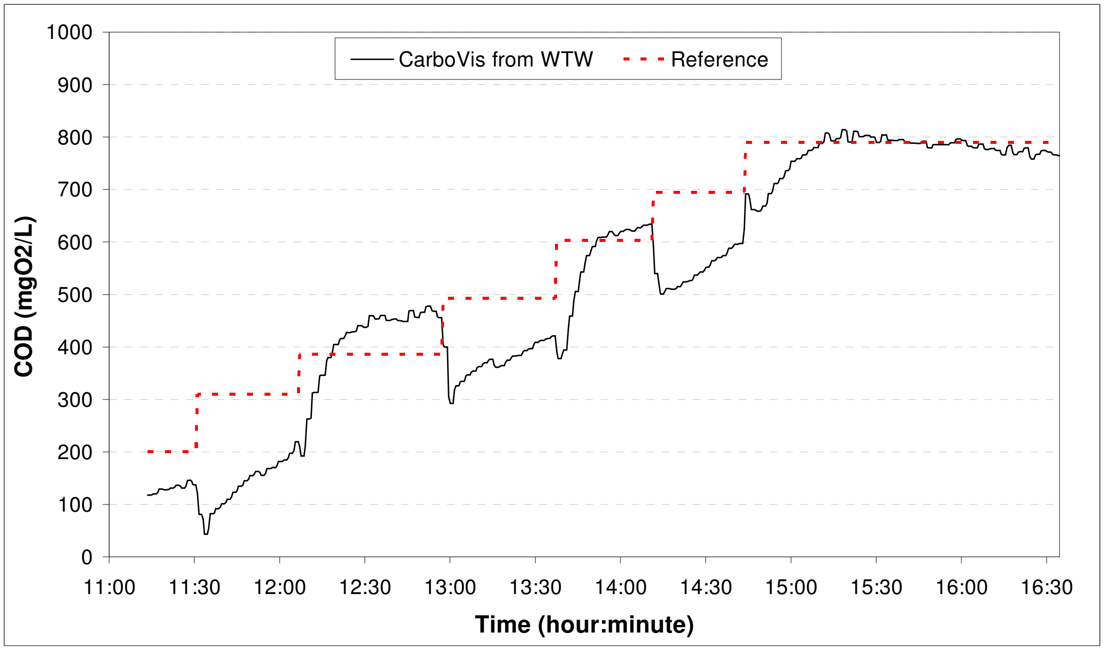


Figure 48: Linearity test of Carbovis from WTW (1/2)

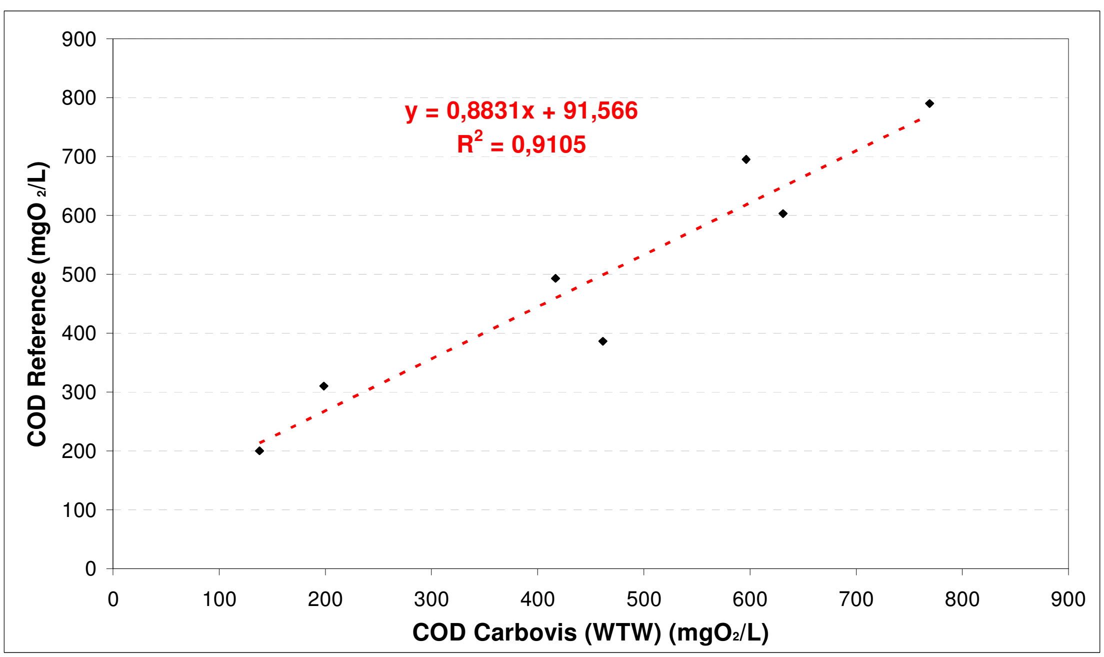


Figure 49: Linearity test of Carbovis from WTW (2/2)

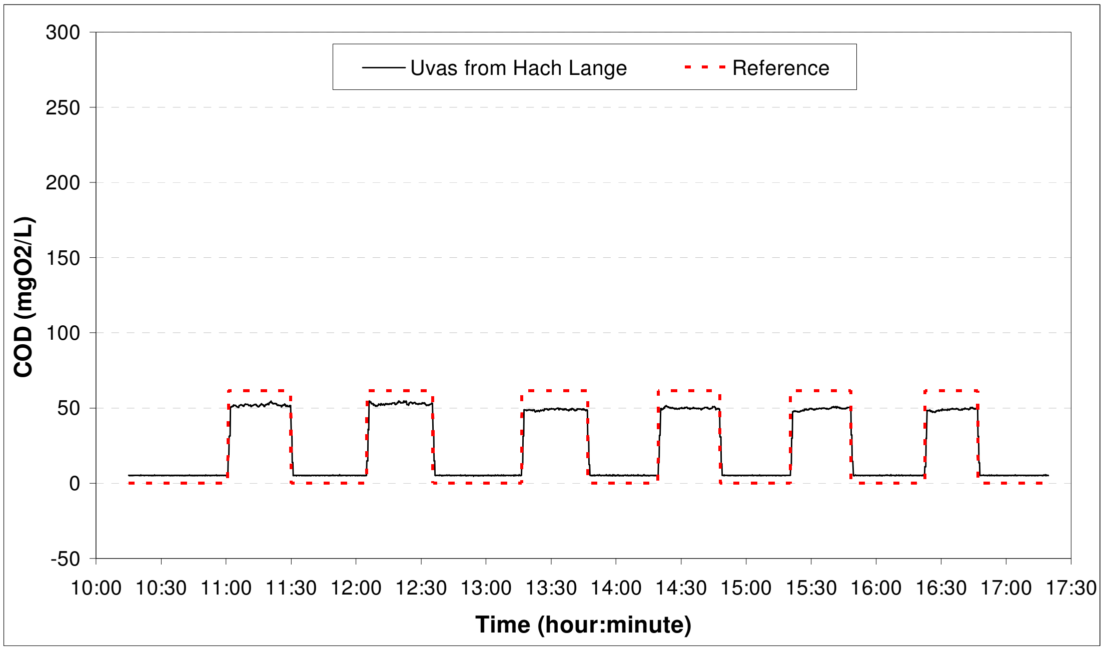


Figure 50: Accuracy test of Uvas from Hach-Lange

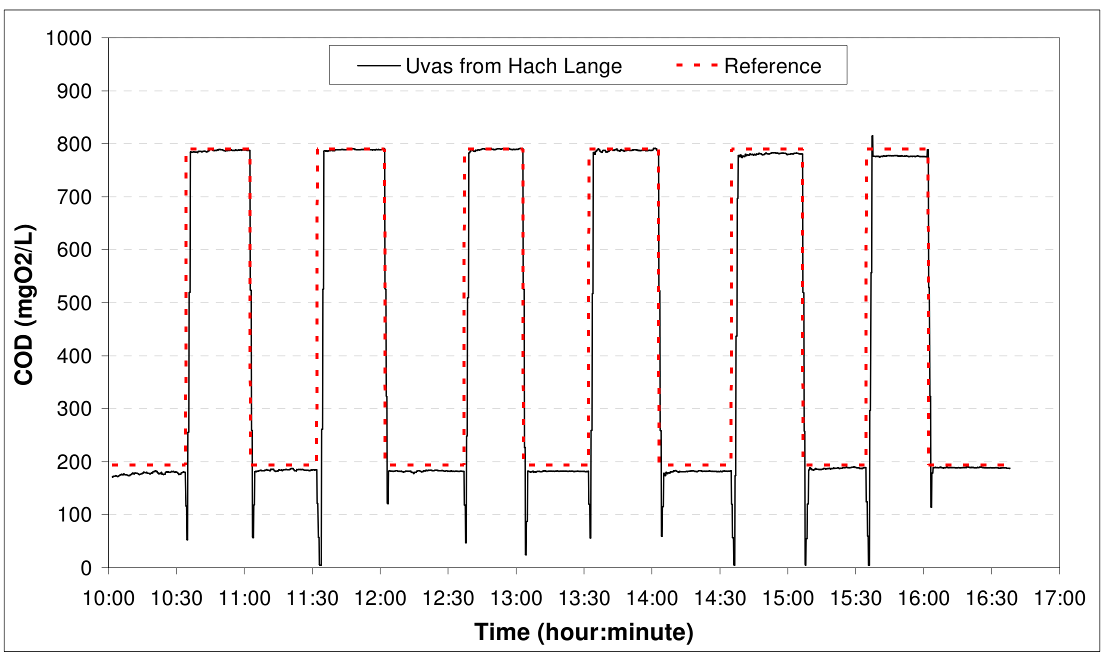


Figure 51: Dynamic test of Uvas from Hach-Lange

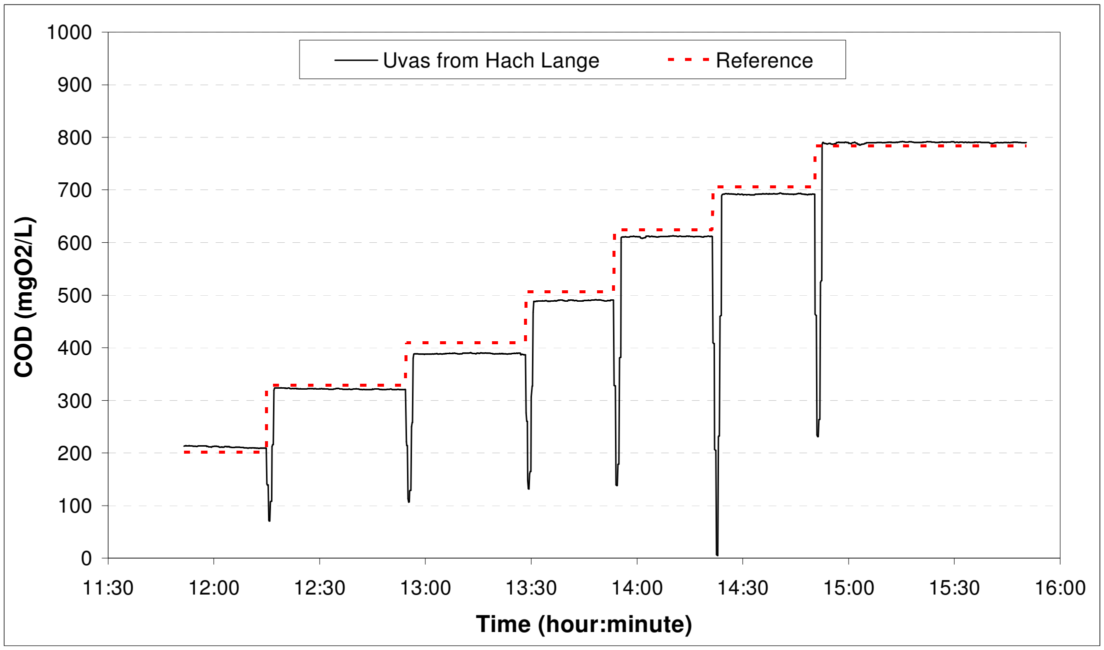


Figure 52: Linearity test of Uvas from Hach-Lange (1/2)

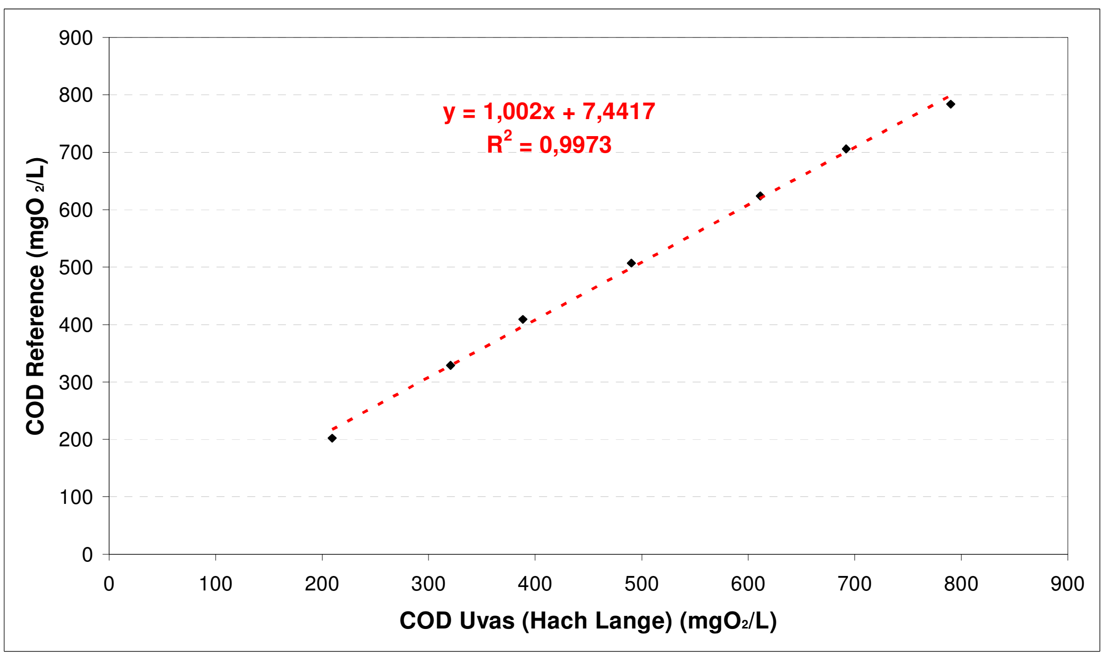


Figure 53: Linearity test of Uvas from Hach-Lange (2/2)

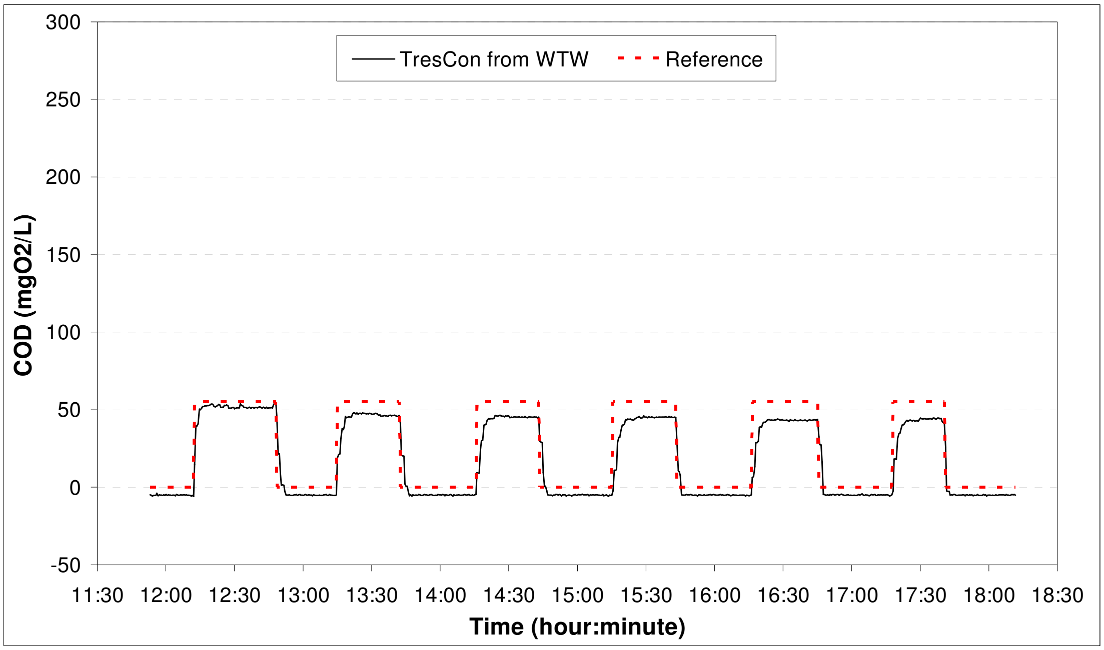


Figure 54: Accuracy test of TresCon SAC OS210 from WTW

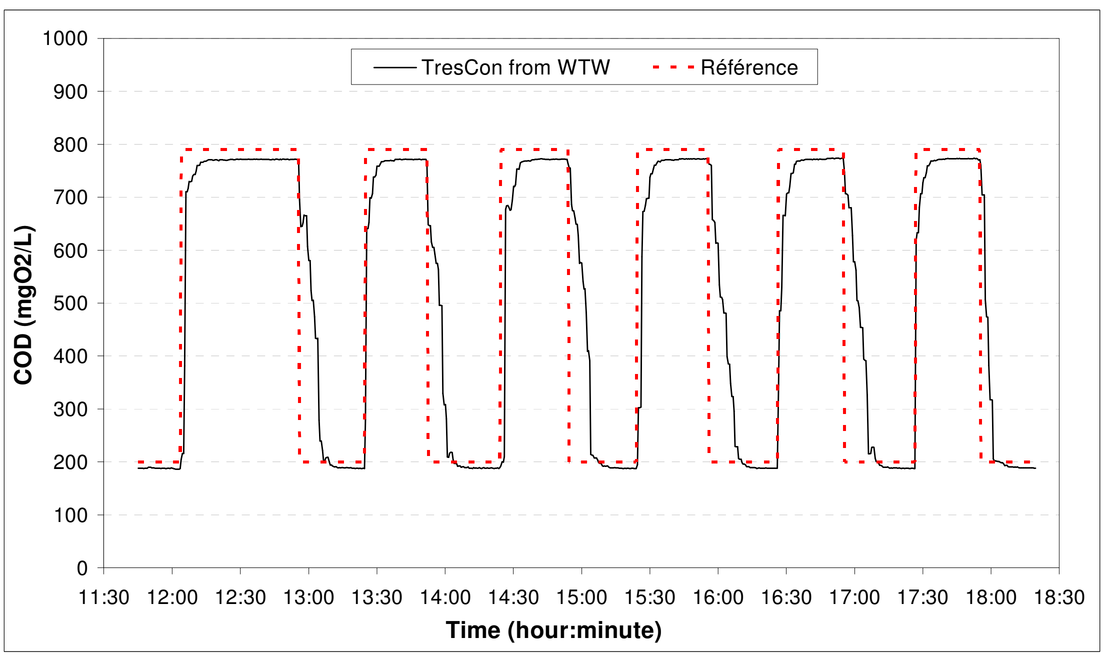


Figure 55: Dynamic test of TresCon SAC OS210 from WTW

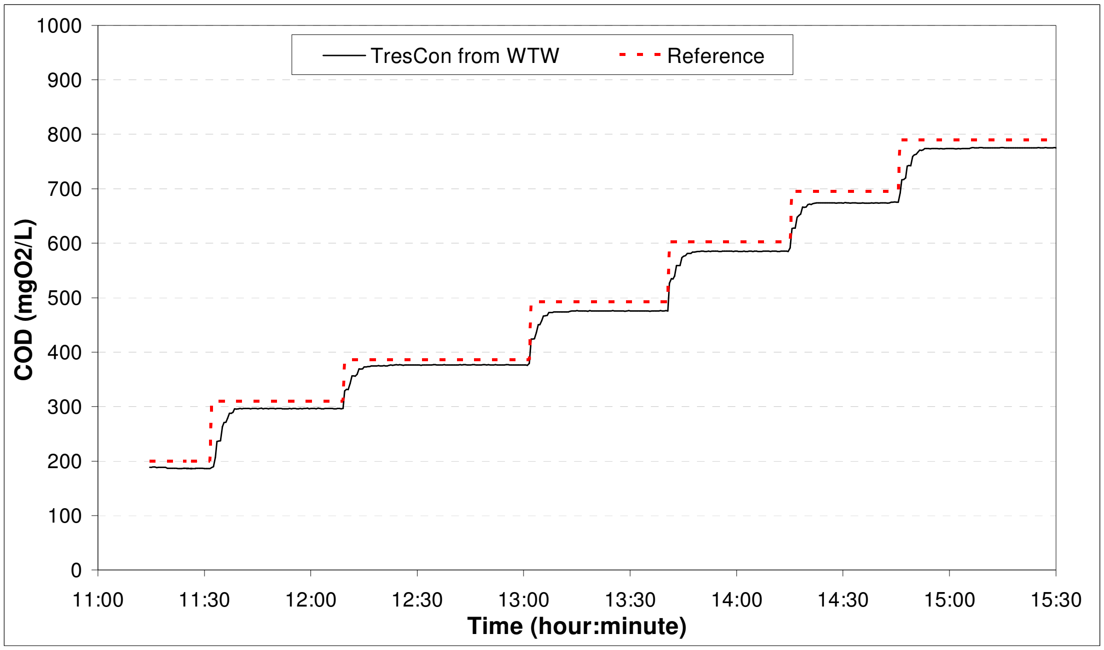


Figure 56: Linearity test of TresCon SAC OS210 from WTW (1/2)

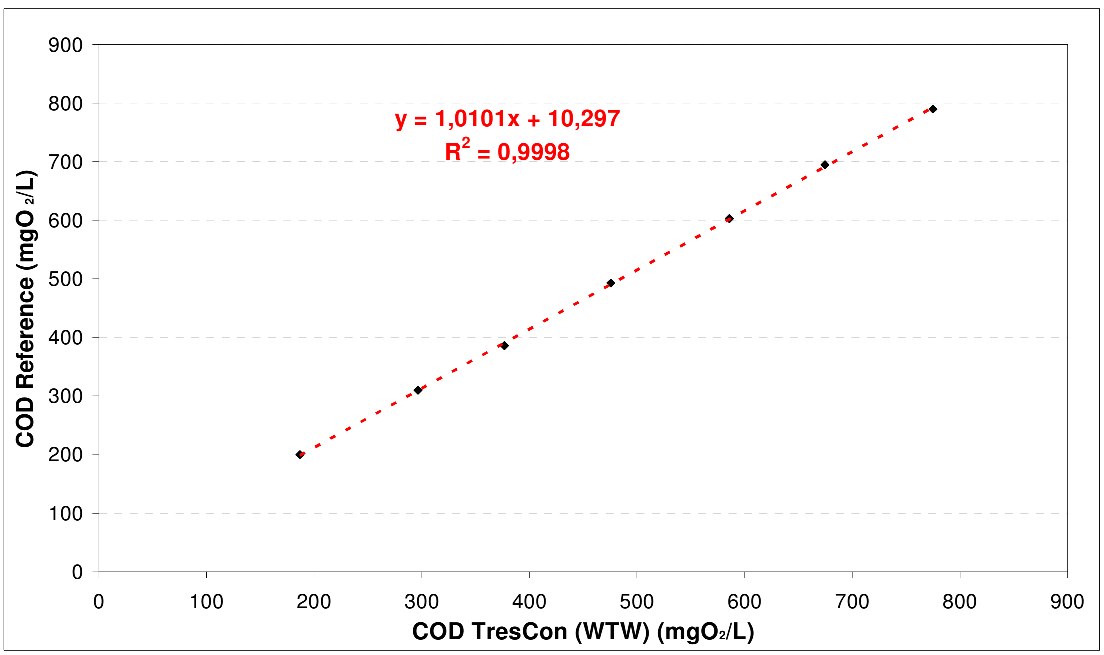


Figure 57: Linearity test of TresCon SAC OS210 from WTW (2/2)

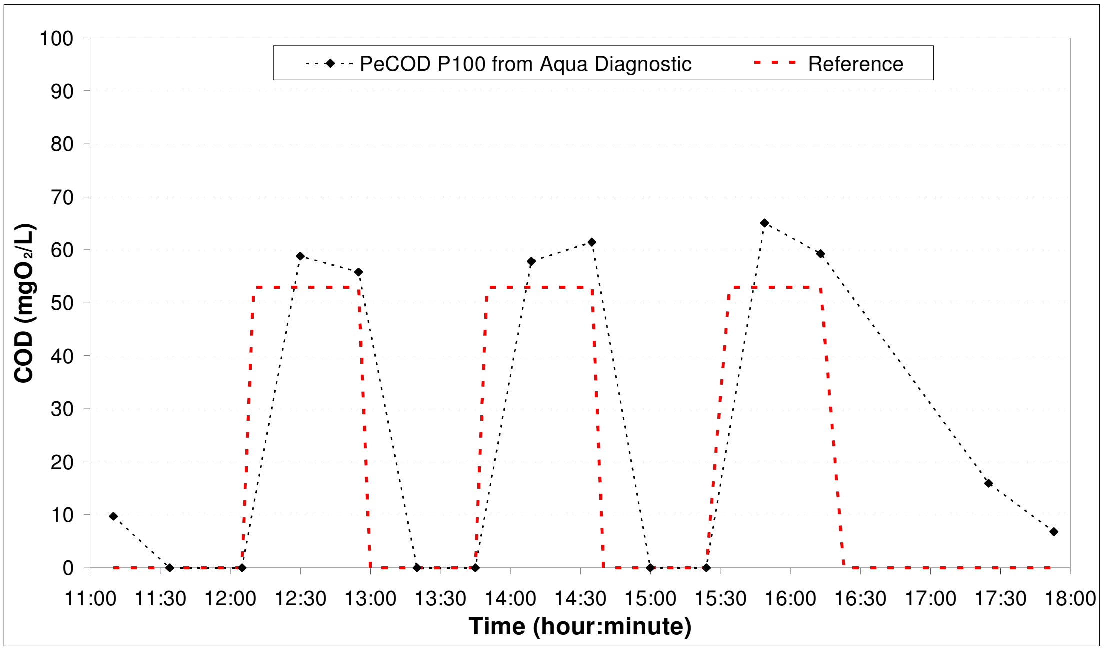


Figure 58: Accuracy test of PeCOD P100 from Aqua Diagnostic

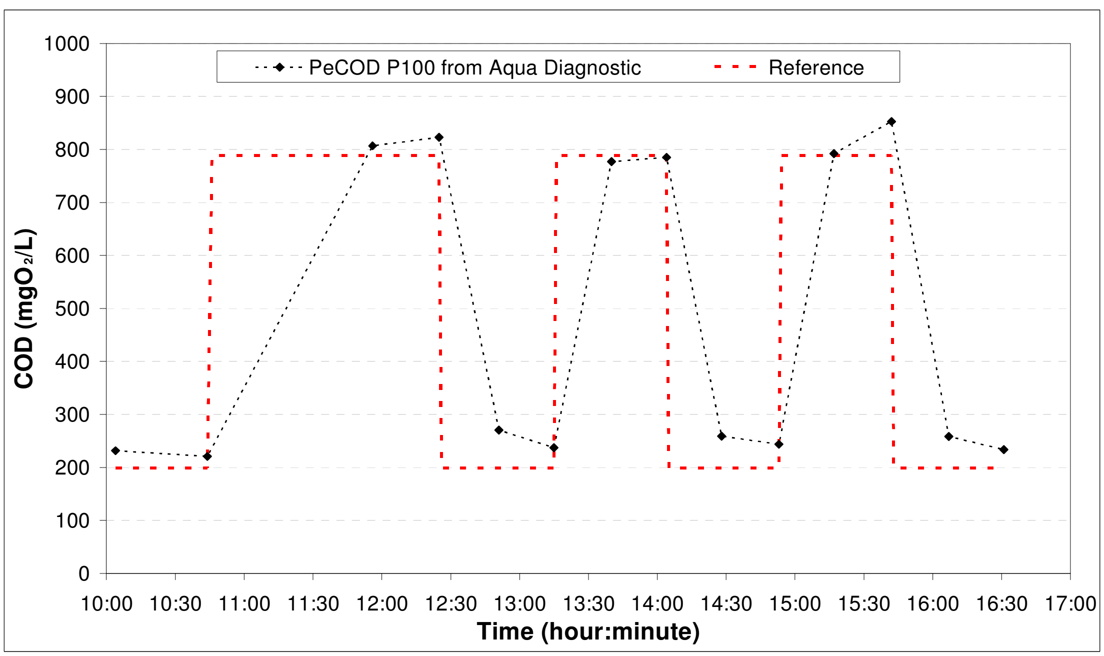


Figure 59: Dynamic test of PeCOD P100 from Aqua Diagnostic

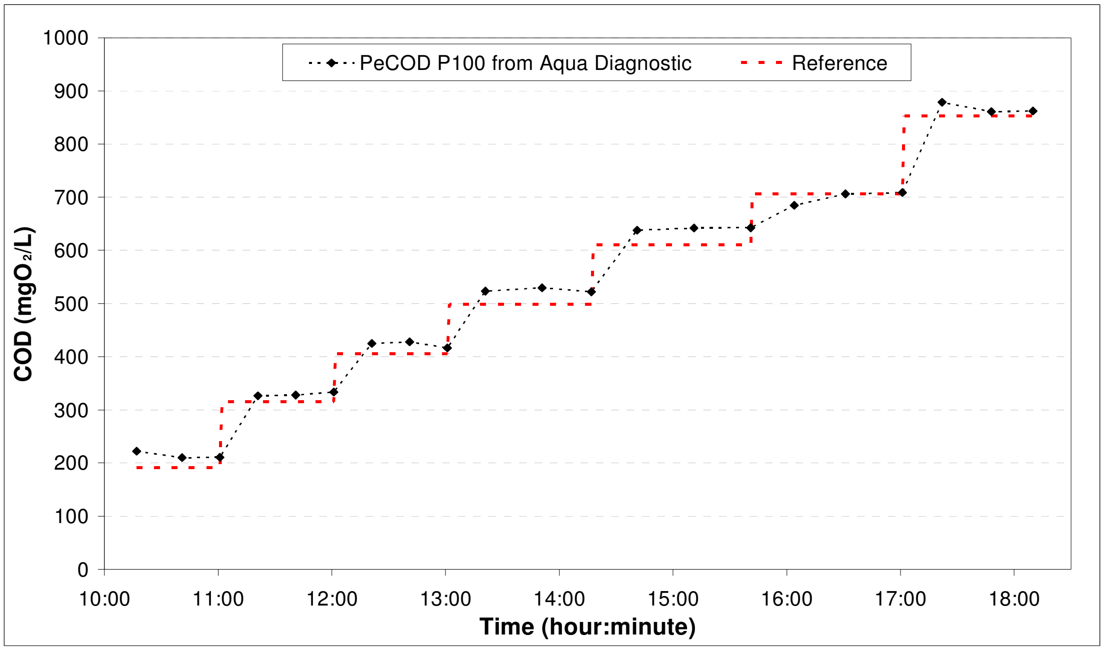


Figure 60: Linearity test of PeCOD P100 from Aqua Diagnostic (1/2)

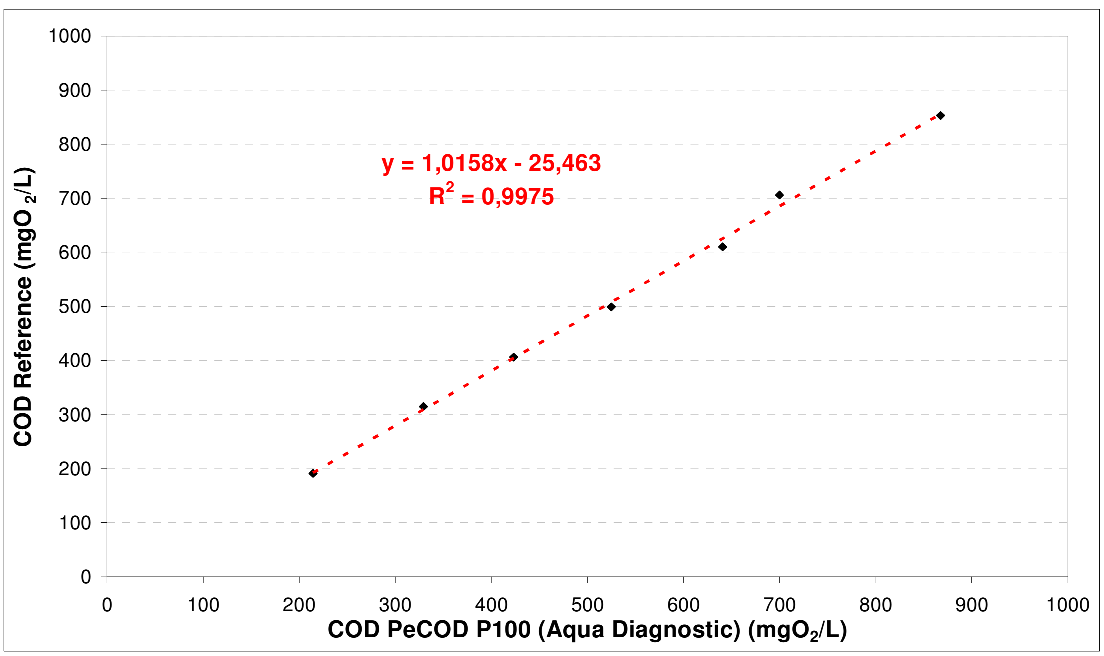


Figure 61: Linearity test of PeCOD P100 from Aqua Diagnostic (2/2)

Ortho-phosphate results

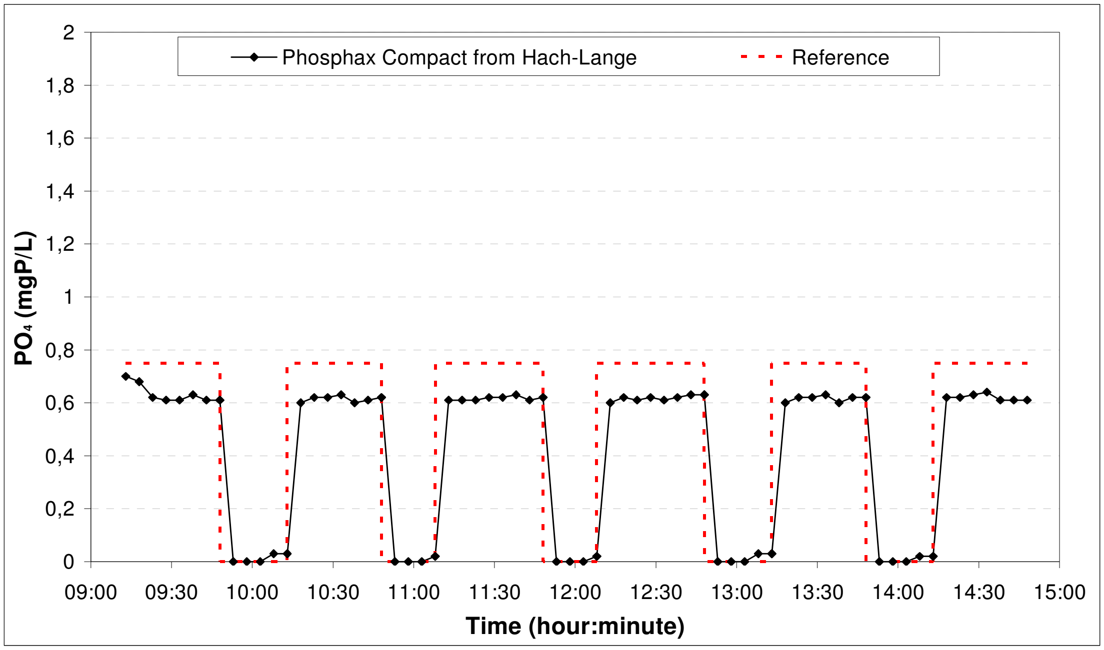


Figure 62: Accuracy test of Phosphax Compact from Hach Lange

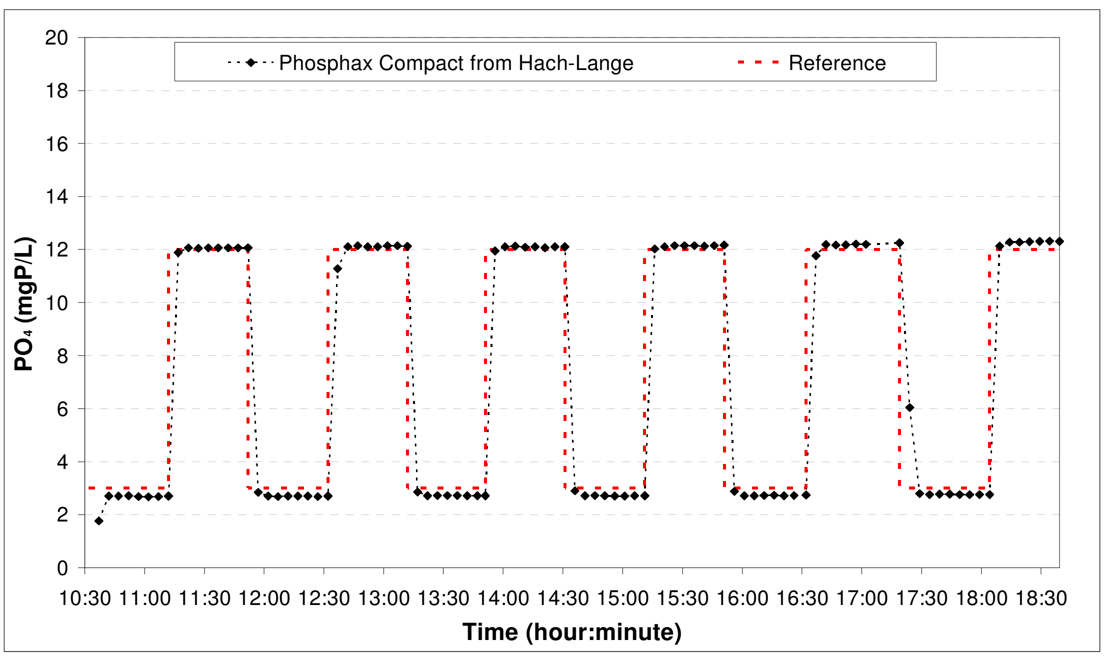


Figure 63: Dynamic test of Phosphax Compact from Hach Lange

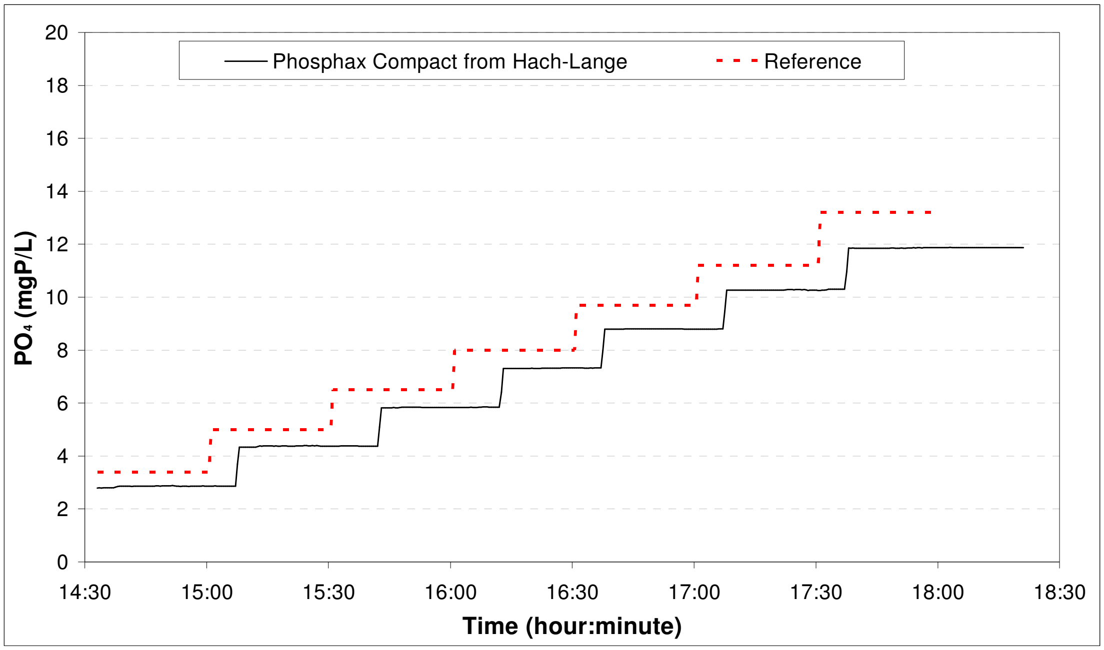


Figure 64: Linearity test of Phosphax Compact from Hach Lange (1/2)

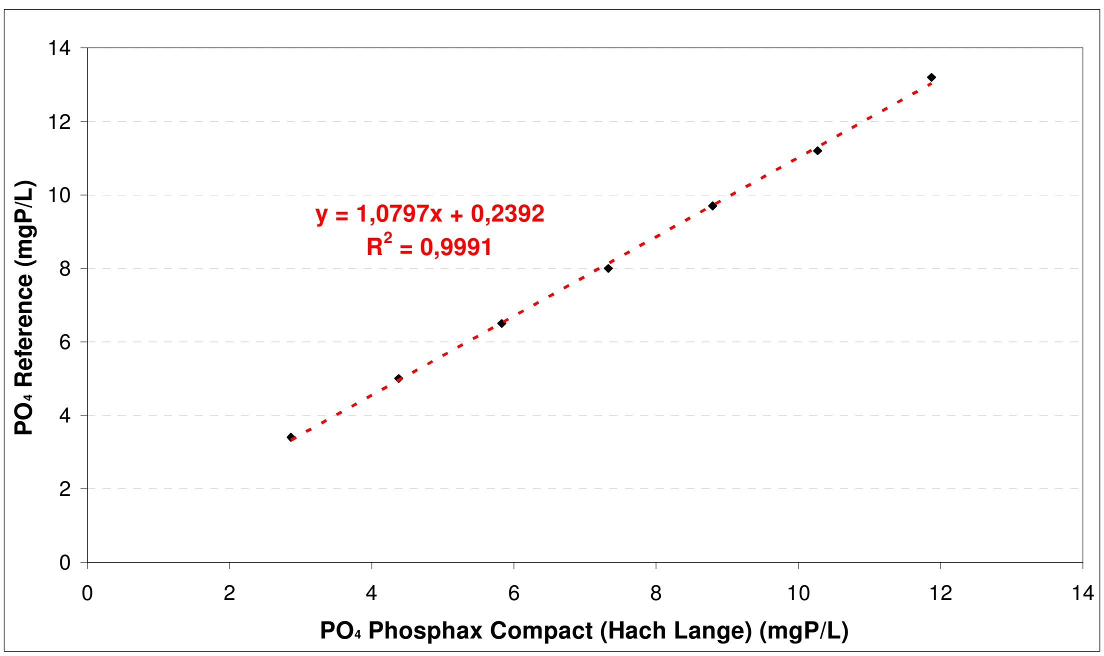


Figure 65: Linearity test of Phosphax Compact from Hach Lange (2/2)

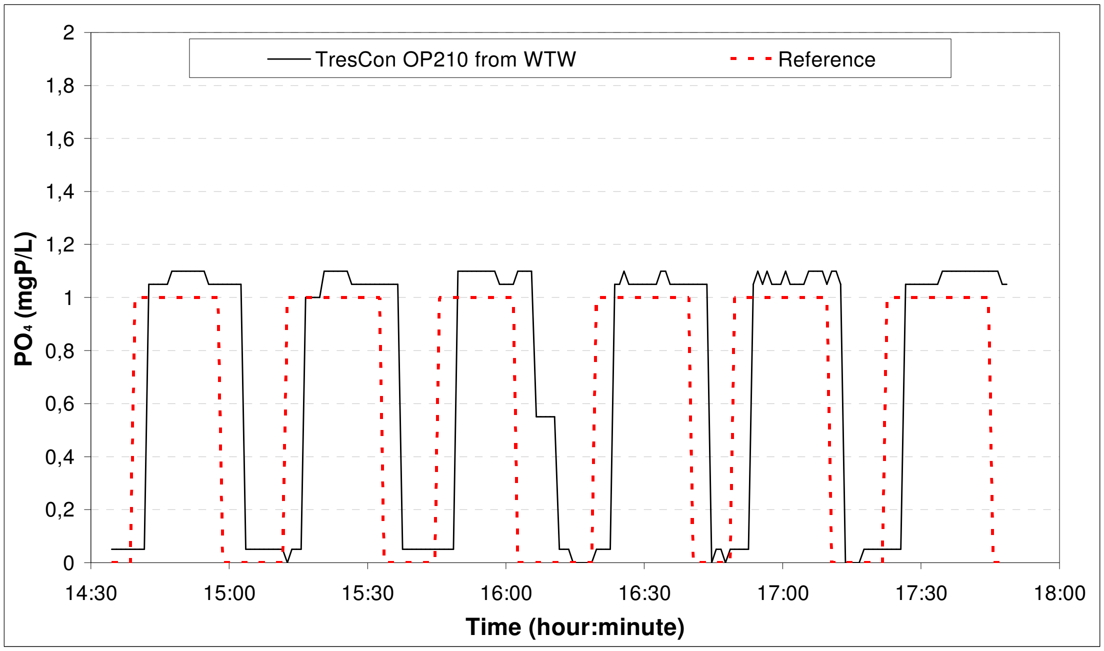


Figure 66: Accuracy test of TresCon OP210 from WTW

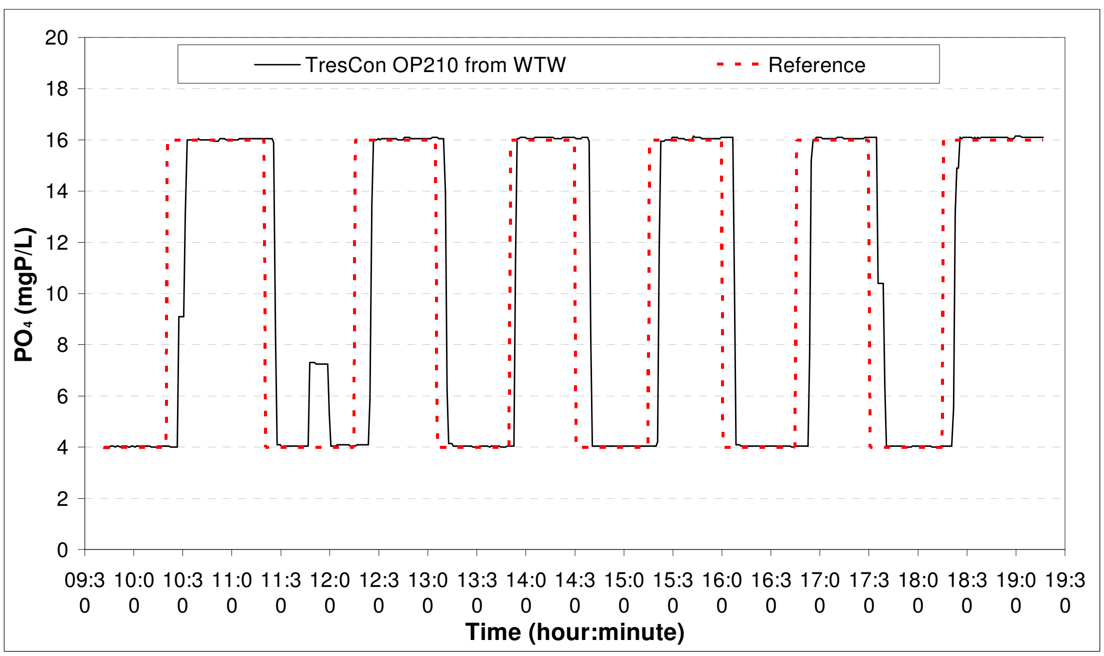


Figure 67: Dynamic test of TresCon OP210 from WTW

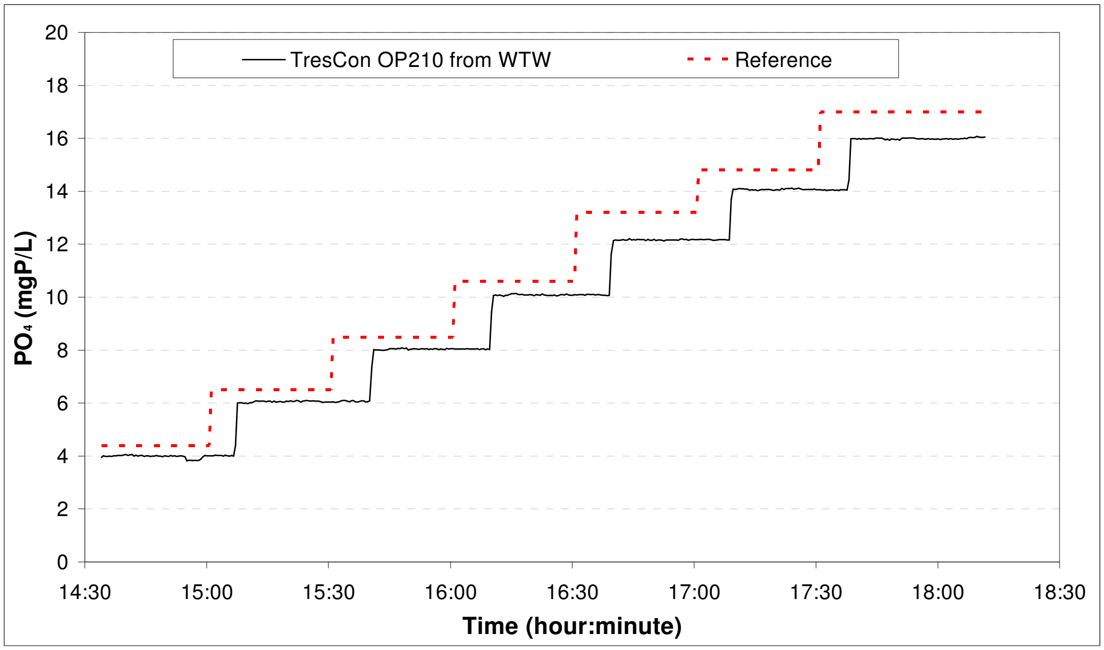


Figure 68: Linearity test of TresCon OP210 from WTW (1/2)

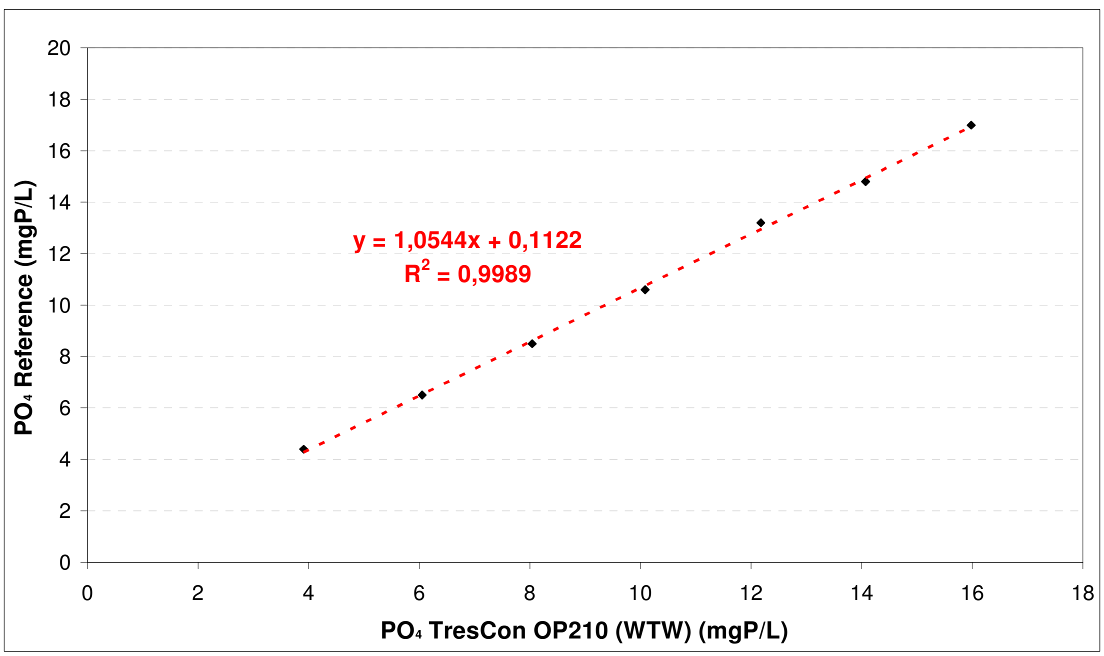


Figure 69: Linearity test of TresCon OP210 from WTW (2/2)