Towards the Automation of Water Quality Monitoring Networks

John B. Copp, Evangelina Belia, Christian Hübner,
Mario Thron, Peter Vanrolleghem and Leiv Rieger

Abstract— **The implementations of water quality monitoring networks have a number of inherent engineering challenges and the automation of the data collection and analysis only adds to that complexity. This paper has been written to discuss the challenges and solutions that have been developed within the framework of an industrial/academic partnership. Water quality monitoring stations are important tools in the area of environmental water science; however, traditional monitoring station installations and their maintenance tend to require more effort than desirable. Common sensors are not easily integrated into fieldbus systems and the lack of storable meta data (status, calibration information, location,…) available from sensor devices in this field, requires additional effort on the part of the owner if a fully utilizable database of meaningful values is to be constructed. An approach is proposed to automate this effort by providing an electronic catalog of predefined devices that can be input by the user during setup or read from the sensor in real-time. Automated data evaluation, alarm triggering and real-time data 'correction' are all being developed with an aim to create fully documented long-term databases of usable and meaningful water quality data. And finally, to initiate improvements in the area of monitoring automation, some thoughts on the future of advanced fieldbus systems are presented.**

I. Introduction

WATER quality monitoring stations are important tools in the area of environmental water science. They are commonly used in rivers and sewage treatment systems to measure and record various properties of the water. The aim of these activities is typically to increase our understanding of the water quality in a particular area and take, if necessary, corrective action if problems exist. The data being collected can also be invaluable if a mathematical model of the water system is being developed. But data collection is only beneficial if the data being collected is accessible, accurate and verifiable when anomalies are recorded.

The selection of appropriate hardware combined with the design and implementation of the station software must be done with care to minimize installation, operation, and maintenance costs. These costs were the focus of the initial development discussions surrounding this monitoring station concept. These discussions quickly revealed several engineering challenges that would contribute to these costs and have to be overcome including the commissioning and parameterization of the fieldbus and the sensor devices as well as the automation of the data collection, evaluation and 'correction' in real time. This paper addresses the approaches that have been developed to resolve these challenges and the automation of the process to encourage future improvements.

Manuscript received March 8, 2010.

John B. Copp is a Principal with Primodal Inc., Hamilton, ON, Canada L8S 3A4 (905-523-8958; e-mail: copp@primodal.com)

Evangelina Belia is a Principal with Primodal (US) Inc., Kalamazoo, MI (e-mail: belia@primodal.com)

Christian Hübner and Mario Thron are with Institut für Automation und Kommunikation e.V. Magdeburg (ifak), 39106 Magdeburg, Germany

Peter Vanrolleghem is at the Université Laval, Quebec, QC., Canada (e-mail: peter.vanrolleghem@gci.ulaval.ca)

Leiv Rieger was at the Université Laval, Quebec, QC., Canada (e-mail: rieger@envirosim.com)

II. The Station Platform

Typical water quality monitoring is still achieved using a data logger in a stand-alone system, but recent research is focusing on the development of monitoring networks that integrate the information from different locations into knowledge about whole water system under study [8]. The development of monitoring networks instead of individual stations leads to new demands on bidirectional data exchange, i.e. various telemetry options, safety issues and accessibility.

Three main reasons tend to limit the use of water monitoring stations [7]: a) the lack of standardization; b) data quality problems leading to data graveyards that do not provide the required information; and, c) insufficient flexibility within the stations leading to problems when new sensors need to be connected or when the focus of the project changes.

The vision of the next generation of water quality monitoring networks is the monitoring station under development and described here. Besides the focus on new automated data evaluation methods, this monitoring network concept combines technology with the highest possible flexibility. This flexibility enables the connection of various and multiple sensors from different manufacturers, different measuring locations and different monitoring goals [7].

A. *Software*

The heart of the monitoring station system is a robust software framework. The software backbone of the station and server network allows the simple connection of various modules through a specified API (Application Programming Interface). Some modules provide basic functionality like data input or output but the main function for this framework is to provide a flexible structure. Access to the framework via defined plug-in interface settings allows for the future incorporation of new developments and the implementation of user-developed data evaluation modules, but by locking the framework (the framework code is not open to the end-users), users are guaranteed robust basic functionality. This framework design combines the necessary controls and robust operation with the required flexibility. Fig. 1 shows the monitoring station concept.

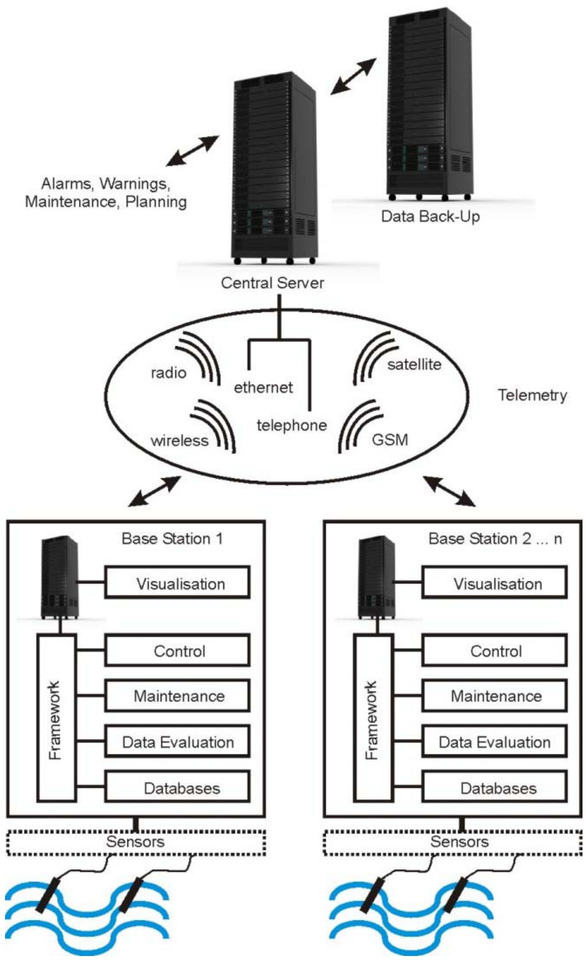

Fig. 1. Set-up of the monitoring station water quality monitoring network.

B. Hardware

The basis for the network concept is that the platform should be the same for all measuring locations. It is recognized that space, energy, and environmental conditions at any one location may dictate the final configuration, but the basic platform is a common design. That is, the basic unit (a box with computational capability and I/O units) will be the same, but the power supply, data transmission, and climate control options will vary. The station itself could be housed in a trailer or delivered as a stand-alone box to be used directly with in-situ probes. Specifications have been developed using the highest standards in terms of durability, robustness and data safety.

The base station is designed to handle multiple sensors irrespective of the manufacturer. Connection of a new sensor automatically triggers the creation of storage capacity and standard visualization. Where possible, meta data from the sensor is used to limit the required installation effort for the new measuring devices. Plug-and-play capability is not feasible (due to standardization problems), but a list of pre-configured sensors facilitates the connection of new devices.

C. Perspectives

The monitoring station system has been designed to provide a high-level platform for all kinds of monitoring tasks, and eliminate the design errors numerous other attempts have experienced. The flexibility of this new monitoring network concept enables different monitoring tasks at different measurement locations. The most commonly used data transmission protocols (between sensor and base station) are provided so the user can select the best suited sensor for the application at hand; all of this being independent of the specific monitoring station capabilities.

Various telemetry options, low energy demand and proactive maintenance concepts enable remote use of the monitoring network. However, the most important step forward is the advanced data quality evaluation. Data quality is paramount and doing data quality evaluation in real time will eliminate the danger of building more data graveyards.

III. Challenges and Achievable Solutions

The design of the monitoring station allows for the integration of communication options (sensor to base station) by means of software modules. Traditional 4-20 mA and serial connections are available, but at this stage of the hardware setup, transmission of the measurement values from the on-line sensors is solely based on PROFIBUS technology [2]. This is the preferred communication option and is justified by the fact that many sensor manufacturers support Decentralized Peripherals (DP) or Process Automation (PA) connections to/from their products [6].

A. Fieldbus and Sensor Commissioning

In a perfect world, it would be ideal if the monitoring station could incorporate a single fieldbus with plug-and-play capabilities. One can envision that a new transmitter would be physically connected to the bus and it would be detected automatically, its measurement channels automatically integrated and all its meta data queried. No prior device-specific configuration information would be required in this scenario and the installation would be controlled and configured by the bus.

In reality though there is no single communication technology that is supported by all the major sensor manufacturers in the water monitoring industry. Even worse some manufacturers do not provide any standardized connection or fieldbus integration for their devices but instead enforce proprietary methods that require tailored engineering solutions. The approach taken in the monitoring station software to cope with this kind of problem consists of defining an interface that is based on abstracting various device access technologies. This approach negates the need to use a particular access method (e.g. PROFIBUS) directly. For each access technology required by any supported sensor device, a separate implementation of this interface must be created. These implementations are provided in the form of separate software modules (e.g. plug-ins) that can be dynamically loaded by the monitoring station on demand without rebuilding the entire software.

One particular implementation of the abstract device access interface currently available realizes PROFIBUS support. This fieldbus was chosen to be the primary communication technology for the monitoring station because it is supported by many sensor manufacturers and has gained wide acceptance in the automation industry. Based on experience from the first stage of the project, the integration and commissioning of new PROFIBUS devices requires much more engineering effort than would be desirable compared to the ideal scenario outlined at the beginning of this section. Whenever the installation engineer wants to add a new PROFIBUS sensor transmitter to the bus, a number of steps have to be followed:

1) A numeric identifier (the bus address) has to be selected and assigned to the device. The engineer has to make sure that this identifier is unique across the entire network of PROFIBUS DP and PA devices either by maintaining a list of assigned identifiers or by checking the identifier setting of each device in the network.

2) Depending on the sensor transmitter type an additional engineering step might be necessary that involves configuring the cyclic telegram, i.e. specifying which measurement values together with their state and quality indicators are placed in what order into the fixed-size data structure that is sent to the PROFIBUS master.

3) On the fieldbus side of the monitoring station further commissioning steps have to be taken. The newly added transmitter device must be included in the PROFIBUS configuration using a device-specific GSD file (Generic Station Description, [5]) that contains the communication characteristics.

4) Additionally, the unique bus address of the new device that has been previously assigned must be entered into the PROFIBUS configuration.

5) The PROFIBUS configuration must also be adjusted to accept the configuration of the cyclic telegram programmed in (2) above. Essentially the same number of bytes and the same telegram structure must be repeated in the PROFIBUS configuration because the syntactical information of the telegram can not be imported from the device but must be duplicated manually.

In the actual monitoring station software the measurement data from the PROFIBUS transmitters is retrieved using a protocol-specific syntax. In the case of this PROFIBUS setup, the numeric identifier (bus address) of the device and the byte offset relative to the start of the cyclic telegram are used to form the identifying syntax. Unfortunately, the meaning and description of the accessible data items cannot be retrieved using this or any other interface because that information is not transmitted in the case of cyclic PROFIBUS communication. This lack of self-descriptive data creates the need for additional configuration steps. In the case of the monitoring station software, the necessary meaning and descriptive information about the data items is provided in a device catalog.

The device catalog describes a set of available devices that can be used in the monitoring station. In this context the term "devices" means a well-defined combination of transmitters (e.g. the PROFIBUS slaves) and connected sensors. The relationship between the station configuration and the entities of the device catalog is shown using the simplified UML diagram in Fig. 2. The main part of the device catalog is the list of device descriptions. Each of these descriptions references a device accessor that specifies the appropriate plug-in for the required implementation of the abstract device access interface described earlier in this section. The device description also describes a set of measurement slots. In this context a slot refers to a named data item that consists of an address, data type and data dimension (size of data vector). The implementation of the corresponding device accessor uses this information to read the current value of a given slot. In the case of PROFIBUS devices the address component is just a number that specifies the offset of the intended data inside of the cyclic telegram.

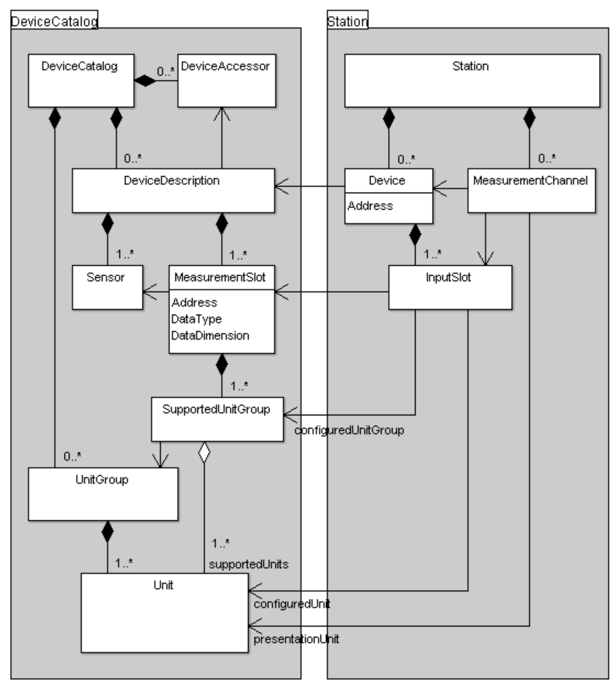

Fig. 2. Simplified UML diagram of station configuration and device catalog

The measurement values typically have an associated unit (^oC if the measurement is temperature for example). To enable unit conversions, all available units are specified in the device catalog. A unit group (e.g. "temperature") defines a set of units (and conversion factors) that can be interchanged (e.g. ^oF, ^oK, ^oC) thus enabling the transparent conversion between appropriate units.

The part of the station configuration relevant for data acquisition is displayed on the right side of Fig. 2. The station configuration describes all variable aspects of the setup of a particular monitoring station, most importantly the installed transmitter devices and measurement channels. A measurement channel defines what sensor data shall be collected and how this collection is performed. When adding a new measurement channel, the input slot and device that will be used to read the data must be specified. Unlike adding a device to the network, the configuration of the devices within the station software is intuitive because the installation engineer only needs to select the appropriate device description from the device catalog and specify the address. Additionally, it might be necessary to specify the configured unit and unit group of each input slot if the corresponding measurement slot supports more than one unit and unit group. The device catalog approach attaches meaning and structure to otherwise anonymous data items and thereby reduces configuration effort.

B. Local and Remote Sensor Parameterization

Apart from the settings needed for the device/station communication there are many other parameters in the sensors and transmitters that influence the measurements, e.g. slope and offset of linear correction functions and physical conditions like salinity of the water or altitude of the measurement location. For this reason the installation engineer must be able to perform parameterization of the devices during commissioning and maintenance operations. Most sensor transmitters, which have been investigated for the monitoring station project, can be directly configured and parameterized using a control panel with display. However, there is still a need for a simple and uniform parameterization facility within the station to minimize effort. In the best case all parameterization would be done using the monitoring station software relying on standardized and automated procedures. This would enable a simple and efficient implementation of remote parameterization, e.g. from the central server which is connected remotely to the base station. Another requirement for the station software in the context of parameterization is that all modifications of the parameter values be detected and saved to a log database. The need to log this information relates to data quality evaluation and the possibility that the sensor parameters have influenced the measurement values.

To date no uniform parameterization approach is available for fieldbus devices in general and sensor transmitters in particular. Even for PROFIBUS there are several competing parameterization approaches with varying support amongst vendors. Some vendors even employ proprietary alternatives for parameterization which makes the development of a generalized approach problematic. The most commonly used parameterization approaches are FDT (Field Device Tool, [1]) and EDDL (Electronic Device Description Language, [3]). In the case of PROFIBUS, both the FDT and the EDDL approach are based on acyclic data access using the DPV1 standard extension. In theory this means that any client software (e.g. the monitoring station software) should be technically able to read and write single parameter values of any DPV1 compatible PROFIBUS device. However, some devices exhibit complex interdependencies between parameters and may require a certain write order. Moreover, additional descriptive information needs to be attached to parameters as the meaning of values is not always self-evident. This is especially true for enumeration-typed parameters that are encoded using an arbitrary integer mapping. If all these difficulties were to be accounted for, a custom solution would essentially resemble major parts of the EDDL approach. Therefore, a fully integrated device parameterization feature is not feasible for the monitoring station at this time.

The consequence of these findings is that external tools have to be used for parameterization. For convenience, these tools can be launched from within the station application. In order to meet the requirement of logging parameter changes, the current parameter values of all configured devices are read (if possible) and compared to the last known value. For reading device parameters, detailed knowledge of parameters (e.g. address and data type) must be available. In the case of PROFIBUS, parameters are read using the acyclic services. Hence, the parameter address consists of four numbers: the slot and index information that refers to a certain data block, the length of this block and the offset of the value inside of the block.

Testing this approach revealed problems when reading acyclic data on different PROFIBUS sensor transmitters. In particular it was found that one DPV1-certified device rejected valid read requests if the queried data length was longer than the length of the available data. The approach used in this case always requested the maximum possible data length (240 bytes) ignoring the device restrictions given in the GSD file during configuration. This compatibility issue might be solved by using a low-level interface of the PCI card. However, this workaround is vendor-specific and when implemented substantially increased implementation complexity. Another observed disadvantage of the chosen approach is that device-level error messages are not accessible. Therefore, special hardware equipment (e.g. a bus monitor) and expert knowledge was necessary to investigate fieldbus problems.

Remote parameterization is desirable; however, the external tools do not tend to allow remote parameterization. A very simple but not always feasible approach is to use VNC (Virtual Network Computing) tools for PC remote control or the Windows built-in Remote Desktop Services. Using these options allows the user to launch the parameterization tools on the station from any PC. This is only possible if a TCP/IP connection with enough bandwidth is available for the station, which will not always be the case. Another idea is based on the manual import and export facility that some parameterization tools provide. The actual parameterization can be done on any PC without the need to be connected to the station. The parameters are exported to a file and transferred to the monitoring station. On the station side the file is imported by the corresponding tool, which writes the parameter values to the device. Unfortunately, in practice, this approach has a number of other complicating aspects that are not easily resolved.

C. Measurement Quality Representation

Besides the raw measurement values, sensor transmitters typically offer additional quality and state information that need to be obtained and considered by the monitoring station during automated data evaluation processes. Currently, there is no common standard for indicting measurement quality. For PROFIBUS devices, a simple quality indicator is returned together with a time stamp and the measurement value. But this quality only refers to the measurement value transportation from the sensor transmitter to the station and does not consider dedicated quality information offered by the sensor transmitter.

D. Data Quality Automation

The three main tasks in environmental monitoring are to obtain sufficient data quality with respect to the objectives, to transfer and store these data and to translate the data into valuable and usable information. In recent years, numerous sensors have been developed and put on the market, but there is still a lack of suitable monitoring concepts to guarantee reliable measurements. If the measuring devices are maintained and monitored in a sensible way, their accuracy, reliability and acceptance will be greatly increased.

The automation of the data evaluation process in real-time is a key component of the monitoring station software. New systems for data evaluation at the station level will result in improvements to the available information which will subsequently result in better data being stored as well as better operation and better planning.

A recently developed data quality monitoring concept requires a minimum of monitoring effort but has a high probability of detecting systematic errors [9]. However, a compromise must always be made between monitoring effort and safety with respect to errors as every single value input into a supervisory or an alarm system is not checked. Different continuous monitoring concepts have been developed recently, but not all errors can be detected with such approaches. Data quality monitoring (Fig. 3) can be broken down into a series of steps:

1) Evaluation of sensor status data: Calibration data and other status data are used to assess the sensor measuring quality and to flag each value with a quality indicator.

2) Automatic check of each single value: This "on-line" concept should guarantee the usability and reliability of the signal for control purposes or alarm systems.

3) Regular analysis of comparative measurements: This "off-line" concept uses reference measurements of grab samples to detect systematic errors and poor calibration. It is needed in addition to the "on-line" analysis because the automatic check can only detect disturbances from a basic state but will not be able to detect a constant offset, for example.

4) The last step is to integrate other measurements into the evaluation by using all available data including flow measurements and routine lab data. Existing redundancies and process knowledge will then be used to identify errors.

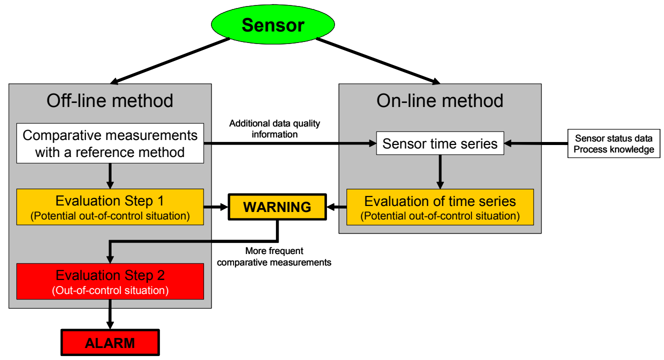

Fig. 3: Generalized data quality evaluation steps

Error and maintenance messaging is an important feature for these monitoring stations and particularly for stations in remote areas. A good messaging system will allow longer service or maintenance intervals with more safety against unrecognized station break downs. Alarm triggering on the central server will be the primary mode of notification, but the potential for connection errors between the base station and central server could require the triggering of messages via a variety of communication technologies such as telephone, pager, e-mail, or SMS to name a few.

Without exact criteria for identifying measuring errors, process upsets or limit violations the measured data are only numbers in a database. An innovative monitoring system has to assess the data with respect to data quality and the aspired objectives.

IV. Derived Requirements For Advanced Fieldbus Systems

According to the experience gained during design and development of the monitoring station, the established fieldbus technology is associated with high engineering costs that appear unnecessary and disappointing from the user's perspective. This section briefly outlines some requirements that advanced fieldbus devices could address in order to improve the situation.

The most obvious requirement is the support for true plug-and-play when connecting fieldbus devices. This would involve the automatic assignment of bus addresses to devices, similar to the DHCP (Dynamic Host Configuration Protocol, [4]) used in IP based networks. Device-specific configuration and description files (e.g. GSD files) should not be necessary. Instead, all connected devices would be queried and provide a human-readable display name together with further attributes like meaningful description and vendor name.

The core requirement for reduced engineering effort is the availability of comprehensive information. Every device connected to the fieldbus should provide all the information required for its operation and control by the network root. In case of a sensor transmitter, the station would query the existing measurement slots and obtain the necessary details to process the values. All available data items (for input and output) would be self-descriptive by containing various metadata like display name, description, data type, valid range, measuring unit, and quality indicator similar to PROFIBUS PA. The data items might be organized in a tree structure to reflect hierarchical relations, e.g. sensors and corresponding measurements. Providing comprehensive information directly from the devices would eliminate the need for the device catalog approach and would consequently minimize the engineering effort.

Information is especially important for the parameterization requirement. Similar to data items, all information about parameters (e.g. name, description, data type, and range of values) should be provided by the device itself. Parameters could be organized in a hierarchical fashion to simplify navigation. Additionally, a formalized description of interdependencies between parameters might be offered using a set of Boolean expressions (e.g. Mode = 'Saturation' implies Unit = '%') that yield true only if the parameter assignment is valid. With the help of these expressions a parameterization tool could automatically infer all valid modifications of a given parameter assignment. The consequence of providing all parameterization information directly by the devices is that no device-specific drivers (e.g. DTM in FDT) or external device descriptions (e.g. EDDL files) would be necessary to perform a parameterization of the device. True plug-and-play functionality would be a vast improvement over the current cumbersome approach that has to be used.

V. Advancing Water Quality Monitoring

The monitoring station approach developed by Primodal and discussed in this paper represents a significant step forward with respect to the storage of fully documented and usable water quality databases. By incorporating flexibility, computational capability, visualization, automated data analysis, alarm triggering, real-time data 'correction' and qualitative user input at source, the approach eclipses the traditional data logger/transmission system and represents a vision of the next generation of water quality monitoring networks. The use of advanced fieldbus systems and bidirectional communication enables these stations to take advantage of the advanced capabilities being incorporated into modern sensors. Data logging type stations are able to record and transmit data, but better data awareness is possible today and this development works towards the development of monitoring networks that integrate information from different locations into knowledge about the whole water system under study with a focus on new automated data evaluation methods to create better water quality databases.

VI. Conclusions

In this paper the experience gained by designing and constructing a monitoring station system for water quality networks has been presented. Several engineering challenges related to fieldbus commissioning and device parameterization were identified and feasible solutions were presented. The project has so far revealed that no single communication technology is supported by all sensor manufacturers and that fieldbus based solutions like PROFIBUS require much more engineering effort than first realized. This additional effort is the direct result of having to manually configure the devices on the network and the complications associated with device parameterization. The presented approach introduced a device catalog that enables the integration of new device access technologies via plug-ins. This approach also simplifies the effort required for the interpretation of the data stream from specific predefined devices. The importance of automated data quality evaluation for water monitoring stations was emphasized and the generalized steps for the evaluation process and maintenance messaging were outlined. Finally, some perspectives on advanced fieldbus systems were offered in an effort to initiate features that would help minimize the engineering effort including true plug-and-play support.

References

[1] FDT JIG (2009). FDT Interface Specification, Version 1.2.1, http://www.fdtgroup.org/en/documents/technical-documents.html

[2] IEC 61158 (2007). Industrial communication networks - Fieldbus specifications - Communication Profile Family 3: PROFIBUS PROFINET

[3] IEC 61804-3 (2006). Function blocks (FB) for process control - Part 3: Electronic Device Description Language (EDDL), Ed. 1.0

[4] Network Working Group (1997). Request for Comments: 2131 - Dynamic Host Configuration Protocol

[5] PROFIBUS INTERNATIONAL (2003). Technical Guideline - Specification for PROFIBUS Device Description and Device Integration, Volume 1: GSD V5.0, Order No. 2.122

[6] PROFIBUS INTERNATIONAL (2004). Profile for Process Control Devices, Version 3.01, Order No. 3.042

[7] Rieger L. and Vanrolleghem P. A. (2008). monEAU: A platform for water quality monitoring networks. Water Sci. & Technol., 57(7), 1079-1086

[8] Strobl, R. O., Robillard, P. D., Shannon, R. D., Day. R. L., and McDonnell, A. J. (2006). A water quality monitoring network design methodology for the selection of critical sampling points: Part I. Environ. Monit. Ass., 112, 137-158

[9] Thomann M., Rieger L., Frommhold S., Siegrist H. and Gujer W. (2002). An efficient moni¬toring concept with control charts for on-line sensors. Water Sci. Technol. 46(4-5), 107-116