

# Crucial elements and technical implementation of intelligent monitoring networks

A. Winkelbauer<sup>a</sup>, R. Fuiko<sup>a</sup>, J. Krampe<sup>a</sup>, S. Winkler<sup>b</sup>

<sup>a</sup> Vienna University of Technology, Institute for Water Quality, Resources and Waste Management, Karlsplatz 13/E226-1, 1040 Vienna, AUSTRIA (Email: awinkelbauer@iwag.tuwien.ac.at)

<sup>b</sup> Hach Lange, Austria

## Abstract:

Growing complexity of water monitoring instrumentation leads to specialized solutions in respect to sensor integration across several measurement device suppliers. Despite efforts of standardization for data interfaces and protocols, problems regarding the combination of several devices to gain the complete picture in terms of water quality remain. This assessment, especially accomplished from the perspective of a catchment area, requires a transition from sole use of data collectors towards an implementation of intelligent measurement networks. Several challenges and bottlenecks concerning distributed data collection are discussed starting with data acquisition up to the user-scope of utilizing data processing software. Finally, experiences using automated data inspection and export tools are discussed and a brief overview of expectable long term data availability is given.

**Keywords:** plausibility assessment; data abstraction; data to information; automated on-line monitoring networks

## INTRODUCTION

Operating water quality monitoring networks implies the requirement of several measurement probes for the assessment of the water quality state, which usually originate from different manufacturers. Even single manufacturers are applying different sensor interfaces in the analogue and digital domain, partially based on manufacturer-customized control software. Additionally, requirements for long-term measurement station operation referring to data stability have to be met. Besides that, automatic event based cycle control, monitoring of proper operation and alerting in case of failure need to be included. Finally, the demands for proper data combination, presentation and export particularly with regard to the end user perspective have to be satisfied.

One approach regarding the increase of data quality through so called “functional analysis” for three parameters is shown in Piñeiro Di Blasi et al. (2013). The paper focuses on a mathematical discourse regarding outlier detection and less on the architecture and instrumentation of on-line stations. In Ellis et al. (2013) a strategy using on-line water quality monitoring for the prediction of bacteriological failures is shown. Albeit, the implementation of data collection tasks and, just to mention one example, the implications regarding time-offsets between probe sampling and data interpretation itself, comes to short. Strobl and Robillard (2007) describe a distinguished approach on network design for water quality monitoring giving efficient guidelines and recommendations, especially regarding the sampling location and frequency. Nevertheless, the remarks concern primarily the design phase of the monitoring network and this paper, in contrast, is mainly based on a practical implementation.

To achieve the mentioned requirements the water monitoring network platform *i<sup>TUW</sup>mon* (intelligent information water monitoring networks) was developed and is described in detail by Winkler et al. (2012). The network platform consists of distributed measurement stations, a central database station and several data

examination workstations, which typically consist of workplace computers. This paper focuses on the pitfalls and bottlenecks during implementation of this technology. The domain is split into three logic layers: “data acquisition”, “data to information” and “data management and presentation”. The layers are described in detail after a brief overview of the therefore required measurement data channel abstraction. Following the description an overview of the implementation of the tool is given.

## INTELLIGENT MONITORING NETWORKS BASED ON THE CONCEPT OF MEASUREMENT DATA CHANNEL ABSTRACTION

The simplified monitoring platform structure is shown in Figure 1. The data acquisition task is commonly performed directly at the monitoring station. In contrast, data from external sources, for example from meteorological stations and monitoring data for the verification of compliance of the corresponding water framework directive released by BMFLUW (2006), are collected or calculated based on already collected data as database-administrated respectively database-calculated and therefore called “virtual channels” inside the main memory. These channels are subsequently displayed directly on-site, together with the sensor data from local probes (screenshot in Figure 2, left side: database icons in the channel list). Data to information takes place directly at the monitoring station (aggregation, plausibility checks and alerting) and in the central database (metadata integration of calibration, maintenance and laboratory data). Finally, data management and presentation is realized with *i<sup>TUW</sup>mon.Examine* (screenshot in Figure 2, right side). This tool is basically intended for data viewing and inspection; export routines to convert data into a standardized, interchangeable file format are also provided. The user interacts with each level of the structure by defining general set-up and site-specific plausibility assessment parameters at the monitoring station and through performing further data inspection and processing data which are obtained on his workplace computer, utilizing the connection of a virtual private measurement network.

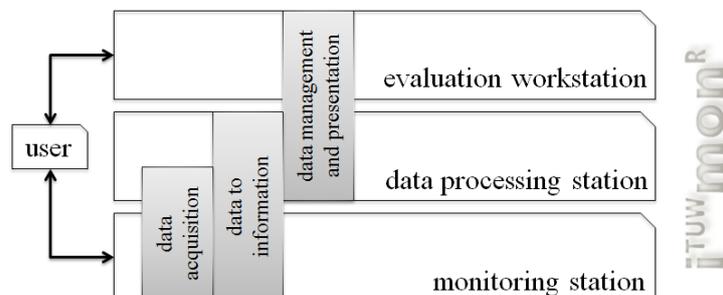


Figure 1 *i<sup>TUW</sup>mon* monitoring platform architecture

The software architecture is implemented using the graphical system design language LabVIEW as an established standard for implementing measurement and automation technology systems. A brief overview of LabVIEW is given by Elliott (2007). Rigorous considerations regarding software modularity and code reuse were obeyed so that for example the use of *i<sup>TUW</sup>mon.Examine* as a data inspection tool is possible for data sets which are imported semi-automatically as so called virtual channels – even without the comprehensive use of a complete monitoring station installation.

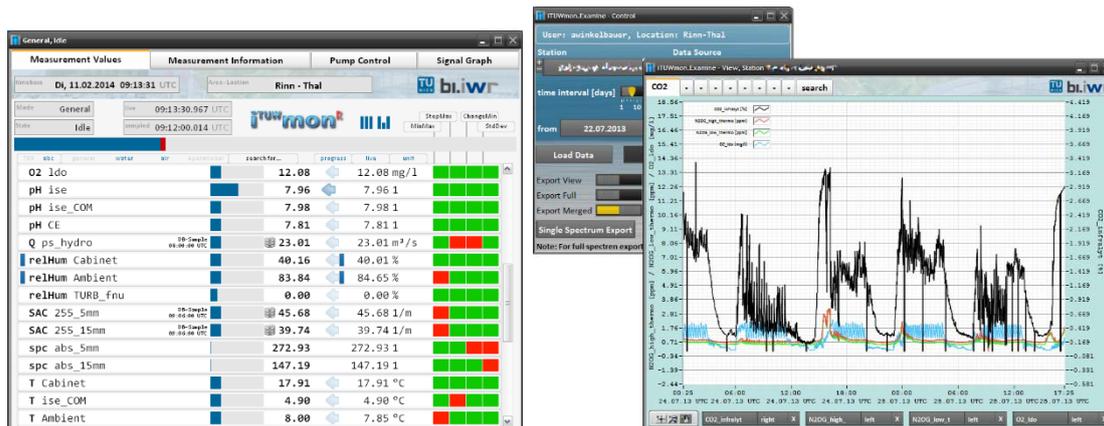


Figure 2 user interfaces: *iTUWmon.Measurement* (left) and *iTUWmon.Examine* (right)

The smallest common denominator to fulfill the monitoring tasks and meet the mentioned aspects is the abstraction of measurement values into abstract data channel objects (see Figure 3), which form the basic data structure in our solution. Every layer of the monitoring platform architecture works with a conglomerate of single data channels as basic units for measurement result processing and for metadata management during station operation. The data channel should be seen as abstract, atomic data object, which includes the measurement timing behaviour, actual conditions of operation and plausibility assessments across the named domains. The structure is instantiated multiple times in form of a software object in all involved systems, starting at the basic data collection system till the end-user toolset for data viewing and export.

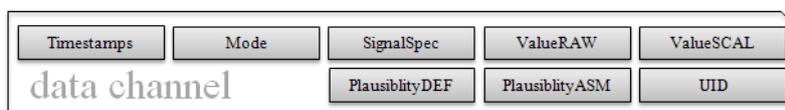


Figure 3 abstract data channel structure

The data channel consists of the following parameters: *Timestamps* contain the measurement timings in three time formats (time-UTC, -local and -MATLAB) and as primary key element it acts as first and most important data descriptor for the data structure. There are three measurement *modes*, which are selected statically in case of modes “General” and “Maintenance” or dynamically in case of exceptional measurement events. A typical event leading to more dense data recording in the time domain is a storm water event. During such events the mode “Trigger” is selected automatically and temporally more dense data sets are collected. Besides that, triggering of an automatic probe sampler is implemented and used to gain probes for further laboratory analysis. The *SignalSpec* (signal specification) contains information regarding measurement area and location, the physical data source, the parameter and unit definition as well as the data source description. The data source is segmented into several types of interfaces, e. g. analogue input, diverse serial interfaces, ModbusTCP, Profibus, spectrometer interfaces or external data (“virtual channels”). The measurement values at a time instance are stored as *raw* and *scaled values* extended by information about configurable data aggregation timeframes, for example automatic aggregation of plausible single measurements to one-hour averages. The structure is implemented as an array of values, enabling efficient storage including absorption spectra where one array element corresponds to an absorption value at a specific

wavelength. Definitions and results of on-site plausibility assessments (for details see Winkler et al. (2013)) are stored in the plausibility structures facilitating automatic data filtering for near real-time reaction on events (*PlausibilityDEF* for definitions and *PlausibilityASM* for assessment results). The type of assessments is user-configurable and extend of calculations is only limited by the processing power on-site. *UID* stands for unique measurement identifier and acts as a data fingerprint for single measurements. Hardware serial numbers, network environment, elapsed time ticks since a reference date and further aspects are used for calculation of the *UID*. In case of data faults the complete measurement dataset might be rebuilt using redundant repositories or databases which contain raw measurement datasets as a sort of hot backup.

### Data Acquisition

The formal tasks of this layer are the sensor interface integration and the operational timing control. In the implemented monitoring network sensor data is processed in terms of abstract monitoring data channel objects as shown above. The aim of device integration is to get as much information as possible on the shortest way from the sensor as possible. It is given, the more digital the device interface is, the more information beside the raw measurement value is transmitted into the measurement system.

Additional information could be, for example, timestamps regarding the exact sampling instant, error messages generated by probe controller systems and calibration/maintenance demand. A wide range of sensor connection possibilities, starting from analogue signal sampling up to complete remote control of sensing software using all-digital data transfer is implemented through integration of data protocols offered by the manufacturer. Analogue probe sensing and simultaneous digital data transmission of measurement channels and operational parameters leads to a higher number of logical channels than physically installed water quality probes, but shows higher data strength against influences like noise and value range exceeding.

In the analogue domain, data acquisition is a straightforward task using common analogue voltage- and current-I/O interfaces (German Standard DIN 66258 (1983)) for sensor connection and signal processing. The drawback of easy integration is given by the necessity of providing further information to interpret the signal. The scaling of the analogue input signal is usually fixed and manual configuration on-site should meet the range of the analogue/digital-converters as effectively as possible to reduce noise impact. Due to the missing option of adjusting the physical signal scaling on-line or automatically and corresponding to the actual signal range on probe-side, the use of analogue interfaces leads to unfavourable results for data quality. For example in the case of rain events, thereby the value ranges are extended multiple times compared to normal operation state and the analogue signal in this case is usually stuck at full-scale levels. Another problematic aspect is the need of additional digital/analogue- and analogue/digital-conversions; therefore additional signal noise is generated. This noise is undesirable especially in case of utilizing complex, up to date probe controllers, which already use digital data processing internally and therefore shall be incorporated into the network using existing digital interfaces.

In the digital domain, several pitfalls mentioned above are addressed through consistent digital signal propagation. Industrial standards for interfaces are applied and, in the majority of cases, the realization is not properly documented. The disadvantage becomes apparent in the lack of conformance in the practical realization of these standards; various dialects and modifications in the implementation of the standards

can be found, although compliance is assured at vendor side. In practice an increased effort for signal integration is the consequence.

In our monitoring solution different digital interfaces like RS232/485, Profibus, ModbusTCP and TCP/UDP-port based access of measurement signals were used and modular software structures for various sensor controller classes were implemented. Besides, the implementation of data exchange protocols using industrial standards, data transfer from the probe controllers is occasionally possible only by using the vendor-supplied software. In the best case communication with this software via programmable software interfaces is found; in the worst case the implementation of text file parsers for the collection of results in order to establish data communication between the measurement aggregator and the probe controller is the only way to achieve proper data communication. In some cases watchdog-mechanisms, which detect and completely re-initialise the vendor-provided software tool in case of software failures, were constructed. This can be described as “remote controlling of literally non-remote-controllable software” and led to time-consuming implementation tasks.

Time synchronization and operational timing control of measurements and cleaning cycles is a fundamental prerequisite for comparability of data within a distributed measurement network. The lack in treatment of the temporal dimension of measurement data is not least shown in missing system time-synchronisation of probe controllers, which is in fact still not provided in the majority of cases. This leads to the situation, that aspects regarding the time zone, daylight saving time and simply the temporal deviation of clocks are not observed. Besides that, information regarding the sampling time and the associated validity of the signal is lost due to quasi-continuous availability of measurement signals on the interface, often combined with the lack of an external measurement trigger. To cope with these pitfalls NTP-time-synchronisation all over the measurement network following the IETF RFC 5905 (2010) was implemented. During measurement operation every signal channel is sampled dependent of the specific delay-time (see International Standard ISO 15839 (2003)) and automatically aligned to general measurement time instants, ignoring the non-synchronized controller timers. The controller time base therefore is interchanged with the network-wide, validated data acquisition timescale and measurement data collected at various points in a catchment area becomes comparable, at all.

The implemented control of measurement-, cleaning- and maintenance-cycles has to insure correct measurement timings and therefore interference of active measurements is impossible through postponing of lower prioritized tasks a certain amount of time. For event based cycle control the measurement cycles as well are adjusted dynamically. To reach this requirement an algorithm based on adaptive scheduling (see Kopetz H. (2011)) was implemented. The adaptive character is given by the possibility of rearrangement of already scheduled tasks due to necessary interventions based on fault removal; e. g. a measurement-cycle is postponed because of a sample supply pump failure. In this case, an automatic pump sieve clean cycle is scheduled before the measurement cycle is caught up.

Measurement abstraction utilizing the proposed data channel structure, modular software architecture for data sampling in the analogue and digital domain and strict consideration of timing aspects enable the transition of data flow into the direction of the data processing layer called “Data to Information”.

## Data to Information

The main challenge in managing extensive data sets is the generation of information out of raw data, possibly utilizing automatic distinguishing of not relevant datasets, preferably in real-time and based on criteria related to local station environment. The logical layer “Data to Information” is responsible for data aggregation and evaluation through plausibility checks, combined with trigger condition generation and/or automatic alerting in case of value range excess. Plausibility checks consist of statistical assessments like allowed minimum/maximum signal range, maximum step respectively minimum signal change per time instant and coefficient of variance and take place directly at the monitoring station. Additionally, quality assessment procedures following dilution behaviours were examined. For detailed information regarding these checks see Winkler et al. (2012) and Winkler et al. (2013).

To get the full picture in regards of data completeness, beside water quality probes, several operating parameters like cabinet temperature, air pressure and pump flow have to be monitored. Together with the raw water quality measurement data, these parameters serve as input signals for event-detection and triggering of fault-correction mechanisms in case of operational errors. Even more, these signals act as the basis of reliable long-term operation. Fault detection, near real-time reaction times and reduced maintenance demand can be achieved through implementation of statistical on-site data plausibility checks called “CoarseChecks (CC)”, combined with automated user notification. Compared to post processing these structures enable early detection of faults during data collection and facilitate the differentiation between real system dynamics and device faults as shown in Winkler et al. (2013). All checks are done on-line, in near real-time and directly at the station. The plausibility assessment is stored together with the raw data in the same data channel objects as used in the “Data Acquisition”-layer. Again, utilizing this approach different data structures, typically vendor supplied, no longer occur, which leads to streamlined data processing throughout the system.

## Data Management and Presentation

In this layer the combination of several data sources within a measurement network and the appropriate presentation of data to the end user including export-functionality take place. Several requirements for proper operation of data synchronisation and long-term provisioning of data archives have to be met. The key requirements, as seen from user perspective, are sufficient data availability and short software reaction times combined with intuitive graphical user interfaces.

For data transfer and remote access, the measurement network is built using a virtual private network based on high-available, industrial 3G-routers installed at every station. Data are stored in local databases and a central data archive in a redundant manner. Asynchronous data replication, independent of measurement cycle time instants, is responsible for balancing of data amounts between the databases. The replication is on the one hand used for post-synchronisation in case of connection errors and on the other hand used for passing measurement data to the project partner’s databases (for mechanisms see Wolfson and Jajodia (1992)).

Short reaction times for data viewing and export utilizing easy to use software toolkits are expected from the end user nowadays. In the particular case of long-term trend examination, which means processing of a huge amount of data in a short period of

time, sufficient dimensioning of database storage machines and software tuning is required.

To overcome these requirements, several database management systems were evaluated and PostgreSQL, issued by The PostgreSQL Global Development Group (2014), was chosen. This open-source enterprise database system provides, among other features, good basic performance, several opportunities of particular tuning and receives good support through the database online-community.

For data inspection and export the tool *i<sup>TUW</sup>mon.Examine* was developed. This software allows inspection of long-term behaviour, based on the abstract channel concept described above. The work routine from end-user scope can be described as “select, examine and export data”.

For data selection the graphical user interface offers all common signal representation and manipulation capabilities using free-transform, multi-axes graphs and optional overlying measurement cursors for highlighting extreme values.

Powerful export-mechanisms allow the automatic combination of arbitrary channels using a two-dimensional table including automatic data alignment of all channel data into a common time-column for further processing. As data basis for table-based data export the user can select between the full dataset containing all channels between start- and end-date and the currently displayed dataset displayed as a graph, revealing a possibly interesting part of signal behaviour. In addition, data export as a picture using common data formats is achievable. Again, timing of measurements play a leading role and data sets across several monitoring stations become comparable in the aspect of signal sampling timestamps.

## **RESULTS: INTELLIGENT INFORMATION WATER MONITORING NETWORKS IN PRACTICE**

In this chapter data availability and user experience of data inspection using *i<sup>TUW</sup>mon.Examine* are discussed. The long term data availability of three stations, one in-situ and two more complex bypass-stations using this data processing tool is compared. A typical monitoring station is equipped with probes for measurement of ammonium-nitrogen, chloride, conductivity, dissolved organic carbon (DOC), nitrate-nitrogen, pH-value, potassium, spectral absorption coefficient (SAC), total organic carbon (TOC) derived from SAC, turbidity, diverse medium and device temperatures and further sensors for proper operation control like compressed air pressure, cabinet temperatures and air humidity as well as power supply measurement variables.

At an in-situ station the sensors are placed on a submersible platform directly into the measurement medium. The bypass stations contain a pump supplying the medium into a measurement tub, where the probes are mounted.

The data availability in 2012 versus station complexity is shown in Figure 4 on the left side. The availability calculation is based on the summation of single measurement instants with applied versus not applied CourseChecks. These checks offer an automatic quality assessment for collected data subsequent to every single measurement directly at the station. They are based on site-specific, statistical tests and provide a first indication of data quality achieved. The results of the checks are available for inspection and post processing in *i<sup>TUW</sup>mon.Examine*. In this specific case, data availability using bypass-stations is notably higher through the lower amount of rejected data sets.

©IWA Publishing [2014]. The definitive peer-reviewed and edited version of this article is published in Water Science & Technology, Volume 70, Issue 12, 1926-1933, 2014, <https://doi.org/10.2166/wst.2014.415> and is available at [www.iwapublishing.com](http://www.iwapublishing.com). This is the accepted version.

Applying the automated CoarseChecks around 15 % of single measurements at the bypass stations compared to 25 % of data at the in-situ station are discarded. The difference in data rejection rate can be explained by:

- Ensured sample stream through flow rate monitoring and pipe trace heating at bypass stations, especially in winter time (frozen surface water),
- environmental-sealed housing and enclosure of measurement apparatuses in measurement containers,
- better intervention options in case of failure (e. g. remote pump clean cycle triggering).

Consequently, it can be said that despite that the higher effort and mechanical complexity of a bypass-system implementation is worthwhile in respect of data completeness. In 2012 on average 60,000 (passed data) respectively 71,500 (raw data) single measurements per channel were collected. In the year 2012, utilizing about 35 data channels per station, a total of 6,200,000 (passed data) respectively 7,600,000 (raw data) data sets in the described abstract channel structure are stored in the databases in a redundant way. These data sets represent the aggregation of all channels of all three measurement stations in the catchment area; typically around 2,000,000 checked datasets are gathered per single station. Raw measurement data are residing in the database and so post-mortem analysis and retrospective refining of statements through the dynamic adaption of plausibility assessments is possible.

Results show, that the availability of stations with increased station complexity through medium bypass situation and automated pump control is at least equivalent, respectively may be higher, than using an in-situ station with reduced complexity. This is applicable especially for “station 2” operating  $i^{TUW}_{mon}$  since 2010. Extensive experiences in long term water quality monitoring have been acquired.

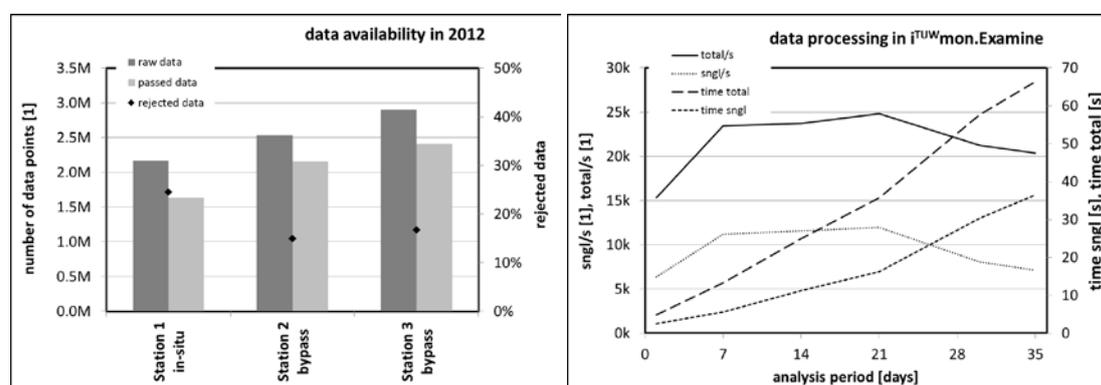


Figure 4 data availability (left) and data processing utilizing  $i^{TUW}_{mon.Examine}$  (right)

The graph on the right hand side of Figure 4 shows the data processing benchmark using  $i^{TUW}_{mon.Examine}$  as analysis tool. The objective of this benchmark is to give an impression of what the end user could expect from data processing within a reasonable amount of time. Depending on the user chosen analysis period an increasing number of data points have to be processed for graphical display and export. The processing task includes loading the complete dataset of the selected time range, filtering regarding the plausibility assessments and, in case of exporting tasks, preparation of export file structure and initialization of data selection, if only a subset of channels should be exported. The automatic alignment of data channel objects into a common timestamp-

column is designed as an on-line task, which takes place inside the database utilizing aggregation functions triggered by data insertion during measurement operation. Therefore, combining of time series is done directly after signal sampling at the station and necessary processing time is multi-partitioned over the complete operation time domain resulting in immediate data availability for the end-user during an export task.

Conventional data are referred as “sngl” and conventional combined with spectral data, stored as vector objects, are referred as “total”. As can be seen in the figure the number of data points increases linear in the first seven analysed days and then remains constant. For analysis periods more than 21 days the processing capacity decreases. As additional information on the right axis the cumulative processing time is shown, which more or less increases linear up to an analysis period of 21 days. For longer analysis periods the processing time increases stronger due to database performance degrading.

The average data loading procedure on a machine equipped with an Intel Core 2 Duo E8400 processor for a three weeks data record including spectra lasts about 38 seconds and results in over 870,000 single measurements, which are processed and formatted for automation-assisted viewing. That means, around 23,000 single measurements are processed per second, whereby spectral values of UV/VIS-spectrometers are processed around three times faster compared to conventional data. Spectral data is processed column-wise and database performance is notably higher for vectorial data sets. Results show, that 21 days as basis for trend analysis offer a good trade-off between processing time and data extent. Similar results are obtained during export of merged data. In this case the memory intensive data display task is bypassed and the processed data are written directly into the export data file, resulting in even higher throughput. Apart from manual adaption of database configuration to achieve higher database throughput and query performance the implementation of solid state-drives as storage basis on the database server was a great leap ahead. These efforts resulted in a performance-factor of approximately 40 for the data processing task compared to the use of standard hard-disks and therefore extensive data records get reasonable tradable.

Using the proposed monitoring platform, actual measurements are available on every data processing workstation within the institute approximately one minute after the on-site sampling takes place and the water quality picture based on plausibility checked datasets is available more or less “live”. Utilizing our proposed structure allows to operate central managed, but physically distributed data acquisition stations in a comfortable and resource efficient way, meanwhile over years. Automatic alarming of faulty system states (e. g. pump failures) and automatic data plausibility assessments (e. g. sensor drifting estimation) offers technicians the ability to plan maintenance cycles and demands.

The specified software architecture allows the use of several software modules as a standalone system; an actual project is based on a small amount of probe controllers on-site and a separated processing workstation located in-house. In this case, the software interface for the end user is available using the virtual private network, in the same transparent manner as before, although the measurement workstation is locally separated from the measurement area.

The novelty character of the described approach can be seen in the sum of separate technical aspects addressed during the implementation since 2010. The developed tool includes comprehensive interfaces to a nearly complete range of market available water quality probes and controllers. The so collected dense, statistically approved and plausible data are available over a long period of time and ready for immediate, “on-line” use.

## REFERENCES

- BMFLUW, Austrian Federal Ministry of Agriculture, Forestry, Environment and Water Management (2006): “*Gewässerzustandsüberwachungsverordnung BGBl. Nr. 479/2006 i.d.g.F.*”; Sektion VII/Abteilung 1, Nationale Wasserwirtschaft
- DIN 66258 (1983): “*DIN 66258: Schnittstellen und Steuerungsverfahren für die Datenübermittlung*“; Deutsches Institut für Normung e.V.
- Elliott C., Vijayakumar V., Zink W. and Hansen R. (2007): “*National Instruments LabVIEW: A Programming Environment for Laboratory Automation and Measurement*”; Journal of Laboratory Automation, p17-24
- Ellis K., Mounce S.R., Ryan B., Templeton M.R. and Biggs C.A. (2013): “*Use of on-line water quality monitoring data to predict bacteriological failures*”; Procedia Engineering 70 (2014) p612-621
- IETF RFC 5905 (2010): “*Network Time Protocol Version 4: Protocol and Algorithms Specification*”; June 2010; 2014-06-06. <http://datatracker.ietf.org/doc/rfc5905/>
- ISO 15839 (2003): “*Water quality – On-line sensors/analyzing equipment for water – Specifications and performance tests*”; ISO International Organization for Standardization
- Kopetz H. (2011): “*Real Time Systems, Design Principles for Distributed Embedded Applications, Second Edition*“; Springer Science + Business Media, LLC 2011; ISBN 978-1441982360
- Piñeiro Di Blasi J.I., Martínez Torres J., García Nietoc P.J., Alonso Fernández J.R., Díaz Muñiz C. and Taboada J. (2013): “*Analysis and detection of outliers in water quality parameters from different automated monitoring stations in the Miño river basin (NW Spain)*”; 2013, Ecobiological Engineering 60 (2013) p60-66
- Strobl R. O. and Robillard P. D. (2007): “*Network design for water quality monitoring of surface freshwaters: A review*”; Journal of Environmental Management 87 (2008), p639-648
- The PostgreSQL Global Development Group (2014), Internet: “*PostgreSQL – The world’s most advanced open source database*”; 2014-06-06, <http://www.postgresql.org>
- Winkler S., Fuiko R. and Winkelbauer A. (2012): “*iTUWmon – A monitoring network platform for automated data plausibility assessment and data integration*”; Conference Proceeding: IWA New Developments in IT & Water, Amsterdam
- Winkler S., Kornfeind L., Fuiko R. and Winkelbauer A. (2013): “*Separating real system dynamics from measurement failures and artifacts by means of automated data quality assessment methods*”; Conference Proceeding: ica2013, 11<sup>th</sup> IWA CONFERENCE ON INSTRUMENTATION CONTROL and AUTOMATION, Narbonne
- Wolfson O., Jajodia S. (1992): “*Distributed algorithms for dynamic replication of data*”; ACM Proceeding: PODS '92 Proceedings of the eleventh ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems, p. 149 163; ISBN 0-89791-519-4