Integrated assessment of receiving water quality based on validated high resolution data and water quality modelling

Évaluation intégrée de la qualité d'un milieu récepteur basée sur des données en continue validées et la modélisation de la qualité

Valentin Gamerith[1], Günther Leonhardt[2], Thomas Hofer[1], Carolina Kinzel[2], Manfred Kleidorfer[2], Wolfgang Rauch[2]

[1]: Graz University of Technology, Institute of Urban Water Management and Landscape Water Engineering, Stremayrgasse 10/I, A-8010 Graz, gamerith@sww.tugraz.at, hofer@sww.tugraz.at
[2]: University of Innsbruck, Unit of Environmental Engineering, Technikerstrasse 13, A-6020 Innsbruck, Austria
guenther.leonhardt@uibk.ac.at, carolina.engelhard@uibk.ac.at, manfred.kleidorfer@uibk.ac.at, wolfgang.rauch@uibk.ac.at

RÉSUMÉ

L'évaluation de la qualité des milieux récepteurs demande une approche intégrant les systèmes urbains et les eaux de la surface, par exemple par la modélisation intégrée. L'effort d'établir ces modèles et la disponibilité des données sont des défis connus. Le projet *IMW3* est dédié à la collection de données dans le bassin versant de Baden (Autriche). Quatre stations de mesure en-ligne sont installées dans la rivière et à l'effluent de la station d'épuration locale. Les données en haute-résolution sont la base pour développer et appliquer des méthodes d'analyse et validation des données et pour établir des modèles conceptionnels. Cette contribution présente i) des méthodes de validation des données et la gestion des méta-informations et ii) des analyses des données pour définir un modèle conceptionnel. On montre qu'il est difficile d'obtenir des séries de données stables même dans un projet dédié et on démontre la nécessite d'une propre gestion des méta-informations. Des méthodes statistiques sont utilisées pour identifier et vérifier les processus importants, mais l'applicabilité dépend fortement de la qualité des données.

ABSTRACT

River basin wide planning for assessing the ecological and chemical status of surface water bodies calls for integrated modelling of the river and the connected urban water systems. Challenges were identified in the effort to set up such models and in general data availability for such studies. The Austrian interuniversity project *IMW3* is dedicated to data acquisition in a pilot catchment in Baden, Austria. Four online measurement stations are installed in a river section and the effluent of an urban drainage system. This contribution focuses on two parts of this project: i) the set-up of a framework for data validation and post processing with focus on implementation of meta information and ii) the transformation of data to information by data analysis and a model implementation. We show that while a lot of effort was put into data acquisition, obtaining reliable data over a longer period is still challenging. We discuss the necessity of a proper meta-information management and show that even with in-depth analysis manual data treatment cannot be completely avoided. Statistical methods are used to identify and verify processes in the receiving water prior to setting up a conceptual model. However, their applicability strongly depends on data quality.

KEYWORDS

Data validation, Integrated approach, Meta-data management, Modelling, On-line monitoring

1 INTRODUCTION

Since the implementation of the EU Water Framework Directive (European-Community, 2000) river basin wide analysis, planning and management has become a necessity to assess the required good ecological and chemical status of surface water bodies. This river basin wide approach has major implications for urban water systems (i.e. drainage systems and waste water treatment plants) due to their influence on receiving water quality. In this complex system the optimum performance of single parts of the system does not necessarily lead to the best overall performance. Therefore an integrated approach accounting for various sources of pollution and impacts on receiving water bodies is required (Rauch and Harremoes, 1997).

For this integrated assessment the use of integrated models has been in discussion in the scientific community since the first INTERURBA conference (Lijklema *et al.*, 1993). Several studies covered this topic in the following decades and were summarized by Rauch *et al.* (2005). As reviewed in Blumensaat *et al.* (2012), today integrated modelling as tool for water quality assessment is also anchored in a number of European guideline documents.

In 2008, recommendations on integrated modelling were issued by the HSGSim group (Muschalla *et al.*, 2009). Therein the complexity of the overall system and a lack of data, which limits the practical application of these models are identified as main bottlenecks in integrated modelling approaches. Schütze and Alex (2004) state that the balance of the processes to be considered, the adequate modelling approaches, and the available data is usually limited by the resources (time and money) available for the study.

Besides data availability, the need of validating the available data is widely acknowledged. Recent publications on this topic in the field of urban drainage focus on the emerging high resolution data, e.g. Mourad and Bertrand-Krajewski (2002) or Winkler *et al.* (2008). Branisavljevic *et al.* (2010) state that un-validated data can result in misleading conclusions and erroneous decisions. Kleidorfer (2009) stresses that data errors have to be identified and removed prior to modelling.

Integrated models are built up of several sub-models. In an urban drainage context those sub-models typically represent sub-systems as the sewer system, the treatment plant and the receiving water. In general, the whole range of modelling approaches (data driven, physically based conceptual, hydrodynamic) is applied to model these sub-systems. However, for integrated modelling, mainly physically based conceptual models are used due to data availability (as mentioned above) and time restrictions in the execution of the software.

For easier handling, software tools to model all sub-systems have been developed, as e.g. *CityDrain* (Achleitner *et al.*, 2007; Burger *et al.*, 2010) or SIMBA (ifak system GmbH). Freni *et al.* (2011) recently applied methods of formal uncertainty analysis to integrated models.

In scope of the Austrian interuniversity Project *IMW3 - Novel measurement techniques in water management* hydraulics and pollutants are monitored with high resolution on-line measurements at four locations in a pilot catchment in the area of Baden, Austria. Besides the design and development of the measurement stations and a database a major effort in this project is put on data validation, data processing and model development for integrated water quality assessment.

This paper presents the framework for data validation and data post processing developed in this project. The methodology and process of data treatment is described and means of metadata storage are presented. In a second part, the procedure to identify important processes from the data prior to setting up an integrated model is discussed. Finally, we briefly describe the modelling tools and concepts.

2 MATERIALS AND METHODS

2.1 Pilot catchment and measurement stations

The pilot catchment of the IMW3 project is situated in the area of Baden, Austria. A river section of the river Schwechat with urban influence is monitored by four measurement stations. These stations are located upstream (IN) and downstream (OUT) of the investigated river section, at the effluent of the local waste water treatment plant (WWTP) of the city of Baden and in a tributary creek to the Schwechat River (the Hoermbach).

At each measurement station water level, temperature, pH, conductivity, chloride, nitrogen (NH_4-N and NO_3-N), turbidity and UV-Vis absorption spectra are measured at time intervals of one to 7.5 minutes. All stations have been continuously measuring for a period of 1½ years. Maintenance and probe calibration is carried out in regular intervals of three weeks (or in the case of a severe failure of a device). Additionally, meteorological data as rainfall, air temperature and duration of sunshine are available in time steps of 10 minutes from official weather stations in the area, operated by the Austrian Central Institute for Meteorology and Geodynamics (ZAMG).

For a detailed description of the pilot catchment, the measurement stations and all measured variables please refer to BMLFUW (in preparation).

2.2 Data validation and post processing framework

As discussed in Bertrand-Krajewski *et al.* (2000) proper analysis and validation of the measured raw data is necessary to exploit the information from the available data, separating "good" from "bad" measurements based on all available knowledge. Especially on-line measurements with high temporal resolution lead to a huge amount of data over time. This limits the applicability of manual or visual data validation and calls for (semi-) automatic validation routines. Nevertheless visual analysis can be useful for a first assessment of the data i.e. to identify obvious errors as gaps, long term drifts etc., and to understand the general system behaviour (Gamerith, 2011).

In this contribution we use the terms *data validation* and *post-processing. Data validation* refers to the process of checking data on errors (i.e. gaps, outliers, drifts etc.) either visually, with automated tests or based on meta-information. *Post-processing* refers to handling the erroneous data and preparing data for a specific purpose (e.g. substituting erroneous data, generating equidistant time series).

In the IMW3-project a step-by-step approach for validating the measured data is followed: First a visual analysis of the available data is carried out to detect obvious errors and to pre-define limits for automated validation tests. Data is first evaluated directly at the measurement station at the time of measurement. This information can be used i.e. for alerting purposes. In-depth validation of the data is then carried out for the data stored in the database within a framework developed in the R software (R Development Core Team, 2011). Based on the results from this validation, data is post processed for further use e.g. in modelling studies. The requirements for the data-validation and post-processing framework were:

- Possibility to use different data sources and to implement (additional) validation tests and post-processing routines.
- Implement additional sources of meta-information as results from visual analysis, log-books etc.
- Log all metadata resulting from validation and post processing to keep track of modifications in the data.
- Basic visualisation of the results.

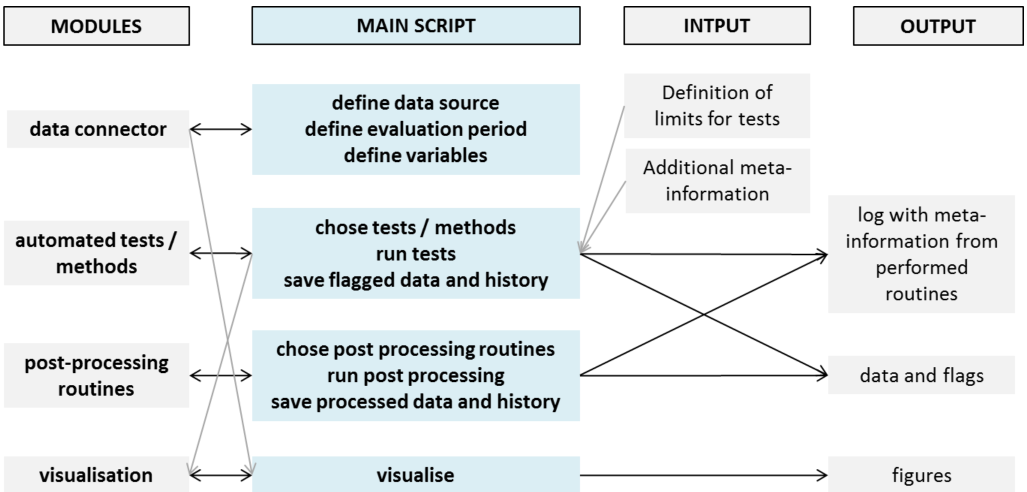

Figure 1: Framework for data validation and post processing – modular design in R

Figure 1 shows the workflow for data validation and post processing within the developed framework. The framework comprises of one main script that links modules for data acquisition and storage, validation tests, post processing and visualisation. These modules consist of sets of generic functions. Data is acquired from any supported format (currently connectors for Postgres, netCDF and csv are available) via the data connector module and is internally stored in an R workspace. Parameterisation of the tests can be set via a configuration file that is easily modifiable. Additional meta-information can be imported from external sources. The main script defines which validation tests and post processing routines are carried out for each variable. Data is flagged according to the test results and the processed data is stored together with the flags. All information on applied tests and post processing is saved in a log file. This allows tracing all modifications from the raw data. The raw data itself is always kept unchanged. The whole process can be automated once all settings are defined. All modules are available as R code on demand.

2.2.1 Classification and methods

For data validation we followed the classification proposed by Mourad and Bertrand-Krajewski (2002) and implemented e.g. by Branisavlijevic et al. (2010). Data is classified (flagged) as A for reliable values, B for doubtful values and C for faulty, outlying or aberrant values based on several automated tests. The values flagged B are then analysed manually and a decision is made based on expert knowledge. Due to the heterogeneity of the data within the IMW3 project (different time intervals, different scales in the measured variables etc.) limits for the automated routines have to be adapted for each measurement station and variable. It is possible to define different system states (e.g. dry weather flow and wet-weather flow) based on any variable value. Different settings for the validation tests can be defined for different system states.

For the presented results, several validation tests as proposed in DWA (2011) and Mourad and Bertrand-Krajewski (2002) were carried out: i) minimum-maximum test ii) test on consistency and/or flat line iv) signal's gradient v) deviation from moving average. In addition automated gap detection was carried out and external meta-information was implemented. The post processing generally substitutes C flagged values by NaN (not a number) in a first step. Next an interpolation of missing values or – for noisy data – a substitution of the values with a moving average is carried out. From this data, equidistant time series with an interval of 5 minutes are generated for modelling.

2.2.2 Meta information management

Meta information can arise from a multitude of different sources. In context of data validation, interest lies in information that i) can influence data quality and ii) is not registered directly with the measurement data. Some examples could be: maintenance, system failures, seasonal variations, snow melt etc. Also simple information can help in data interpretation, e.g. the indication of (heavy) storm events. To organise all the available meta-information, a digital calendar was set up for each measurement station. Information from maintenance log books, visual data analysis and other sources is centrally stored in this calendar. Data is classified via colour coding (green: informative content only, A flagged; yellow: possible error in data, B flagged; red: data erroneous, C flagged). The calendar can be accessed via the standardised CalDAV protocol. We used Google calendar. However, any calendar based on the CalDAV protocol can be used. These calendars are easy to use, can be accessed by mobile devices, shared among different users, and force the user to define start- and end date of the meta-information. All data can be exported via the CalDAV protocol and can be linked to the validation routines for data flagging.

2.3 Data analysis and modelling

As described above, the focus of the monitoring in the current research project is on the receiving water. Due to limited resources, not all fluxes from urban drainage systems could be monitored (i.e. combined sewer overflows) but only the effluent of the WWTP. The main focus of the modelling is thus also on receiving water quality.

Prior to model setup, we performed an analysis of the data to identify physical and biochemical processes in the river stretch. For this analysis, visual inspection, mass balances and methods of descriptive and multivariate statistics are applied. The aim is to ensure that the model is as simple as possible and simulated processes can be validated from the data. Whereas methods as multiple regression analysis are also used to derive statistical models, we are interested in identifying significant processes. Those processes were then implemented in the conceptual model.

2.3.1 Data analysis and process identification

Data analysis should comprise of two steps, first the Exploratory Data Analysis and second the Confirmatory Data Analysis (Hipel and McLeod, 1994). Exploratory Data Analysis starts by common visual inspection of the available data. Very obvious processes can already be identified by such simple means. For less dominant or more complex phenomena, statistical methods should be used. For this statistical analysis, it is crucial to use validated data, as otherwise results could be disturbed by data errors. Additional post processing and the availability of metadata are also extremely helpful, as some methods for statistical analysis require evenly spaced data.

Different sub-sets of parameters for various periods were selected and prepared for further analysis. Selection was performed with respect to meteorological characteristics (e.g. different seasons or dry weather vs. rainfall events) and the suspected processes.

The sub-sets of flow data from different periods were first evaluated with respect to flow travel time and mass balance. As flow data is derived from various other measurements, uncertainty was first estimated according to the law of propagation of uncertainties (JCGM, 2008) and considered in the analysis.

One method we applied is stepwise linear regression. This algorithm aims to select the most significant explanatory variables among a number of possible ones in a multilinear regression model. We used stepwise regression with two aims: The first one is to check if well-known processes and interdependencies between parameters are reflected by the data. The second aim is to identify further interdependencies, i.e. if a certain parameter is significantly related to one or several others in a larger predefined set. If the model proposed by the algorithm is not physically or chemically meaningful or statistically weak, the related physical or biochemical process is if course not included in the conceptual model.

As the data originates from a non-steady system, we used a method described by Dürrenmatt and Gujer (2012) to include dynamic effects in the regression analysis. The original dataset is enlarged by multiple copies shifted by different lags. The enlarged dataset is then analysed by stepwise linear regression. A further extension can be achieved by deriving new quantities from the original data, as e.g. cumulated rainfall during an event or cumulated sunshine during the day. Similar approaches are applied in stormwater quality modelling (Sun and Bertrand-Krajewski, 2012). By including this lagged and cumulated data, it is possible to identify dynamics and “memory” effects (e.g. the effect of very sunny days in relation to partly overcast conditions).

2.3.2 Conceptual integrated model

The integrated model was set up in the software CityDrain 3 (Burger *et al.*, 2010), which is a follower of CITYDRAIN (Achleitner *et al.*, 2007). It allows building a model using pre-defined small scale conceptual sub-models as e.g. subcatchments, sewer reaches or combined sewer overflow (CSO) structures. For this project, a sub-model for the simulation of river reaches with respect to both flow and water quality was developed.

The newly developed river-module combines the conceptual Muskingum routing scheme (Chow *et al.*, 2001) with a generic but flexible water quality simulation function. It is capable to simulate transport and conversion of several substances based on the concepts of the River Water Quality Model No. 1 (Reichert *et al.*, 2001; Shanahan *et al.*, 2001). The river stretch is represented as a cascade of reactors, which are assumed to be completely mixed during each time step. To allow for a maximum of flexibility, the choice of substances as well as the conversion equations are not pre-defined, but have to be selected by the user. Interaction with groundwater can be represented by lateral in- or outflows. The river module further allows for build-up and degradation of sediments. For details concerning the other sub-models used in CityDrain 3 see Achleitner *et al.* (2007).

As no in-sewer measurements are available, possible emissions from storm sewers and CSO structures could only be roughly estimated based on water quality data from the receiving water. The same applies for groundwater infiltration or exfiltration.

The setup of the hydrological model was followed by the selection of water quality parameters and conversion processes.

3 RESULTS AND DISCUSSION

3.1 Data validation and processing

Overall it proved challenging to provide complete data sets from all four measurement stations and all sensors over longer periods. Stability issues of the sensors themselves, clogging, drifts over time as well as non-functional equipment led to incomplete data sets. Figure 2 shows a screenshot of the online-calendar used to compile all meta-data from various sources as maintenance logs, log books and visual data analysis for June 2012. The colour coding allows identifying periods of rainfall events (green), maintenance (blue), warnings (yellow) and errors (red). It can be seen that for this example no periods without remarks exists.

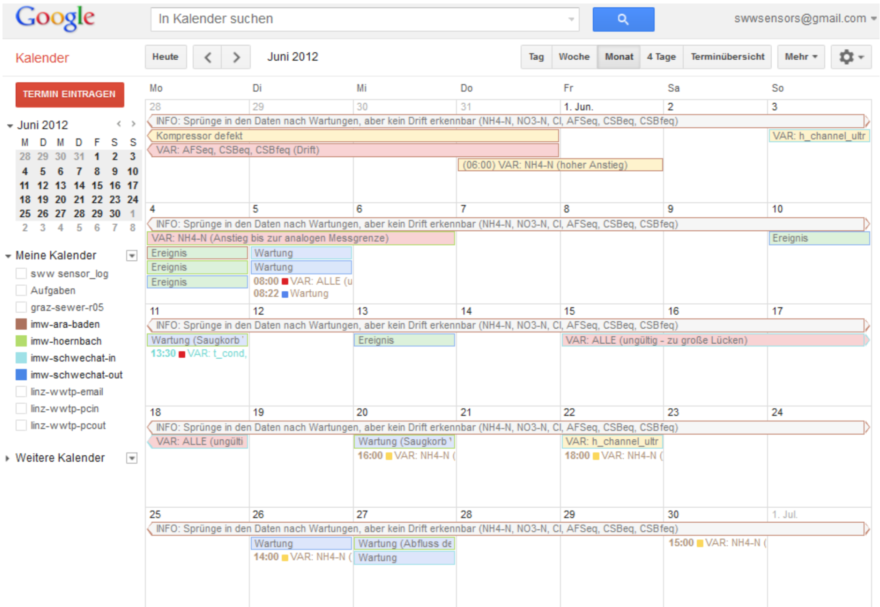

Figure 2: Overview of online-calendar with meta-data for June 2012. Color coding correspond to : rainfall events (green), maintenance (blue), errors (red) and warnings (yellow)

Figure 3 shows results from a validation and post processing run for the chemical oxygen demand (COD) concentrations of the WWTP effluent for the period of January to October 2012. For the modelling part, stable periods where most of the sensors work satisfactorily were chosen.

This example allows deriving some general remarks and identifying challenges in automated data validation:

- Outliers are in general the most easily identifiable errors, either by using a minimum-maximum or a gradient test.
- The importance of additional meta-information is shown e.g. for a period in June where the compressor for cleaning the sensor was damaged due to a power failure – this period is excluded from the validated data by connection to the online-calendar information.
- One main difficulty proved to be the parameterisation of the validation routines. For example in the period of March to May data is unstable. Setting too strict limits would result in more ‘good’ data being identified as erroneous. So either a “compromise solution” has to be chosen or the period in question needs to be evaluated separately.
- Some errors remaining after applying the automated routines can only be removed or flagged manually.

3.2 Data analysis and modelling

Very obvious processes were already identified by visual analysis of plots of different short-term periods. Typical diurnal dynamics of dissolved oxygen, pH and temperature can be observed. The correlation of these parameters was also confirmed by applying stepwise regression. Furthermore, seasonal variations can be easily detected by box-and-whisker plots. Therewith the seasonal variations of the daily mean concentrations of the examples chloride, temperature, pH and dissolved oxygen were demonstrated.

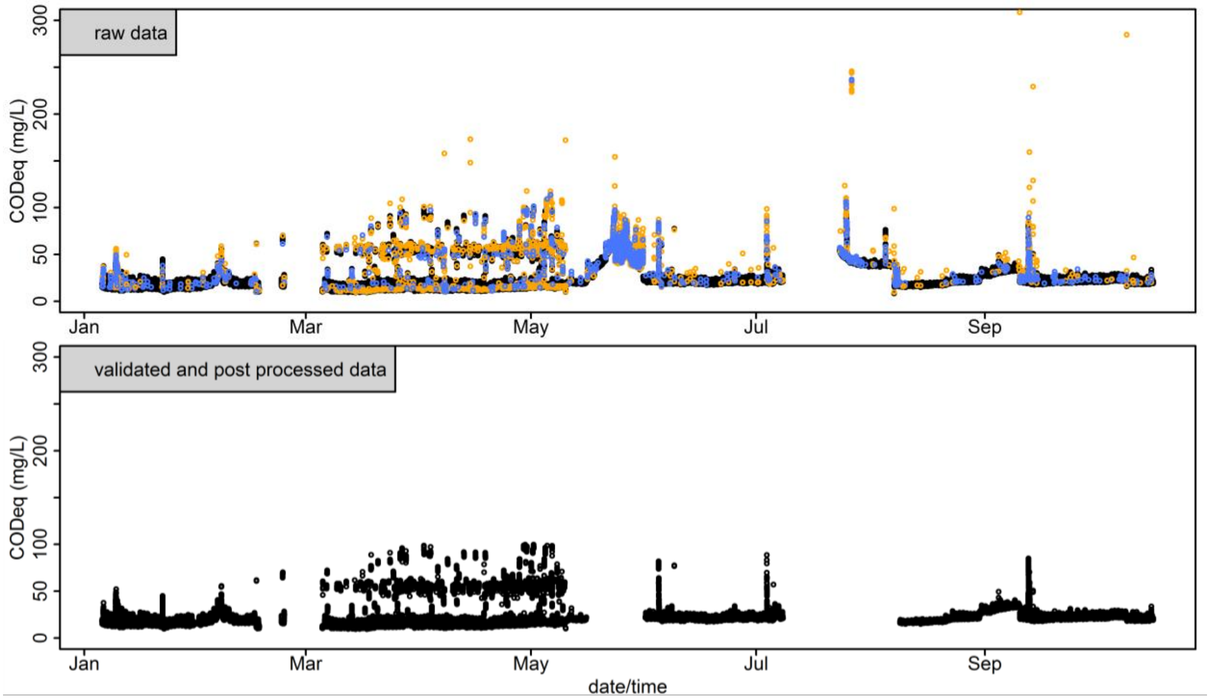

Figure 3: Example of data validation and post-processing results for COD concentrations in the WWTP effluent for January to October 2012: raw data in the upper part with validation flags (black: A-flagged, blue: B-flagged and orange: C flagged) and post-processed data in the lower part.

Reliable flow data is crucial, not only to model the hydrologic behaviour. Flow data is also necessary for establishing mass balances which are needed for understanding the processes in the river. After the validation of the flow data, mass balances of the conservative substance chloride were computed. Both, the hydrological mass balance and the mass balance of chloride show that only a part of the sources are accounted for with the measurement stations. Large parts (20 – 60% for water and around 50% for chloride) remain unknown, possible processes could be groundwater exfiltration, surface runoff or unobserved discharges from combined sewer overflows and separate sewer systems. Figure 4 shows flow data and according uncertainties as well as the mass balance deficit (sum of unknown inflows divided by the sum of known outflow) derived from the hydrologic mass balance for a rainfall event. The distribution of the relative deficit was derived by Monte Carlo Simulation from 100 000 flow data samples.

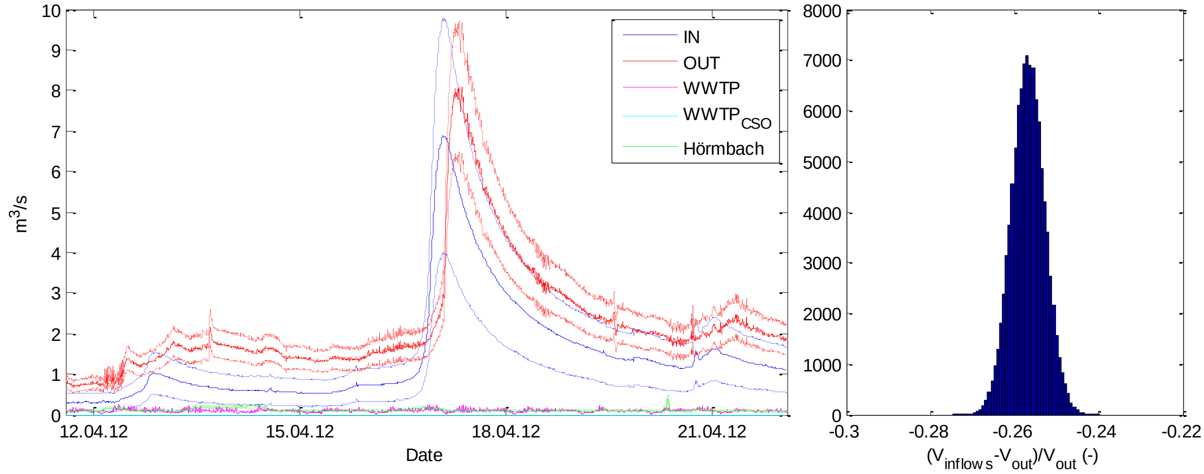

Figure 4: Evaluation of flow data for a rainfall event; left: observed flow data and corresponding uncertainty (2.5- and 97.5-percentiles); right: distribution of mass balance deficits calculated from the flow data samples.

To identify the influence of the different inflow sources (upstream catchment IN, WWTP and the Hörmbach) on the conditions at the downstream boundary of the investigated river stretch, water quality parameters were analysed for different meteorological conditions. Stepwise regression was used to identify the drivers for dynamics of conductivity at the end of the investigated river stretch. The analysis was performed separately for different weather conditions – a dry period and a rainfall event. Results are shown in Figure 5. For dry periods, variations at the end of the investigated river stretch (OUT) are related to the dynamics at all known inflows, i.e. both the WWTP effluent and the Hörmbach seem to have significant impact. On the other hand, for a rainfall event, the conductivity at the end of the stretch is mainly related to the inflow from the upstream catchment (IN). No significant influence of surface wash-off in urban areas could be observed for this event. However, despite one and a half years of monitoring, complete datasets are only available for a few events. This reflects both, the relatively dry observation period and the fact that certain problems at monitoring stations are likely to occur during harsh conditions.

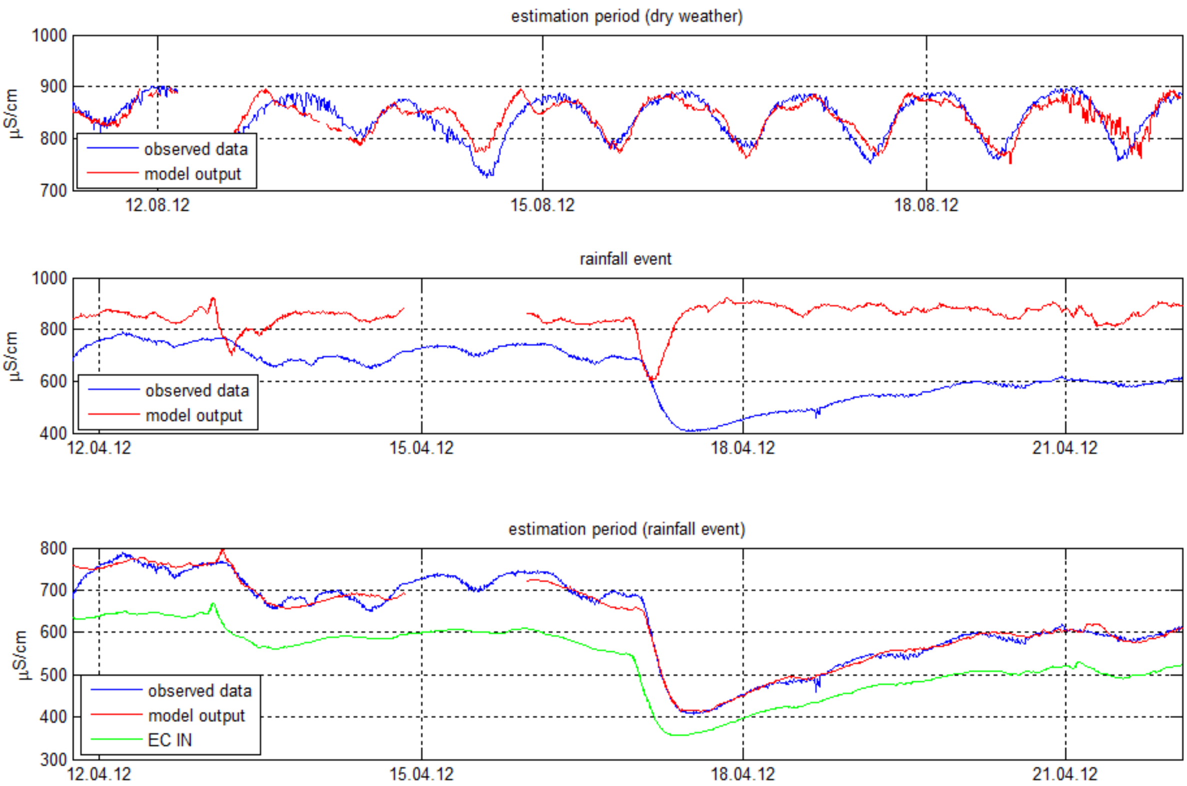

Figure 5: Multivariate linear regression model for electric conductivity at OUT: model estimated for dry weather period (top), model applied to rainfall event (middle) and model estimated for rainfall event mainly based on conductivity at IN (bottom).

When including shifted copies of the dataset into stepwise regression (as proposed by Dürrenmatt and Gujer, 2012), small lags can result in unrealistic models. The lag times should therefore be chosen with care, e.g. related to flow travel times, which are in the order of two to five hours.

Based on the data analysis, the following processes were selected to be represented in the conversion model to simulate water quality in the investigated river stretch: 1) growth of sessile algae, 2) respiration of sessile algae, 3) nitrification, 4) growth of heterotrophs, 5) lysis of nitrifiers and 6) lysis of heterotrophs. Exchange of heat energy with the atmosphere is also modelled.

Despite the fact that stormwater and combined sewage (from CSO structures) is discharged into the river stretch, no severe effects regarding physical water quality parameters (oxygen deficit, NH4-N peaks) were observed during large rainfall events at the downstream measurement station. The modelling was therefore focused on dry weather periods.

4 CONCLUSIONS

- Within the IMW3 project a lot of effort was put into data acquisition. Nonetheless, it was shown that the high amount of data, different sensors and sensor locations led to problems to obtain stable and reliable data over a longer period. This lets us conclude that even with a dedicated project and non-negligible personal and monetary efforts a reliable data acquisition for integrated modelling is still challenging and the need to validate the measured data is incontestable.

- Overall a visual analysis to identify the system behaviour and obvious data errors as well as the input of meta-information (e.g. maintenance information and log-books) proved of highest importance for the validation task. For this the possibility of coupling an open calendar protocol to automated validation routines proved to be a flexible and promising solution.
- While the automated tests allowed to gain significantly better data quality, application in detecting erroneous data in practical conditions is limited as no generic parameterisation over all sensors and measurement stations is possible e.g. due to different system states, different ranges at different measurement locations etc. Even with in-depth analysis data remains that have to be treated manually.
- Even after a sound data validation, the choice of applicable methods for statistical analysis remains limited, as only some methods can deal with data gaps. However, results have to be examined critically. A longer dataset would be desirable to verify identified phenomena.
- Only a limited number of processes affecting water quality could be clearly identified from the data. Thus, even with great effort made for measurements, only limited insight into the condition of the receiving water can be provided.

ACKNOLEDGEMENTS

The IMW research project was funded by the Federal Ministry of Agriculture, Forestry, Environment and Water Management (BMLFUW) together with water authorities of Lower Austria and Styria.

Hydrometeorological data was provided by the Austrian Central Institute for Meteorology and Geodynamics (ZAMG) and Hydrographischer Dienst Niederösterreich.

Special thanks to the laboratory and IT staff of all project partners for their continuous effort in maintenance of the measurement stations.

LIST OF REFERENCES

Achleitner S., Möderl M. and Rauch W. (2007). CITY DRAIN © - An open source approach for simulation of integrated urban drainage systems. *Environmental Modelling & Software*, **22** (8), 1184-1195.

Bertrand-Krajewski J.-L., Laplace D., Joannis C. and Chebbo G. (2000). *Mesures en hydrologie urbaine et assainissement*. Editions Tec&Doc, Paris, France.

Blumensaat F., Staufer P., Heusch S., Reussner F., Schutze M., Seiffert S., Gruber G., Zawilski M., Rieckermann J. and Grp H. (2012). Water quality-based assessment of urban drainage impacts in Europe - where do we stand today? *Water Science and Technology*, **66** (2), 304-313.

BMLFUW (in preparation). *IMW3 - Innovative Messtechnik in der Wasserwirtschaft. Integrierte Betrachtung eines Gewässer-(abschnitt)s auf Basis kontinuierlicher und validierter Langzeitmessreihen - Endbericht*. Vienna, Austria.

Branisavljevic N., Prodanovic D. and Pavlovic D. (2010). Automatic, semi-automatic and manual validation of urban drainage data. *Water Science and Technology*, **62** (5), 1013-1021.

Burger G., Fach S., Kinzel H. and Rauch W. (2010). Parallel computing in conceptual sewer simulations. *Water Science and Technology*, **61** (2), 283-291.

Chow V. T., Maidment D. R. and Mays L. W. (2001). *Applied hydrology*. McGraw-Hill series in water resources and environmental engineering, New York [et al.]; McGraw-Hill,

Dürrenmatt D. J. and Gujer W. (2012). Data-driven modeling approaches to support wastewater treatment plant operation. *Environmental Modelling & Software*, **30** (0), 47-56.

European-Community (2000). Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Official Journal of the European Communities*, **43** (L327), 1- 71.

Freni G., Mannina G. and Viviani G. (2011). Assessment of data and parameter uncertainties in integrated water-quality model. *Water Science and Technology*, **63** (9), 1913-1921.

Gamerith V. (2011). *High resolution online data in sewer water quality modelling.* PhD thesis, Institute of Urban Water Management and Landscape Water Engineering, Graz University of Technology.

Hipel K. W. and McLeod A. I. (1994). *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, ISBN 0 444 89270-2.

JCGM (2008). *Evaluation of measurement data – Guide to the expression of uncertainty in measurement (JCGM 100:2008, GUM 1995 with minor corrections)*. Working Group 1 of the Joint Committee on Guides in

Metrology.

Kleidorfer M. (2009). *Uncertain calibration of urban drainage models. A scientific approach to solve practical problems.* PhD thesis, Unit of Environmental Engineering, University of Innsbruck.

Lijklema L., Tyson J. M. and Le Souef A. (1993). Interurba - Interactions between sewers, treatment plants and receiving waters in urban areas. *Water Science and Technology*, **27** (12), 1-244.

Mourad M. and Bertrand-Krajewski J.-L. (2002). A method for automatic validation of long time series of data in urban hydrology. *Water Science and Technology*, **45** (4-5), 263-270.

Muschalla D., Schutze M., Schroeder K., Bach M., Blumensaat F., Gruber G., Klepiszewski K., Pabst M., Pressl A., Schindler N., Solvi A. M. and Wiese J. (2009). The HSG procedure for modelling integrated urban wastewater systems. *Water Science and Technology*, **60** (8), 2065-2075.

Rauch W. and Harremoes P. (1997). Acute pollution of recipients in urban areas. *Water Science and Technology*, **36** (8-9), 179-184.

Rauch W., Seggelke K., Brown R. and Krebs P. (2005). Integrated Approaches in Urban Storm Drainage: Where Do We Stand? *Environmental Management*, **35** (4), 396-409.

Reichert P., Borchardt D., Henze M., Rauch W., Shanahan P., Somlyódy L. and Vanrolleghem P. A. (2001). *River water quality model No. 1 . Scientific and Technical Report No. 12*. IWA Publ., London.

Schütze M. and Alex J. (2004). Suitable Integrated Modelling - Based on Simplified Models. *6th International Conference on Urban Drainage Modelling*, Dresden, Germany.

Shanahan P., Borchardt D., Henze M., Rauch W., Reichert P., Somlybody L. and Vanrolleghem P. (2001). River water quality model no. 1 (RWQM1): I. Modelling approach. *Water Science and Technology*, **43** (5), 1-9.

Sun S. and Bertrand-Krajewski J.-L. (2012). On calibration data selection: The case of stormwater quality regression models. *Environmental Modelling & Software*, **35** (0), 61-73.

Winkler S., Zessner M., Saracevic E. and Fleischmann N. (2008). Intelligent monitoring networks - transformation of data into information for water management. *Water Science and Technology*, **58** (2), 317-322.