Michael MAIR

DYNAVIBE

Dynamic Virtual Infrastructure Benchmarking

DISSERTATION

eingereicht an der

LEOPOLD-FRANZENS-UNIVERSITÄT INNSBRUCK
FAKULTÄT FÜR TECHNISCHE WISSENSCHAFTEN

zur Erlangung des akademischen Grades
DOKTOR DER TECHNISCHEN WISSENSCHAFTEN

Innsbruck, November 2015

Erster Beurteiler:	Univ.-Prof. Dipl.-Ing. Dr. techn. Wolfgang Rauch Universität Innsbruck Institut für Infrastruktur Arbeitsbereich Umwelttechnik
Zweiter Beurteiler:	Univ.-Prof. Dr.-Ing. Dirk Muschalla Universität Graz Institut für Siedlungswasserwirtschaft und Landschaftswasserbau

Hauptbetreuer: Univ.-Prof. Dipl.-Ing. Dr. techn.
Wolfgang Rauch
Universität Innsbruck
Institut für Infrastruktur
Arbeitsbereich Umwelttechnik

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht. Die vorliegende Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Magister-/Master-/Diplomarbeit/Dissertation eingereicht.

Innsbruck, 21. März 2016

Michael Mair

ABSTRACT

Dynamic Virtual Infrastructure Benchmarking (DynaVIBe) is an approach for the algorithmic generation of virtual case studies of the entire urban water infrastructure (e.g. water supply, drainage and sewer systems) to investigate its behaviour and impact on the urban environment and socio-economy over time (mid- and long-term perspective in the context of e.g. urban development and climate change). Furthermore scientific propositions in this context are therefore case study unspecific. Identifying and analysing processes of the urban water cycle can be done with the help of real world case studies. However, data observation and preparation for such studies is a time consuming task and therefore the number of free available case studies for research is limited. The approach of DynaVIBe closes this gap.

The focus of this thesis is to develop algorithms and implement software tools for the algorithmic generation of virtual case studies and evaluating their system performance. This includes the generation and simulation of infrastructure systems on a spatial and temporal scale to map the dynamics of urban water infrastructures into computing models and simulations.

The strength of the algorithms developed in this thesis is the flexibility to use different level of details as input data. Virtual case studies can be generated by using only a few input parameters up to using detailed data sets (generating semi-virtual case studies). Input data sets with different detail levels can explicitly control the validity of model simulation results so that various research questions in the field of urban water management can be investigated. Due to the high flexibility of the generation algorithms surrogate data can be used as input. This approach was applied on street network data sets which are of good quality and free accessible (e.g. OpenStreetMap). Further, the history of an urban infrastructure network can be recreated by extracting street network data out of historical orthophotos. It is proved that the developed algorithms and tools are capable to be applied on real case studies for finding optimal mechanisms for developing and controlling urban water infrastructure systems.

KURZFASSUNG

Dynamic Virtual Infrastructure Benchmarking (DynaVIBe) ist ein Ansatz für die algorithmische Generierung von virtuellen Fallstudien der gesamten städtischen Wasserinfrastruktur (z.B. Wasserversorgung, Entwässerung und Kanalisationssysteme). Dabei kann in weiterer Folge das Verhalten und die Auswirkungen auf die Umwelt und Sozioökonomie im Wandel der Zeit (mittel- und langfristige Perspektive von z.B. Stadtentwicklung und Klimawandel) untersucht werden. Des Weiteren sind wissenschaftliche Aussagen in diesem Sinne Fallstudien unabhängig. Das Identifizieren und Analysieren von Prozessen des städtischen Wasserkreislaufes kann mit Hilfe von realen Fallstudien durchgeführt werden. Datenerhebung und Aufbereitung für Untersuchungen sind ein zeitaufwendiges Vorhaben und daher ist auch die Anzahl der frei verfügbaren Fallstudien für Forschungszwecke beschränkt. Der Ansatz von DynaVIBe wirkt dem entgegen.

Der Schwerpunkt dieser Arbeit ist die Entwicklung von Algorithmen und Implementierung von Software-Tools für die automatische Generierung von virtuellen Fallstudien, und der Bewertung ihrer Systemleistung. Dazu gehört auch die automatische Erstellung und Simulation von Infrastruktursystemen in räumlicher und zeitlicher Auflösung um die Dynamik der städtischen Wasserinfrastruktur abzubilden.

Die Besonderheit der in dieser Arbeit entwickelten Algorithmen liegt in der Flexibilität der Verwendung von Eingabedaten mit unterschiedlicher Detailliertheit. Virtuelle Fallstudien können mit nur wenigen Eingabeparametern bis hin zur Verwendung von bereits detaillierten aber unvollständigen Datensätzen (semi-virtuelle Fallstudien) generiert werden. Durch die unterschiedliche Detailliertheit der Datensätze kann die Aussagekraft von Modellsimulationen gezielt gesteuert, und damit auch verschiedenste Fragestellungen im Bereich der Siedlungswasserwirtschaft, untersucht werden. Die Flexibilität der Generierungsalgorithmen ermöglichte es auch Surrogatdaten als Eingabedaten zu verwenden. Insbesondere wird dies am Beispiel von Straßennetzwerkinformationen gezeigt, welche von guter Qualität und frei zugänglich sind (z.B.: OpenStreetMap).

Des Weiteren kann dadurch auch die Historie eines städtischen Infrastrukturnetzwerks durch das Ableiten von Straßennetzwerkdaten aus historischen Orthophotos nachgebildet werden. Es wird gezeigt, dass die entwickelten Algorithmen und Software-tools Untersuchungen an realen Fallstudien ermöglichen und in weiterer Folge fallspezifisch optimale Mechanismen hinsichtlich der Änderung und Steuerung der Wasserinfrastruktur ermittelt werden können.

ACKNOWLEDGMENTS

I want to thank my parents for giving me the chance and a good base to get such an education. Thank you for all expensive education years.

Thank you to my professor and supervisor for giving me the chance to make this thesis happen and for giving me the opportunity to present my work all over the world at sunny places (except of Kiruna).

I say thank you to Robert as well, my co-supervisor and friend. Hopefully, many working years will follow.

And finally, thank you to my girlfriend. She made me to complete this thesis.

Thank you

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Part A

THESIS

1 INTRODUCTION

Infrastructures are crucial for the development of a settlement area. Especially urban water infrastructures (e.g. piped water supply and sewer system) for supplying fresh water to and disposing polluted water from the settlement area are one of the most important infrastructures in an urbanized area. We are expecting that the system components (e.g. pipes) of our piped urban water infrastructure have a lifetime of many decades up to 100 years. Adaptations of the network structure and system components which we perform today, have an impact on the system performance over a long time period. This adaptations can be triggered by various urban development processes (e.g. population growth). Therefore, it is essential to develop tools for engineers, operators and stakeholders which support them during decision making processes by evaluating new technologies, strategies or measures. However, new knowledge gained from such evaluations are often case study specific and can hardly be transferred to other case studies or generalized. The approach of 'Dynamic Virtual Infrastructure Benchmarking' (DynaVIBe) takes care of this shortcomings. The main objective is to automatically generate virtual case studies of the urban water infrastructure taking into account the dynamics of urban areas (e.g. change of the water demand over time due to population fluctuations). By investigating the system performance of the virtual case studies under different scenarios (e.g. different population growth scenarios) case study unspecific scientific proposition can be made. In this sense, the impact of adaptation strategies on the system performance of the urban water infrastructure can be investigated and tested (Benchmarking). Specific topics within the DynaVIBe approach are handled in 14 scientific publication in international journals and conference proceedings and presented here in form of a cumulative thesis. Six publications are published as first author papers. 13 of 14 are already published and one is 'under review' at the time of writing this thesis (see table 1.1). In eight presentations parts of this thesis were presented at international events (see Chapter 21). This thesis deals with scientific questions from the area of computer science and urban water management modelling and fur-

14 publications, 6 first author publications and 3 software tools ther software development. Therefore three developed software tools (Dynamind, DynaVIBe-Web and Achilles) are part of this work. Most of the work presented here is part of the DynaVIBe project funded by the Austrian Science Fund (FWF): P 23250-N24.

The thesis is split in three main parts which are 1) a short and general 'Thesis' part giving an overview of urban water infrastructure modelling and its computational challenges , 2) the main 'Publication' part containing all relevant publications referenced in the 'Thesis part' and 3) Appendix. Due to character of a cumulative work the second part 'Publications' should be seen as main part.

Aim: algorithmic generation of virtual case studies, development of software tools, evaluating system performance and scenario testing

The main focus of this thesis is to develop algorithms and implement software tools for the algorithmic generation of virtual case studies. This includes the generation and simulation of infrastructure systems on a spatial and temporal scale. Moreover, the generated case studies are tested and system performance is evaluated under different scenarios (Benchmarking).

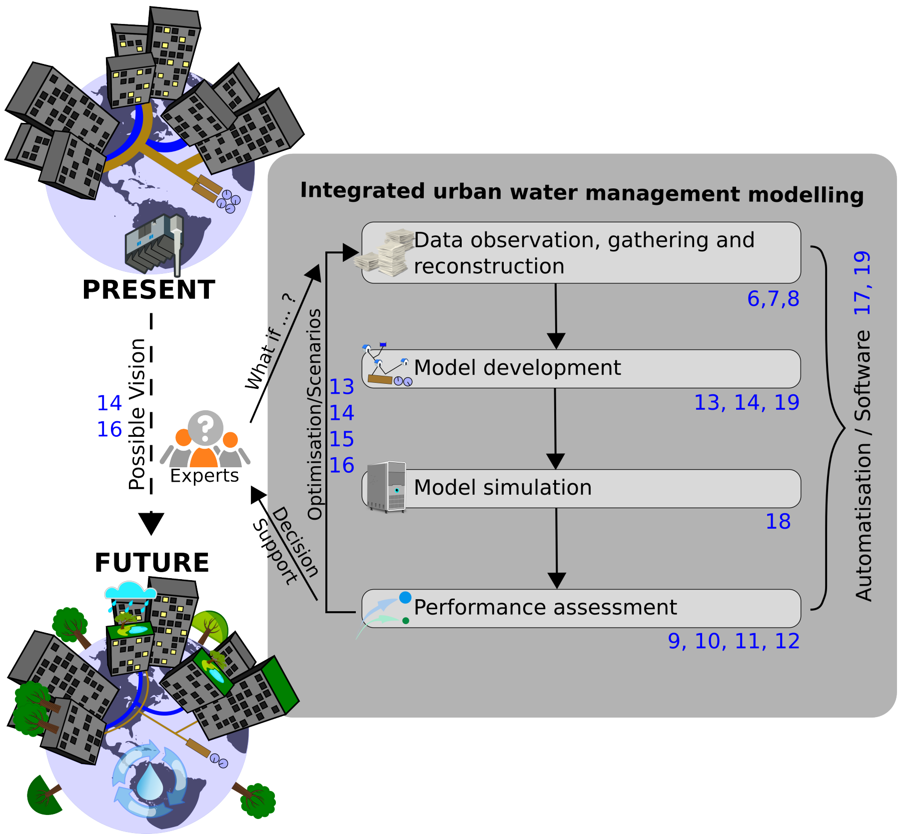

Figure 1.1: Integrated urban water management — Publications in the context of long-term urban water management presented in this thesis — *Related publications/chapters are marked in blue*

Figure 1.1 shows the big picture of the scope of this thesis. The main aim is to model, simulate and benchmark the dynamics of the entire water cycle of an urbanized area over a timespan of several decades. For example the inhabitants of a city have the vision of a city with more green areas, less pollution and in general more attractive places. It is clear that such visions can only be realized in a long-term perspective. The question is: How can we make such visions real, and how do we get there safely? Such scientific questions are investigated in the work 'Stochastic performance assessment and optimization strategies of the water supply network transition of Kiruna during city relocation' (Chapter 16). This work deals with the vision of relocation huge parts of the city 'Kiruna' due to mining activities below the city. In this example we have to deal with a deconstruction and construction of the urban water infrastructure at the same time. Although many challenges are case specific and individual especially in daily operation work of carriers, there are research questions regarding long-term impacts of system adaptations on the urban environment, which can be answered on an abstracted ('virtual') level of the urban water infrastructure ('Modelling Dynamic Expansion of Water Distribution System for New Urban Developments' (Chapter 14)). At the time of writing this thesis no software tool exist which is able to answer such question automatically. A state-of-the-art approach is trying to reach such visions by realizing them in form of 1) a master plan developed by experts (e.g. Engineers and Stakeholders) and 2) implementing it with the permission of decision-makers. Crucial for that is the assurance of the master plan's reliability. This implies some kind of 'What if ... ?' questions where we try to answer the impact of changes in urban structures on the entire urban water infrastructure. The answers of such questions may help the experts in their decision-making process. All these questions are part of the topic 'Integrated urban water management' (Figure 1.1 - grey box). Investigations in this area are often realized by assessing the system performance over time with the help of model simulations. Crucial steps thereby are 1) data observation, gathering and reconstruction, 2) model development, 3) model simulation and 4) performance assessment. Depending on the simulation results and the system performance assessment respectively, these four steps can be repeated for the purpose of the system performance optimization. This thesis contributes with various publications in all five modelling steps.

How can we make visions real, and how do we get there safely?

DATA OBSERVATION, GATHERING AND RECONSTRUCTION: The type, quantity and quality of the real world data needed for model development, simulation, calibration and verification is depending on the modelling aim and therefore can vary. Hence, data observation can easily be a time intense and expensive task. It is of great interest to minimize this effort. One promising approach is to use all already available data sources that are associated to the data required for modelling. The paper 'Where to find water pipes and sewers? – On the correlation of infrastructure networks in the urban environment' (Chapter 6) investigates the correlation of urban infrastructures (street, sewer and water supply networks) and the usability to use street network data as surrogate data for sewer and water supply network data. From a pure spatial viewpoint of system components it is not guaranteed that this data describes a functioning system (e.g. all pipes of a water supply network are connected to one network). The papers 'Improving Incomplete Water Distribution System Data' (Chapter 8) and 'Spanning Tree Based Algorithm for Generating Water Distribution Network Sets by Using Street Network Data Sets' (Chapter 7) are presenting methods for data reconstruction with the focus on water supply network modelling. The basis of the reconstruction methods are graph theory based algorithms.

MODEL DEVELOPMENT: In general this task itself includes many sub-tasks beginning with model setup, calibration and ending with verification of the model simulation results. It is not always clear how much input data is needed such that the developed model fulfils its purpose. The work 'Assessing Model Structure Uncertainties in Water Distribution Models' (Chapter 13) investigates the impact of the quantity of observed pressure points within a water supply system on the quality of the simulated pressure surface. The topic 'Model development' ranges from setting up models in existing software tools (e.g. EPANET2 or SWMM5 for water supply network and storm water management modelling, respectively) to developing completely new innovative models and tools to extend the range of possible applications. The work 'Modelling Dynamic Expansion of Distribution Systems for New Urban Developments' (Chapter 14) investigates the impact of spatial and temporal dynamics of projected water demand on the system performance of water supply systems. The basic approach here is to integrate a land use and population model into an infrastructure development model.

MODEL SIMULATION: With increasing model complexity and size of used data often the runtime of model simulations increases as well. Long-term urban water management simulation are commonly based on the approach to test a city's master plan under various scenarios. In the work 'Performance improvement with parallel numerical model simulations in the field of urban water management' (Chapter 18) the approach of parallelizing independent model simulations to decrease the overall computational runtime is assessed.

SYSTEM PERFORMANCE ASSESSMENT: Model simulation results have to be summarized for further usage. The publication presented in the chapters 9, 10, 11 and 12 investigate different local and global system performance indicator. These can be used for graphical representation (e.g. colored GIS maps) in decision making processes or as objective functions for the purpose of optimizing a system.

OPTIMISATION / SCENARIO TESTING: This task represents the application of the developed model and its simulation results (Chapters 13 to 16). For example the publication 'Stability of Traditional Urban Water Systems — Integrated Assessment of Transitions Scenarios' (Chapter 15) presents a model with the aim to represent urban water systems in urban areas. Its application is to show the impact of different population development scenarios on the overall system performance (Scenario testing — Benchmarking).

In the context of integrated urban water management modelling many tasks have to be solved by hand. Therefore it is worthwhile to automatize as many as possible in form of algorithms and further a software tool. The paper 'DYNAMIND - A Softwaretool for Integrated Modelling of Urban Environments and their Infrastructure' (Chapter 17) presents a software dealing with all aspects in the context of integrated urban water management modelling. The main approach thereby is to model an entire city with its water cycle and making one comprehensive simulation. An important point for developing and setting up new models is the availability and usability of software tools. The rapid change and development of new technologies for managing huge digital data sets and delivering new software products in form of 'on-demand' services opens new opportunities for data management and modelling processes in the field of urban water management modelling. These new technologies may help to reduce DYNAMIND's complexity and increase the usability (Model developers need knowledge in programming and GIS). The paper titled 'The application of Web-geographic information system for improving urban water cycle modelling' (Chapter 19) demonstrates a prototype of a web based modelling and simulation software designed for experts with less programming and GIS skills.

Chapter	Scopus	Publication title	Published	Status
6	yes	**Where to find water pipes and sewers? – On the correlation of infrastructure networks in the urban environment**	CEUS	R
7	yes	**Spanning Tree Based Algorithm for Generating Water Distribution Network Sets by Using Street Network Data Sets**	EWRI	P
8	yes	**Improving Incomplete Water Distribution System Data**	PI	P
9	yes	**GIS-based applications of sensitivity analysis for sewer models**	WST	P
10	yes	GIS Based Applications of Sensitivity Analysis for Water Distribution Models	EWRI	P
11	yes	Cascade vulnerability for risk analysis of water infrastructure	WST	P
12	yes	Identifying Hydropower Potential in Water Distribution Systems of Alpine Regions	EWRI	P
13	yes	Assessing Model Structure Uncertainties in Water Distribution Models	EWRI	P
14	yes	Modelling Dynamic Expansion of Distribution Systems for New Urban Developments	EWRI	P
15	yes	Stability of Traditional Urban Water Systems — Integrated Assessment of Transitions Scenarios	PI	P
16	yes	Stochastic performance assessment and optimization strategies of the water supply network transition of Kiruna during city relocation	EWRI	P
17	no	DYNAMIND - A Softwaretool for Integrated Modelling of Urban Environments and their Infrastructure	HIC	P
18	yes	**Performance improvement with parallel numerical model simulations in the field of urban water management**	JHI	P
19	yes	**The application of Web-geographic information system for improving urban water cycle modelling**	WST	P

Table 1.1: Publication list — Titles written in **boldface** are first author papers — *Scopus:* yes = listed in Scopus search egnine; no = not listed in Scopus search engine; — *Published in Journal:* WST = Water Science and Technology; JHI = Journal of Hydroinformatics; CEUS = Computers, Environment and Urban Systems; PI = Procedia Engineering; — *Published in Proceedings paper:* EWRI = World Environmental and Water Resources Congress; HIC = International Conference on Hydroinformatics; — *Status:* P = Published; R = Under review; S = Submitted

2

INTEGRATED URBAN WATER MANAGEMENT MODELLING

One of the most important fundamentals of a modern civilization and their development is the access to water with an accurate quality at any time and place. This can be seen by the fact that human settlements can always be found next to a huge water source (e.g. springs, rivers, huge groundwater bodies). With increasing population of a settlement and consequently the population density also the need for an efficient water supply system increases. The development started from simple groundwater wells, where each inhabitant gets the drinking water from central points within the settlement area, up to modern supply systems, which supply the water directly to the point of usage (e.g. in each household). It is not surprising that the development and operation of such systems was and still is one of the most complex engineering challenges. Water is needed for various applications (e.g. toilet flushing, washing machine, usage by industry) which are often resulting in polluted water and further in a polluted environment. Due to the increasing pollution of the environment, the need of systems arose, which dispose polluted water from a settlement area, with the focus on a controlled ascription of urban water back to the natural water cycle. This includes the discharge of storm water (grey water) for flood avoidance and disposal/treatment of heavy polluted water (black water).

Both systems (System for bringing the water into a settlement area and system to bring the used water back to the natural water cycle) are developed/constructed and operated by human beings to fulfil basic requirements for a positive development of the society, hence these systems are also known as urban water infrastructure.

2.1 INTEGRATED URBAN WATER MANAGEMENT

A state-of-the-art piped water infrastructure has a lifetime of several decades (up to 100 years and more), hence this period of time is the planning horizon which has to be considered by engineers. It is crucial that future infrastructures have to perform at least as good as our current infrastructures, but it has also to be mentioned that changes we are performing today have an impact on the future urban water infrastructure. This may not always result in an intended behaviour. For example if we plan a future water supply infrastructure with the assumption that there is a constant population growth, and in reality a population shrink occurred may result in a high water age and consequently in a bad water quality. Sitzenfrei et al. (2015) investigated the impact of water supply network development decision on the system performance of the city Innsbruck over a timespan starting in 1900 and ending in 2010. In this case study an unforeseen high water demand occurred due to a malpractice of the inhabitants (they never closed the water-tap) and further led to minimum pressure issues in the system. There exist a multitude of studies which investigate this topic from a wider viewpoint by taking into account the whole urban water infrastructure and the expectations of the population in the future. Worth mentioning in this context is the work of Brown et al. (2009) which investigated the city transition when moving its water infrastructure towards sustainable urban water conditions.

This shows that urban water management deals, beside daily operational tasks of the urban water infrastructure, also with managing a city's water infrastructure with a planning horizon of several decades. Conventionally the field of urban water management can be split up into three different rough topics which are:

- Management of systems for supplying water
- Management of systems for discharge storm water
- Management of systems for disposal/treatment polluted water

It easily can be observed that the three systems differ in their basic usage. The development took place at different places and times on earth. For example approximately 3000 B.C. the first open channel drainage system was built in Lothal (India). The development was mainly driven due to a previous flood event where huge parts of the settlement area were destroyed (De Feo et al., 2014). The first piped water distribution system was installed 1500 B.C. in Crete (Walski et al., 2003). The first installation of 'modern' sewerage systems were documented in the 19th century, mainly driven by increasing concerns of the public health in London (Benidickson, 2011). Maybe because of the historical development, today's understanding of urban water management is still split up in the same categorization.

However, research of the last several years has shown that it is necessary to analyse the urban water cycle from a more integrated view point. In detail this means it is necessary to investigate the urban water cycle from the viewpoint of one huge system, including 1) all three previous demonstrated systems (Pikaar et al., 2014; Rauch and Kleidorfer, 2014) and 2) systems which have an impact on the development of a city (e.g. climate, socio-economics) (Mitchell et al., 2001; Rozos and Makropoulos, 2013; Willuweit and O'Sullivan, 2013). In the last years research focused on the transition from current/traditional urban water infrastructures to sustainable urban water infrastructures in a long-term perspective (Marlow et al., 2013; Ferguson et al., 2013; Makropoulos et al., 2008). Important during transition processes is to guarantee operability of all systems all the time. This implies that the transition strategies and systems are flexible enough to capture a broad range of different scenarios (Basupi and Kapelan, 2013; Urich and Rauch, 2014). Chapter 16 'Stochastic performance assessment and optimization strategies of the water supply network transition of Kiruna during city relocation' demonstrates the benefits of including the aspect of climate change, urban development and social changes in urban water management and investigates system performance of the water supply system during the transition process to identify critical system states.

2.2 SYSTEM ANALYSES

Numerical models and numerical model simulations are a state-of-the-art tool to analyse and investigate processes in urban water infrastructures (Sitzenfrei et al., 2014). The advantage is to answer so called 'What if ... ?' questions (e.g. What if a certain pipe collapses ?) without altering the real world system. Moreover, model simulation results can be used as basis for decision-making (e.g. Define a rehabilitation order for pipe replacement according to the impact on the overall system performance in case of pipe bursts) or to proof the efficiency of systems in terms of service quantity and quality (Carstensen et al., 1997). Bach et al. (2014) summarised the historical evolution of integrated modelling starting in the year 1970 up to present day. The most popular simulating programs for water supply and urban drainage infrastructures are EPANET2 (Rossman, 1999) and SWMM5 (Huber et al., 2005), respectively.

The general workflow for developing models are as followed:

1. Defining the modelling aim
2. Select the software (e.g. EPANET2, SWMM5, develop a new one)
3. Collect model parameters (e.g. diameter, length, building year of pipes)
4. Collect model input data / data observation (e.g. rainfall data)
5. Collect real system behaviour data (e.g. Inflow and rain series of a combined sewer system at the waste water treatment plant during a rain event)
6. Enter data into the selected program
7. Select uncertain model parameters (e.g. pipe roughness)
8. Perform a sensitivity analyses on uncertain model parameters
9. Split real system behaviour data into a (S1) calibration set and (S2) validation set
10. Perform rough calibration of highly sensitive uncertain parameters on the subset S1
11. Perform fine calibration on lower sensitive uncertain parameters on the subset S1
12. Validate model results with subset S2
13. Use model according to its modelling aim (Model simulations)

Developing a model is an iterative process, because in most cases several uncertain or even unknown model parameters are included, which have to be changed during model calibration. This means step 8 to 11 have to be repeated until the deviation between model simulation results and real system data is minimized. All piped network models representing one case study are useful for simulating e.g. daily operations, energy management in general. It is crucial for such applications to use detailed data with a high quality as model input. Albeit it is possible to model short and mid-range real world system behaviour with that approach, the question arises how to model more complex long-term system development (e.g. for master planning), where detailed data for model development, calibration, validation and usage is not available? Moreover, in the context of integrated modelling this whole model development process gets more complicated. Here the challenge is to couple models from several scientific disciplines to one huge accurate and valid model (Voinov and Shugart, 2013).

2.3 DYNAMIC VIRTUAL INFRASTRUCTURE BENCHMARKING

Analyses of case studies with the aid of model simulations are the driving force of research in the field of urban water management. Therefore, Jolly et al. (2014) developed a research database of water distribution system models summarizing and storing all information and data of real, hypothetical models as well as tools for automatically generating artificial models with certain characteristics. With such a database, the scientific community has the ability to develop and validate new algorithms on the basis of a huge model data set. A limited number of public accessible models for research purpose where early available. One of the first was the New York Tunnel system in 1969 (Schaake and Lai, 1969). Over the years other real and artificial models followed (e.g. hypothetical two loop system (Alperovits and Shamir, 1977), Anytown (Walski et al., 1987), Network 1 and 2 (Ostfeld et al., 2008), C-Town (Ostfeld et al., 2011), Net2 and Net3 from EPANET and Micropolis and Mesopolis (Brumbelow et al., 2007)). Albeit these models fulfil the requirements to develop and test new scientific methods in the context of water distribution system modelling, the scientific results are case study specific due to the limited set of free available models. It is worthwhile to test new methods on as many case studies as possible to prove its correctness and gain credibility. The main reason for the limited number of case studies is the time consuming and complex task to create a model of a real case study — which can be easily estimated with several man-months. The time intense part can be revert to data observation for model setup, calibration and evaluation. Although it is essential for many investigations to build a detailed model, especially if they are case study specific (e.g. impacts of a water supply pipe collapse, 1D/2D flooding investigations) there exist questions which can be answered on more 'coarse grained' models. Thus time for data observation can be decreased. Sitzenfrei et al. (2010a) investigated a new approach of automatically generating virtual case studies (Virtual Infrastructure Benchmarking - VIBe) of urban structures including all water infrastructures. Moreover, they presented a graph-based approach to generate an unlimited set of water distribution system models (Sitzenfrei et al., 2010a; Möderl et al., 2011). Urich et al. (2010) presented an agent-based method to generate virtual sewer systems. The generated models represent systems with one state in time. This means mid-term (some years) and long-term development (some decades) of the water infrastructure are not considered. The approach of Dynamic Virtual Infrastructure Benchmarking (DynaVIBe) also focuses on the dynamics of urban settlement areas, its urban development and therefore changes in all water infrastructures (Sitzenfrei et al., 2010b). In this context the system performance over time of the water infrastructure can be assessed under different scenarios (Benchmarking).

2.3.1 *Data requirements for modelling and potential data sources*

Another aspect of the DynaVIBe approach embraces the time reduction for model development by investigating the potential of using already available and easy accessible surrogate data sources. The abstract figure 2.1 shows possible data sources for models to reach accurate model simulations for answering specific questions in the field of urban drainage modelling. Similar considerations can be made for water distribution network modelling.

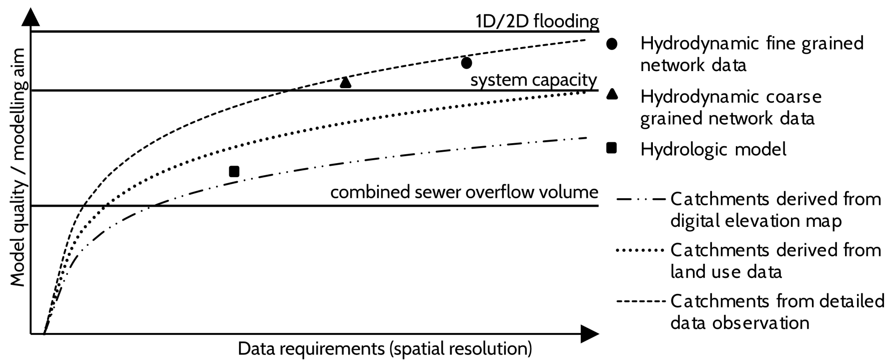

Figure 2.1: Amount of data needed for reaching a certain model quality according to a certain modelling aim (Hypothetical image)

The amount of data (X-Axes) can be interpreted as the spatial resolution of the data source, where an increase of the special resolution implies an increase of the model quality in general (Y-Axes). However, it does not imply an increase of the validity. Depending on the type of data sources and level of detail (e.g. catchment data derived from digital elevation, land use data or direct data observation) the resulting models have maximum boundaries for its validity (Figure 2.1 dashed, dotted and dashed-doted parabolic line). Model simulations using a limited amount of input data (rough spatial resolution) may deliver accurate model simulation results and an increase of the spatial resolution is expected to have a parabolic impact on the model quality. However, core message is that with increasing the spatial resolution of the input data we cannot expect to change the validity of a model and its simulations. In the same manner this is assumed to be valid for other piped infrastructure types like water supply networks. In the VIBe approach this idea goes one step further, where model input data are only boundaries describing a set of case studies with common characteristics. These boundary conditions are further used for automatically generating needed input data for more detailed models like EPANET2 or SWMM5. VIBe enables research on systems at a specific point in time (e.g. pressure distribution of a virtual water supply network case study at steady state). However, with the focus on urban water management including long-term planning (master planning) the integration of the temporal dimension into the overall model is essential for modelling the development of an urbanized area with its infrastructure. This approach is the focus of Dynamic Virtual Infrastructure Benchmarking (DynaVIBe). Here the virtual case studies are automatically generated according to predefined boundary conditions with the addition of using input parameters representing the dynamics of an urban environment and their influencing phenomenon (e.g. population growth, climate change).

Beside the main aim (use as minimal data as possible for generating virtual case studies) the developed models are also highly responsive to use more detailed data for generating more realistic case studies and therefore resulting in more accurate infrastructure models. Due to the strong correlation in layout design of piped infrastructure and street network data, the latter one can be used as surrogate data for automatically generating information on the layout of any network infrastructure. Fundamentally for that is the knowledge about the strength of the relationship (e.g. likelihood of a water supply pipe to be underneath a street).

2.3.2 *Piped network layout generation*

In literature several algorithms exist for automatically generating valid piped network layouts. Haghighi (2013) developed a loop-by-loop cutting algorithm for generating urban drainage systems. Urich et al. (2010) propose an agent-based algorithm for generating virtual sewer systems. Work has also been done in the development of tools for water supply layout generation (De Corte and Sörensen, 2014; Muranho et al., 2012; Trifunović et al., 2012; Sitzenfrei et al., 2013). The main aim of most tools is to generate huge sets of virtual/artificial water infrastructure models for certain model aims to overcome the lack of real world case studies. However, most of them are using explicit input data directly related to the model domain (e.g. data explicit observed for the usage in water supply modelling). In chapter 6 'Where to find water pipes and sewers? — On the correlation of infrastructure networks in the urban environment' the correlation between street, water supply and sewer networks is investigated. A strong correlation between street and piped network layout could be found. Hence it is reasonable to investigate the usage of street network data as surrogate data for water infrastructure information in such algorithms. In a first step this enabled the development of an algorithm presented in chapter 7 'Spanning Tree Based Algorithm for Generating Water Distribution Network Sets by Using Street Network Data Sets'. Its correctness is proven in chapter 8 'Improving Incomplete Water Distribution System Data'. A responsive algorithm is presented where an arbitrary sized set of known piped network data (information about location and dimension) for water distribution systems can be used. Missing data is automatically generated out of street network data (layout reconstruction). The latter is easy to access for nearly each region worldwide (e.g. Open Street map and Google maps), hence there is a huge potential to decrease time needed for modelling. Especially, if we think on modelling the dynamics of historical water distribution infrastructures, street network data can be derived from historical ortho photos. After generating the layout of the network, the pipe sizing can take place.

2.3.3 *Pipe sizing*

To develop a proper model the properties of network elements like pipe diameters, pipe roughness and pipe profile have to be observed or estimated. The later one implies highly uncertain model parameters and hence highly uncertain model simulation results. According to the result of investigations performed in the paper 'Where to find water pipes and sewers? – On the correlation of infrastructure networks in the urban environment' (Chapter 6 exist no correlation between street types (e.g. primary or secondary roads) and network element properties of urban water infrastructures (e.g. pipe diameter). Hence using street network data as surrogate data for water infrastructure networks is only valid in terms of network layout reconstruction/generation. All other network properties have to be observed (if the developed model represents a real world system) or approximated (if it is a virtual model representing a set of real world models). The method used for property approximation strongly depends on the infrastructure type. This can be illustrated by the hydraulic pressure property of water supply systems (system under pressure) and urban drainage systems (free surface flow under normal conditions).

Focusing on the diameter distribution of pipes, country specific technical design methods can be used (Muranho et al., 2012). Based on such design methods algorithms for automatic pipe sizing can be developed. Blumensaat et al. (2012) presented a method for sewer modelling to approximating the diameter distribution based on a surface flow accumulation algorithm and technical design methods for hydraulic dimensioning.

Saldarriaga et al. (2011) investigated a method for predetermining a pressure surface and further designing water distribution systems. Based on this method the algorithm presented in 'Spanning Tree Based Algorithm for Generating Water Distribution Network Sets by Using Street Network Data Sets' (Chapter 7) was developed. In this work realistic virtual water supply network models were generated with minimal data requirements. When using additional information, like real world pressure measurements at junctions of the real system, the generated models are even more close to reality. When using this additional information a more realistic pressure surface can be approximated, hence pipe diameters close to reality are calculated. This work was presented in 'Assessing Model Structure Uncertainties in Water Distribution Models' (Chapter 13).

2.4 PERFORMANCE ASSESSMENT

Depending on the modelling aim performance assessment of water infrastructure networks can be based on local (spatial resolution e.g. maximum pressure at a specific junction in a water supply system) and global assessments (e.g. combined sewer overflow efficiency in sewer models) (Yazdani and Jeffrey, 2011). The latter is commonly used when assessing 1) short/mid-term (range of several hours/-days) infrastructure performance of a system stressed with different scenarios (e.g. impact on the reliability of a water supply system in case of a pipe burst (Möderl et al., 2007)) or 2) long-term (range of several decades) infrastructure performance simulations driven by a transformation of the system due to e.g. climate change and urban development. Publications dealing with the first case are 'GIS-based applications of sensitivity analysis for sewer models' (Chapter 9) and 'GIS Based Applications of Sensitivity Analysis for Water Distribution Models' (Chapter 10). Here, the impact of successively altering spatial model parameters (e.g. change the roughness coefficient of one pipe) on the global system performance is investigated. Each performance indicator is spatial join to the element which was altered for simulation. After normalizing all indicators a network map shows system elements with high and low impact on the overall system performance. This approach can also be applied to determine the vulnerability of water infrastructures to investigate one system failure (e.g. one pipe burst) at a time (chapter 10 and 9) and simultaneous system failures (e.g. pipe burst and pump failure at the same time). Latter is demonstrated in the publication titled 'Cascade vulnerability for risk analysis of water infrastructure' (chapter 11). Parts of this are implemented in a software tool called Achilles, which is realized as a module set for the open source SAGA GIS (Figure 2.2).

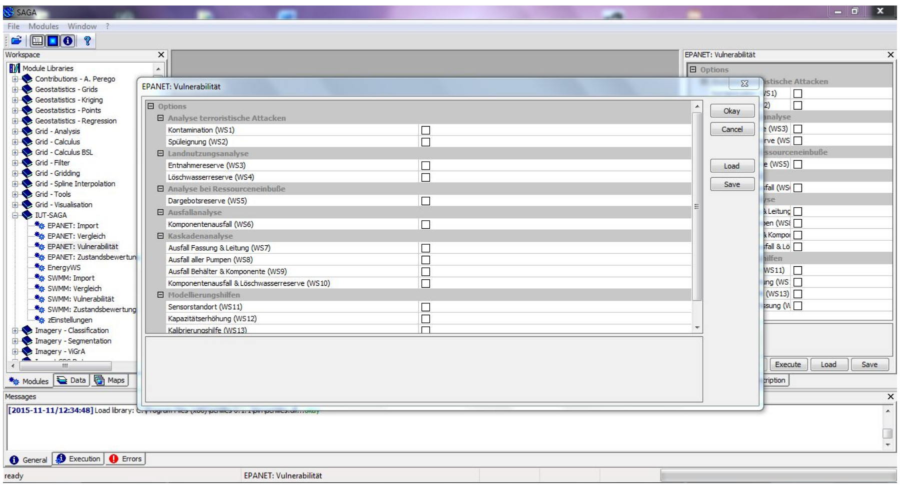

Figure 2.2: Screenshot of the software tool 'Achilles'

2.5 APPLICATIONS

The range of possible applications using the methods presented in this thesis are versatile. Following publications are examples for possible applications of GIS-based sensitivity analyses and the DynaVIBe approach.

- Identifying Hydropower Potential in Water Distribution Systems of Alpine Regions — chapter 12
- Assessing Model Structure Uncertainties in Water Distribution Models — chapter 13
- Modelling Dynamic Expansion of Distribution Systems for new Urban Developments — chapter 14
- Stability of Traditional Urban Water Systems — Integrated Assessment of Transitions Scenarios — chapter 15
- Stochastic performance assessment and optimization strategies of the water supply network transition of Kiruna during city relocation — chapter 16

3 SOFTWARE

With the approach to represent a complete urban environment — 'digital city' and simulate its development, several requirements for a computational realization can be met. In this chapter the basic concept for implementing urban water management models and simulation in the context of DynaVIBe are summarized, challenges are identified and possible solutions are suggested.

3.1 DynaMind - A framework for modeling the dynamics of an urban environment and their infrastructures

DynaMind is an open source scientific workflow engine operating on data with a resolution in space and time. It is optimized for model simulations in the field of urban water management (see chapter 17). The workflow engine is comparable with the Model builder of ArcGis or QGIS. However, they still have some downsides at managing data with a time dimension. The code base of DynaMind was developed over years driven by various scientific projects with the intention to reuse programming code/algorithms, create a high performance system (regarding model simulation runtime and memory management) and a flexible code base for fast prototyping at the same time. Code fragments regarding the optimized runtime and memory management issue are realized in C++, whereas the capability for a fast prototyping is realized by embedding a Python interpreter.

In DynaMind, modules (small working tasks) are linked together to a workflow for simulating an arbitrary urban environment. Modules are either implemented in C++ or Python by realizing the DynaMind module interface class. DynaMind dynamically loads all availably modules during runtime for further usage during model development and simulation. Each module has in-ports to process data into the module and then further to out-ports, thus being able to represent a complex network. These modules can be linked together to a model by using either the simulation interface or the graphical user interface. Figure 3.1 shows an example of a basic DynaMind application, where modules are assembled to a DynaMind workflow representing a urban water management model.

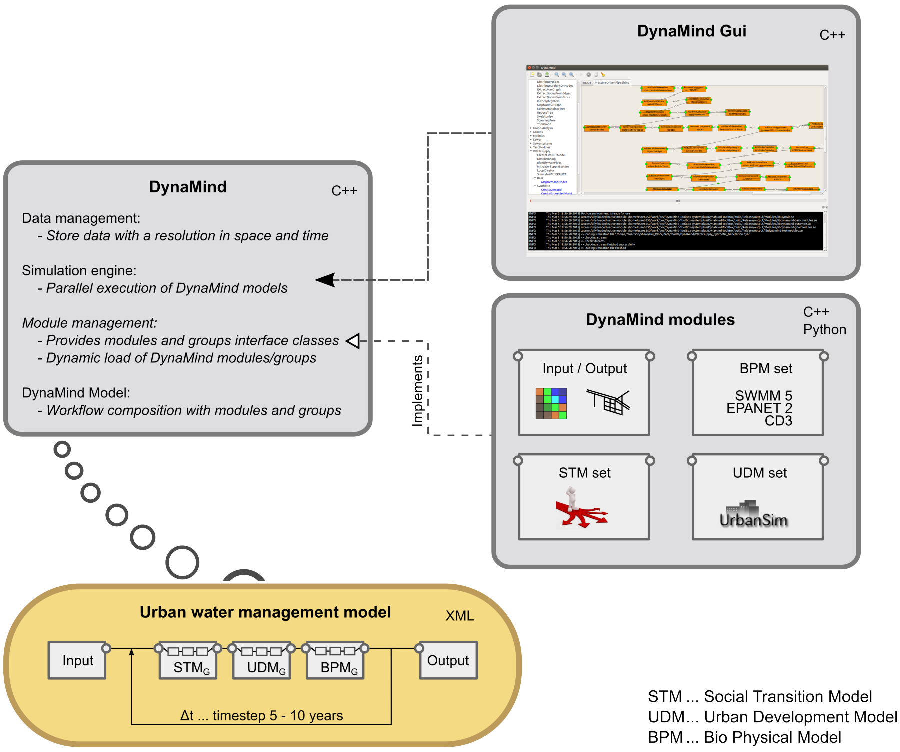

Figure 3.1: Basic components of DynaMind

Every urban water management simulation has its own data storage. Before the urban system can be evolved into the future the data storage needs to be initialized with the urban system at the current state (Figure 3.1 'Input' module). Next the generic modules of the STM, UDM and BPM are loaded from an external library and added to the workflow. These modules can access/create/modify predefined and specific data from the workflow data storage. The access rights can be hard coded in the module or set by the model developer during model development. Introducing access management in this framework enables the simulation engine 1) to evaluate data availability for a model prior to simulation and 2) perform a parallel execution of the models at the level of the workflow definition / model description. The simulation interface of DynaMind provides access to all available data storages at any time of model simulation. This is important for both, the scenario input module and the reporting and presentation module. The last one allows data visualization in form of plots or date export for other programs.

The complexity of the modules can strongly vary, hence runtime of each module varies as well. State-of-the-art computers have many central processing units on one tile to decrease the overall runtime. The aim of the workflow engine is to fully utilize all available hardware during model simulation to ensure an acceptable simulation runtime. The publication titled 'Performance improvement with parallel numerical model simulations in the field of urban water management' is dealing with this topic (see chapter 18).

3.2 THE POTENTIAL OF WEB SERVICES

The W3C (World Wide Web Consortium) defines a Web service as 'a software system designed to support interoperable machine-to-machine interaction over a network'. The implementation of such systems has a huge potential for various use cases where a direct communication between different machines is necessary. Many of us are already using web-services in form of e.g. web-portals for booking flights. Here, flight information between airlines and the web-portal is automatically exchanged via web-services for finding and booking the best suitable flight connection. The usage of web-services is an established technology in daily business processes with a variety of benefits like:

- Web-services are an open and standardized method to realize machine communication via network. This is fundamental to reach various different computational platforms.
- Data can be stored, accessed and modified remotely, enabling computational intense operations on huge data sets on nearly any computing device (e.g. mobile devices).

Model simulations in the field of urban water management are often computational time intense calculations on huge data sets. The benefits from using web-services are in line with these characteristics, hence it is beneficial to implementing model simulations based on web-services. In the work 'The application of Web-geographic information system for improving urban water cycle modelling' in chapter 19 web-service concepts were successful applied on DynaMind models. For the realization the well known standards WMS (Web Map Service), WFS (Web Feature Service) for data exchange and WPS (Web Processing Service) for data manipulation defined by the OGC (Open

Geospatial Consortium) were used. The overall operativeness of the concept was proved on two use cases dealing with 1) identifying the impact of additional catchment areas (in terms of flooding return periods) on an existing sewer system and 2) designing infiltration measures. In the following section another use case based on generating virtual water supply networks with minimal data is demonstrated.

3.2.1 *DynaVIBe-Web*

DynaVIBe-Web, is a free available and online accessible web frontend application for generating virtual water infrastructure networks for any place on earth. The tool has been developed as part of this thesis and the DynaVIBe project and was successfully applied on several cities. In most cases the main application was the generation of virtual water supply network models (EPANET2 models) for testing newly developed or existing optimiziation algorithms and/or network hydraulics solvers. The needed data for DynaVIBe-Web is a digital elevation map, service area, the total demand and position of water sources. Using the strong correlation between street and water distribution network (see chapter 6), an arbitrary sized set of technical feasible water distribution models with different characteristics can be generated and used for benchmarking and other scientific applications. The easy accessible platform also includes an archive to share all model sets of different case studies with the whole scientific community (Figure 3.2).

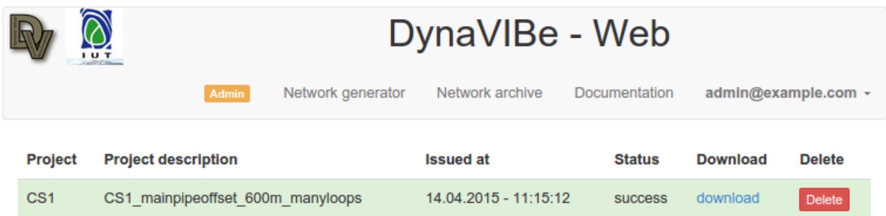

Figure 3.2: DynaVIBe web status page

DynaVIBe-Web is a proof of concept for running complex model simulations in the web. By outsource computational calculation (simulating DynaMind models) to dedicated servers the client side can be unburdened. Besides that, from the perspective of software development and deployment it is a big advantage to rollout software updates fast to guarantee that all clients are updated instantly.

Figure 3.3 shows the input mask of DynaVIBe-Web with an open street map (left side) and a parameter input section (right side). The list after the figure describes all input parameters in detail.

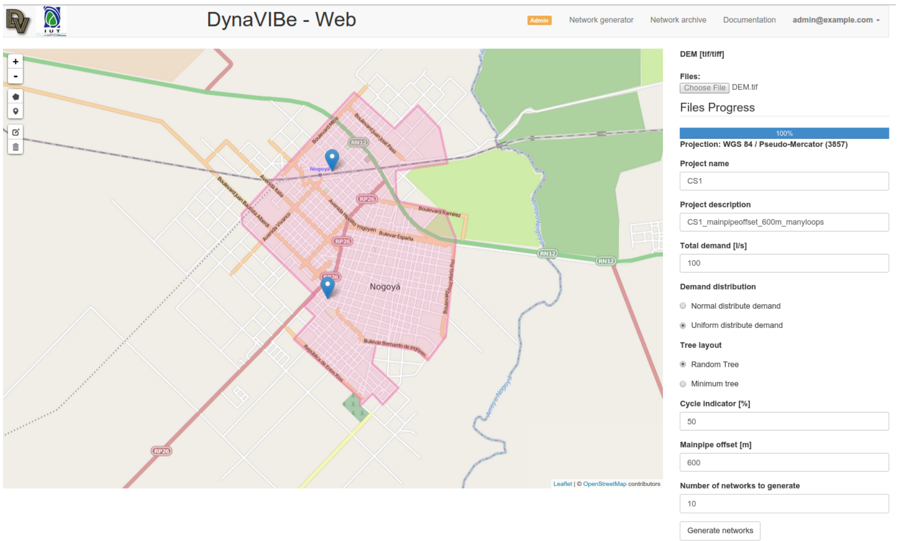

Figure 3.3: DynaVIBe web input mask

TOTAL DEMAND (L/S) Total amount of water supply in the defined service area polygon in litre per second.Total amount of water supply in the defined service area polygon in litre per second.

DEMAND DISTRIBUTION (-) The total demand is distributed on demand points which are automatically generate within the service area. You can choose the type of the point distribution. If you have a high urbanized area with a high population density in the center, you may choose "Normal distribute demand

TREE LAYOUT (-) The minimal operating water supply network layout is automatically generated based on a "Random" or "Minimum" spanning tree using OpenStreetMap data as input.

CYCLE INDICATOR (%) Defines the condition when adding an additional flow path between two water supply system nodes. E.g. 50% Means only additional flow paths (water supply pipes) are added if the alternative path length between two nodes is smaller than 50% of the flow path length in the spanning tree.

MAIN PIPE OFFSET (M) Places pipes with a wider diameter next to a defined polygon. This polygon is calculated by creating an offset polygon of the area of interest.

NUMBER OF NETWORKS TO GENERATE (-) Number of virtual water supply network models to generate (EPANET2 input files).

Once all input data is valid, the networks can be generated. Figure 3.4 shows one possible network generated with DynaVIBe-Web (left pressure distribution and right elevation map). The red circle in both images shows the main pipes of the network. They are automatically generated in dependency of the 'Main pipe offset' parameter.

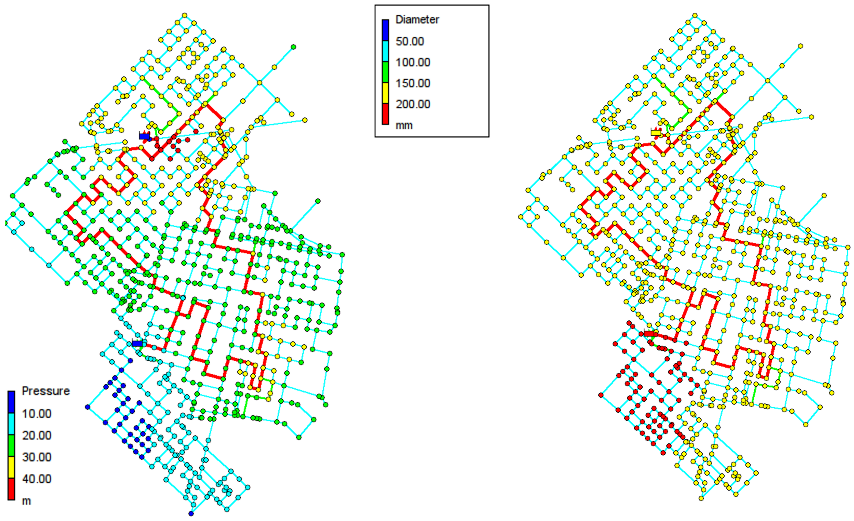

Figure 3.4: Generated water supply model with DynaVIBe web

4 SUMMARY, CRITICAL REFLECTION AND FUTURE RESEARCH

Dynamic Virtual Infrastructure Benchmarking (DynaVIBe) is an approach for the algorithmic generation of virtual case studies of the entire urban water infrastructure to investigate its behaviour and impact on the environment and socio-economy over time (mid- and long-term perspective). This includes the integration of water supply and sewer systems as well as projections of the urban development and climate change. Results from such investigations can be used by the scientific community, stakeholders and urban water infrastructure carriers to better understand the impact of system changes (e.g. implementing decentralized solutions) we are performing today, on the system performance in the future. With virtual case study model simulation and benchmarking it is possible to make case study unspecific scientific propositions. Moreover, this approach closes the gap of missing free available case study data sets for research purpose, which can be frequently traced back to the huge effort for data observation/preparation and legal aspects.

The aim of this thesis was the development of algorithms and software tools in the context of the DynaVIBe approach. 14 publications contribute to topic 'integrated urban water management and modelling' and more specifically to 1) data observation, gathering and reconstruction, 2) model development in general, 3) model simulation and 4)performance assessment of water infrastructure systems (Chapter 1). Developed algorithms and software tools are attributable to all four categories at the same time. Hence they are combined in one integrated model and realized/implemented in the DynaMind framework. DynaMind is a scientific workflow engine designed for handling huge geo-referenced data sets with a temporal dimension. This enables model developer to digital represent an urban water infrastructure over time (digital city) and assess the system performance with model simulations (Chapter 3).

Special focus was on developing algorithms for infrastructure layout generation and pipe sizing in virtual and semi-virtual urban structures with limited input data. Therefore, the algorithm parameters represent 1) information of systems with common characteristics (e.g. cities with 100,000 inhabitants) or 2) more accurate data sets which are free available (e.g. open street map). The results are huge data sets of virtual/semi-virtual urban water infrastructure network models for the scientific community (Chapter 2). In this context the applications are diverse ranging from benchmarking and validating new algorithms (e.g. layout generation, pipe sizing, hydraulic and hydrologic simulations, optimization algorithms, calibration algorithms, parallel implementation of existing or new algorithms) to complex investigation with the purpose to explore and expand the knowledge of system interactions in urban structures from an integral viewpoint.

Where most of the developed algorithms and tools were initially intended to be used by the scientific community, also investigation on real world case studies could be carried out (e.g. Modelling the transition of the city movement of Kiruna (Chapter 16)). Here, the intention is to use the developed software tools as decision support tools for engineers, carriers and stakeholders.

Most of the work presented here must be assigned to the DynaVIBe project, which was funded by the Austrian Science Fund (FWF). It has been successfully completed and outstanding evaluated.

This thesis presented algorithms and software tools, proved their correctness and demonstrated their applicability in the context of urban water management.

It must be clearly stated that the software tools are not ready for daily use by target groups outside of the scientific community, but they successfully demonstrated their potential based on real world case studies. The range of potential target groups of the DynaVIBe approach can be increased by further developing the software tools, but also improve the methods and algorithms. Therefore, future research should focus on increasing the model diversity and decreasing software complexity at the same time but still keep the capability of investigating new scientific ground. This implies following topics:

Increase data availability: One of the initial aims of the DynaVIBe approach is to generate virtual case studies because of lacking real world case studies. Important is that the generation algorithms are controlled with a minimal number of parameters describing boundary properties of systems with equal behaviour (e.g. all water supply networks within the alpine region). To verify the correctness and strengthen the credibility of these algorithms it is essential to demonstrate and apply them on real world case studies. The cause of the lack of real case study data are 1) the low accuracy, misconception or even missing standardisation of digital data management (Sonnenberg et al., 2013) in the context of spatial and temporal data acquisition of urban structures and the marginal availability of data because of legal aspects. Both points are embraced in the INSPIRE directive (European Parliament, 2007). Although the technical infrastructure will be developed in this project, it is questionable to which extent data owners can be forced to share their data (What is the benefit for a water supply company to share detailed data of their infrastructure network with the public ?). However, from a scientific point of view, focus has to be on developing data management infrastructures in the context of urban water management modelling, which are capable to manage spatial and temporal data and verify them at time of observation and retention.

INCREASE MODEL DIVERSITY: The presented algorithms for the automatic generation of the case studies have to be developed further to capture all aspects and system components of modern urban water infrastructures. The approach of using surrogate data (e.g. street network data) for network generation implies that the developed methods are capable to use the information accordingly. However, the core algorithm is graph theory based and therefore independent from the used data source type. The important thing is that in this case the input data represents a network structure. It does not matter if this structure is derived from street networks or generated automatically with another algorithm (e.g. urban development algorithm). The main aim is to develop models which are capable to 'seamless' use various types of input data. This concept potentially enables the usage of all developed algorithms outside of the scientific community. This means the gap between virtual case studies and real case studies decreases and questions regarding model validity in correlation to needed data for modelling can be assessed. The publication 'Assessing Model Structure Uncertainties in Water Distribution Models' (Chapter 13) presents some initial work in this direction.

DECREASE SOFTWARE AND MODEL COMPLEXITY: To ensure usability of existing and future tools it is of high importance to simplify their handling by hiding computational aspects. The aim of such tools is to develop models in the context of urban water management for target groups which may not be skilled in programming and GIS environments. DynaMind models are data, computational and time intense simulations. Therefore, the approach to outsource simulations to dedicated servers next to the data sources via web-services should be pursued. In a broader sense this approach enables the usage of state-of-the-art cloud computing infrastructures (e.g. Google or Amazon cloud). Another point which strengthen this statement is the fact that the development of standardized data management infrastructure for spatial data seems to go in the direction of web-service based technologies as well (e.g. WMS and WFS).

Just because we can

However, more and more data get freely available and now we are able to 'flood' our models with huge data sets (The accuracy of the data sets may be questionable). In the same extent we get a flood of model simulation results. But, what is the additional benefit in a scientific context? Is it counteracting if we use highly detailed models for questions which could be answered with less detailed? On the other hand, from an engineer's perspective, these huge data sets are holding great potential in terms of applying the developed methods and algorithms on realistic case studies (e.g. cities with several million inhabitants).

5 BIBLIOGRAPHY

Alperovits, E. and Shamir, U. (1977). "Design of optimal water distribution systems." *Water resources research*, 13(6), 885–900.

Bach, P. M., Rauch, W., Mikkelsen, P. S., McCarthy, D. T., and Deletic, A. (2014). "A critical review of integrated urban water modelling – urban drainage and beyond." *Environmental Modelling & Software*, 54, 88 – 107.

Basupi, I. and Kapelan, Z. (2013). "Flexible water distribution system design under future demand uncertainty." *Journal of Water Resources Planning and Management*, 141(4), 04014067.

Benidickson, J. (2011). *The Culture of Flushing: A Social and Legal History of Sewage*. UBC Press.

Blumensaat, F., Wolfram, M., and Krebs, P. (2012). "Sewer model development under minimum data requirements." *Environmental Earth Sciences*, 65(5), 1427–1437.

Brown, R., Keath, N., and Wong, T. (2009). "Urban water management in cities: historical, current and future regimes." *Water Science & Technology*, 59(5), 847–855.

Brumbelow, K., Torres, J., Guikema, S., Bristow, E., and Kanta, L. (2007). "Virtual cities for water distribution and infrastructure system research." *World Environmental and Water Resources Congress*, 15–19.

Carstensen, J., Vanrolleghem, P., Rauch, W., and Reichert, P. (1997). "Terminology and methodology in modelling for water quality management—a discussion starter." *Water Science & Technology*, 36(5), 157–168.

De Corte, A. and Sörensen, K. (2014). "Hydrogen: an artificial water distribution network generator." *Water Resources Management*, 28(2), 333–350.

De Feo, G., Antoniou, G., Fardin, H. F., El-Gohary, F., Zheng, X. Y., Reklaityte, I., Butler, D., Yannopoulos, S., and Angelakis, A. N.

(2014). "The historical development of sewers worldwide." *Sustainability*, 6(6), 3936–3974.

European Parliament (2007). "Directive 2007/2/ec of the european parliament and of the council of 14 march 2007 establishing an infrastructure for spatial information in the european community (inspire)." *Official Journal of the European Union*.

Ferguson, B. C., Frantzeskaki, N., and Brown, R. R. (2013). "A strategic program for transitioning to a water sensitive city." *Landscape and Urban Planning*, 117, 32 – 45.

Haghighi, A. (2013). "Loop-by-loop cutting algorithm to generate layouts for urban drainage systems." *Journal of Water Resources Planning and Management*, 139(6), 693–703.

Huber, W. C., Rossman, L. A., and Dickinson, R. E. (2005). "Epa storm water management model, swmm5." *CRC Press: Boca Raton, FL*, 339–361.

Jolly, M., Lothes, A., Sebastian Bryson, L., and Ormsbee, L. (2014). "Research database of water distribution system models." *Journal of Water Resources Planning and Management*, 140(4), 410–416.

Makropoulos, C., Natsis, K., Liu, S., Mittas, K., and Butler, D. (2008). "Decision support for sustainable option selection in integrated urban water management." *Environmental Modelling & Software*, 23(12), 1448 – 1460.

Marlow, D. R., Moglia, M., Cook, S., and Beale, D. J. (2013). "Towards sustainable urban water management: A critical reassessment." *Water Research*, 47(20), 7150 – 7161.

Mitchell, V., Mein, R., and McMahon, T. (2001). "Modelling the urban water cycle." *Environmental Modelling & Software*, 16(7), 615 – 629.

Muranho, J., Ferreira, A., Sousa, J., Gomes, A., and Marques, A. S. (2012). "Waternetgen: an epanet extension for automatic water distribution network models generation and pipe sizing." *Water Science & Technology: Water Supply*, 12(1), 117–123.

Möderl, M., Fetz, T., and Rauch, W. (2007). "Stochastic approach for performance evaluation regarding water distribution systems." *Water Science & Technology*, 56(9), 29–36.

Möderl, M., Sitzenfrei, R., Fetz, T., Fleischhacker, E., and Rauch, W. (2011). "Systematic generation of virtual networks for water supply." *Water Resources Research*, 47(2).

Ostfeld, A., Salomons, E., Ormsbee, L., Uber, J. G., Bros, C. M., Kalungi, P., Burd, R., Zazula-Coetzee, B., Belrain, T., Kang, D., et al. (2011). "Battle of the water calibration networks." *Journal of Water Resources Planning and Management*, 138(5), 523–532.

Ostfeld, A., Uber, J. G., Salomons, E., Berry, J. W., Hart, W. E., Phillips, C. A., Watson, J.-P., Dorini, G., Jonkergouw, P., Kapelan, Z., et al. (2008). "The battle of the water sensor networks (bwsn): A design challenge for engineers and algorithms." *Journal of Water Resources Planning and Management*, 134(6), 556–568.

Pikaar, I., Sharma, K. R., Hu, S., Gernjak, W., Keller, J., and Yuan, Z. (2014). "Reducing sewer corrosion through integrated urban water management." *Science*, 345(6198), 812–814.

Rauch, W. and Kleidorfer, M. (2014). "Replace contamination, not the pipes." *Science*, 345(6198), 734–735.

Rossman, L. A. (1999). "The epanet programmer's toolkit for analysis of water distribution systems." *ASCE 29th Annual Water Resources Planning and Management Conference*, 39–48.

Rozos, E. and Makropoulos, C. (2013). "Source to tap urban water cycle modelling." *Environmental Modelling & Software*, 41, 139 – 150.

Saldarriaga, J., Takahashi, S., Hernández, F., and Escovar, M. (2011). "Predetermining pressure surfaces in water distribution system design." *Proceedings of the World Environmental and Water Resources Congress 2011, ASCE*.

Schaake, J. and Lai, D. (1969). "Linear programming and dynamic programming applications to water distribution network design. report 116, hydrodynamics laboratory, department of civil engineering.

Sitzenfrei, R., Fach, S., Kinzel, H., and Rauch, W. (2010a). "A multilayer cellular automata approach for algorithmic generation of virtual case studies: Vibe." *Water Science & Technology*, 61(1), 37–45.

Sitzenfrei, R., Fach, S., Kleidorfer, M., Urich, C., and Rauch, W. (2010b). "Dynamic virtual infrastructure benchmarking: Dynavibe." *Water Science and Technology: Water Supply*, 10(4), 600.

Sitzenfrei, R., Kleidorfer, M., Meister, M., Burger, G., Urich, C., Mair, M., and Rauch, W. (2014). "Scientific computing in urban water management." *Computational Engineering*, Springer International Publishing, 173–193.

Sitzenfrei, R., Mair, M., Tscheikner-Gratl, F., Hupfauf, B., and Rauch, W. (2015). "What can we learn from historical water network transition?." *World Environmental and Water Resources Congress 2015*, 907–916.

Sitzenfrei, R., Möderl, M., and Rauch, W. (2013). "Automatic generation of water distribution systems based on gis data." *Environmental Modelling & Software*, 47, 138 – 147.

Sonnenberg, H., Rustler, M., Riechel, M., Caradot, N., Rouault, P., and Matzinger, A. (2013). "Best data handling practices in water-related research." *Water Practice and Technology*, 8(3-4).

Trifunović, N., Maharjan, B., and Vairavamoorthy, K. (2012). "Spatial network generation tool for water distribution network design and performance analysis." *Water Science & Technology: Water Supply*, 13(1), 1–19.

Urich, C. and Rauch, W. (2014). "Exploring critical pathways for urban water management to identify robust strategies under deep uncertainties." *Water Research*, 66, 374 – 389.

Urich, C., Sitzenfrei, R., Möderl, M., and Rauch, W. (2010). "An agent-based approach for generating virtual sewer systems." *Water Science & Technology*, 62(5), 1090–1097.

Voinov, A. and Shugart, H. H. (2013). "'integronsters', integral and integrated modeling." *Environmental Modelling & Software*, 39, 149 – 158.

Walski, T., Brill, E., Jr., Gessler, J., Goulter, I., Jeppson, R., Lansey, K., Lee, H., Liebman, J., Mays, L., Morgan, D., , and Ormsbee, L. (1987). "Battle of the network models: Epilogue." *Journal of Water Resources Planning and Management*, 113(2), 191–203.

Walski, T. M., Chase, D. V., Savic, D. A., Grayman, W. M., Beckwith, S., Koelle, E., et al. (2003). *Advanced water distribution modeling and management*. Haestad press.

Willuweit, L. and O'Sullivan, J. J. (2013). "A decision support tool for sustainable planning of urban water systems: Presenting the dynamic urban water simulation model." *Water Research*, 47(20), 7206 – 7220.

Yazdani, A. and Jeffrey, P. (2011). "Applying network theory to quantify the redundancy and structural robustness of water distribution systems." *Journal of Water Resources Planning and Management*, 138(2), 153–161.

Part B

PUBLICATIONS

6

PAPER I

Where to find water pipes and sewers? – On the correlation of infrastructure networks in the urban environment

Authors: **Michael Mair**, Jonatan Zischg, Wolfgang Rauch and Robert Sitzenfrei

Date: submitted
Place: Computers, Environment and Urban Systems
Type: Journal

7

PAPER II

Spanning Tree Based Algorithm for Generating Water Distribution Network Sets by Using Street Network Data Sets

Authors: **Michael Mair**, Wolfgang Rauch and Robert Sitzenfrei

Date: 2014
Place: World Environmental and Water Resources Congress
Type: Proceedings paper

8

PAPER III

Improving Incomplete Water Distribution System Data

Authors: **Michael Mair**, Wolfgang Rauch and Robert Sitzenfrei
Date: 2014
Place: Procedia Engineering
Type: Journal

Available online at www.sciencedirect.com
ScienceDirect
Procedia Engineering 70 (2014) 1055 – 1062
Procedia Engineering
www.elsevier.com/locate/procedia

12th International Conference on Computing and Control for the Water Industry, CCWI2013

Improving incomplete water distribution system data

M. Mair[a]*, W. Rauch[a], R. Sitzenfrei[a]

[a] *Unit of environmental engineering/University Innsbruck, Technikerstr. 13, 6020 Innsbruck, Austria*

Abstract

Data for water distribution systems (WDS) are often not available or of poor quality. One promising approach for improving these is collecting data sets with strong coherences to the WDS and reconstructing possible WDS by using these data sets. An example for such a strongly correlated data set is the street network which can easily be accessed (e.g. open street map). The aim of this paper is to systematically analyze the impact of data improvement from alternative sources for creating WDS models. Investigations showed that hydraulic WDS models with a mean pressure error of three meters can be created by knowing 30% of pipes with a diameter ≥250 mm.

Keywords: Improving data; water supply system modeling; stochastic data reconstruction; street network data; hydraulic modeling

1. Introduction

For water distribution system analysis, data from case studies is crucial. But depending on the modeling aim and therefore the required accuracy of the model output, different levels of data quality are required (Hellbach et al., 2011). Contrary to practical applications where the aim is to describe a specific system sufficiently and very accurately, for research purposes it is often more important to gather information on many different water distribution systems with different characteristics in order to obtain case unspecific results from evaluations (Sitzenfrei et al., 2013). Therefore, recently researchers started to create few virtual water distribution systems

* Corresponding author. Tel.: +43-512-507-62129.
E-mail address: michael.mair@uibk.ac.at
The presentation of this paper, delivered by Robert Sitzenfrei, received the *CCWI 2013 Early Career Award*

doi:10.1016/j.proeng.2014.02.117

manually (Torres, 2006) or even automatically in great number (e.g. Möderl et al., 2007; Sitzenfrei et al., 2010b; Möderl et al., 2011; Trifunovic et al., 2012; Muranho et al., 2012). All these approaches are capable to create different kinds of water distribution systems. But a lack of all these approaches is that the generated network layouts are only a simplified representation of real systems. E.g. the approach presented in Möderl et al., 2011) has the limitation that only four pipes can be connected to one junction in the cardinal directions. This results in a rectangular grid of junctions in which all pipes have the same lengths. Likewise only reservoirs and junctions are regarded, therefore only gravity driven water supply can be represented in those models. Hence, more accurate approaches which mimic the layout characteristics of real world systems are required. More recently, researchers also investigated the evolvement of water distributions systems over time (e.g. Yazdani et al., 2011; Chang et al., 2012). For these research tasks like identifying and modeling optimal and robust future expansion strategies for water distribution systems the issue of data availability is even more challenging. One way to address the lack of data in this context is also to come back to data of virtual water distribution systems. Sitzenfrei et al., 2012) presented an approach for modeling dynamic expansion of water distribution systems for new urban developments. In that work, only virtual data was regarded because for calibration and validation of such an approach, historical data is required. But collecting historical network data for analyzing the dynamics of a network is a complex and time consuming task (Sitzenfrei et al., 2010a). In order to address the issue of simplified representation of real systems for the automatic generation of virtual water distribution systems and also to gain knowledge and sufficient data on the historical development of water distribution systems, new approaches are required. One promising approach for improving insufficient data sets is 1) collecting easy to access data sets with strong coherences to the water distribution system and 2) reconstructing possible water infrastructure data with a stochastic approach from these data sets. Sitzenfrei et al., 2013) presented an approach in which water distribution systems can be generated based on GIS data of elevation, population and housing densities. With that approach also different types of networks can be generated (e.g. looped or branched networks) and promising results were obtained for the pressure distribution in the investigated supply area. For sewer model creation, Blumensaat et al., 2012) presented an approach for generating possible representations of sewer models (Rossman, 2004) which is based on the street layout. In Mair et al., 2012) coherences in capacity, design and layout of water infrastructures are analyzed and described by comparing street, water supply and sewer network data. According to this study, there is a strong coherence for the layout between street network and water distribution network.

The aim of this paper is to analyze the impact of data improvement from other available data sources (e.g. street network data, population density) for creating water distribution models. Different data sets with different quality are compared with regard to the hydraulic performance, layout and asset costs between model results provided by a) improved incomplete data (e.g. only poor knowledge on spatial layout of the water distribution system improved with street network data) and b) the complete water distribution network model. By systematically altering data sets of incomplete water distribution network and water demand data and with a successive comparison of the improved incomplete network model with the complete network model, the impact on e.g. hydraulic performance can be quantified. All results presented in this paper are based on a detailed case study in the Alps with approximately 120 thousand inhabitants.

Nomenclature

DEM	(m) Digital elevation map
DN	(mm) diameter
J	(-) Set of supply junctions in a water distribution model
#J	(-) number of elements in a set J
p	(m) hydraulic pressure
PI	performance indicator
PI_1	(-) Performance indicator for the relative pressure difference between models
PI_2	(m) Performance indicator for mean absolute pressure difference between models
Q	(m^3/s) water demand

2. Material and methods

The principal idea of this work is use easy to access data like street layout, elevation, and population densities to create semi-virtual hydraulic models. The presented generation procedure is repeated by using different amount of the available data for creation of the semi-virtual hydraulic models (Rossman, 2000). In a next step, the results of the hydraulic modeling of each created model are compared with traditional assembled and calibrated hydraulic water supply models (which is assumed to be the exact representation of the system).

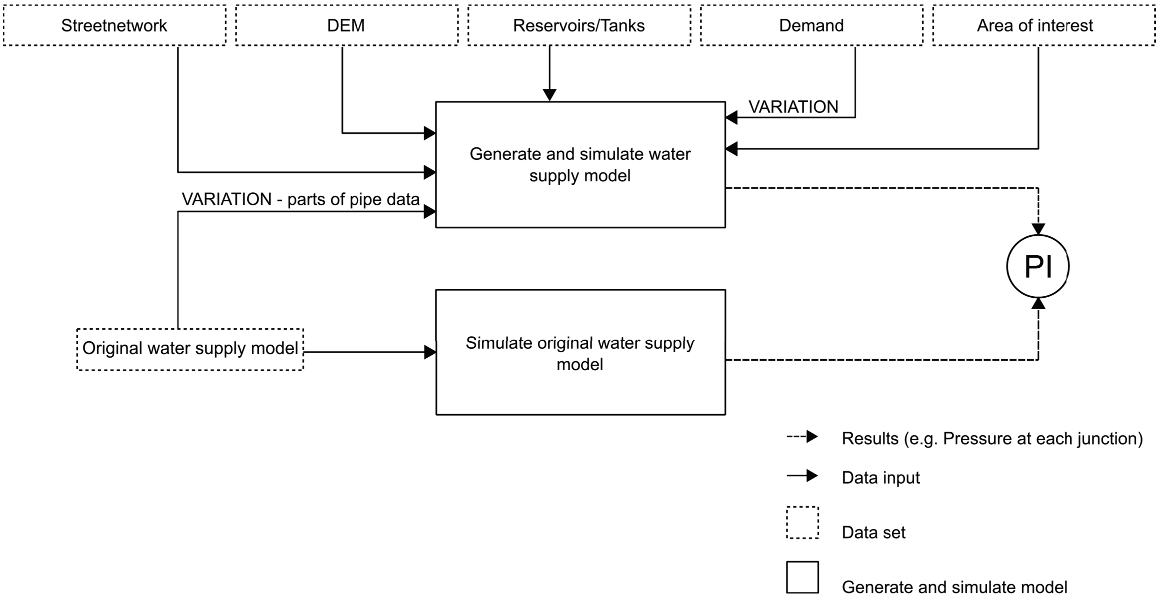

Fig. 1. Workflow setup for benchmarking artificial generated water supply models.

For that, a benchmark system was developed to automatically generate water supply models (EPANET2 Rossman, 2000) and evaluating their performance according to two different performance indicators (PI). The whole benchmark system was set up in a scientific workflow engine called DynaMind (Urich et al., 2012). Fig. 1 shows the general concept of the benchmark system. In the context of this work we have identified five different data sets (Fig.1 top row boxes), which are at least needed to automatically generate and simulate hydraulic models. The accuracy of the automatic generated model is determined by comparing simulation result of the generated model with simulation results from an already existing exact model. To determine the impact of additional data (e.g. location, diameter, roughness of real system pipes or demand with spatial resolution) the benchmark simulations are repeated by successively increasing the amount of additional data derived from the exact model.

2.1. Water supply network layout and pipe sizing

In Mair et al., 2012) it was demonstrated that up to 78 percent of a street network are containing up to 81 percent of the total length of a water supply network. According to that fact it is obvious to use street network data during the modeling process because of its easy accessibility (e.g. open street map, Google maps). This data is used to determine the layout of a water supply network by means of graph theory, which is the starting point for automatically generating a water supply model. The basic steps of the algorithm to generate a water supply network model in this work are:

- Trim the street network data set according to the area of interest
- Join additional pipe data set with the street data set
- Generate a minimum spanning tree out of the joined data set (Kruskal, 1956)
- Distribute demand points over the spanning tree
- Remove all leafs which have no demand points (A leaf is a vertex with only one connected edge)
- Generate loops
- Connect all reservoirs and/or tanks to the looped graph
- Pipe sizing of the generated graph based on a simple pipe sizing algorithm (Sitzenfrei et al., 2013)

A detailed description of the graph theoretical generation procedure can be obtained in a future publication. The output of this algorithm is an EPANET2 model which can be simulated and compared with the original EPANET2 model within the area of interest.

2.2. Performance indicators – PI

For evaluating the accuracy of the generated water supply models two different performance indicators (Eq.(1) and Eq.(2)) are used. Both of them are comparing the pressure at different locations between the generated and original EPANET2 model. Table 1 shows the exact meaning and definition of each variable and function occurring in both equations (Eq.(1) and Eq.(2)).

Table 1. Definition of variables and functions within both performance indicators.

Function or variable name	Set of all junctions with a demand greater zero in the original water supply model
J	Number of elements in the set of junctions
#*J*	Pressure at junction *i* within the original model
p_i	Pressure at coordinate x and y of p_i (pressure within original model) projected in the area of interest of the generated model. This projection is network layout unspecific and only depending on the real location within the area of interest.
proj(p_i)	Set of all junctions with a demand greater zero in the original water supply model

Eq.(1) (PI_1) is an indicator for the relative pressure difference between both models. Values smaller than one indicate the prediction of too low pressure values in the generated model. Values greater than one are indicating an over estimate of pressure. In contrast to that Eq.(2) (PI_2) indicates the mean absolute pressure difference.

$$PI\_1 = \frac{\sum_{0}^{i=\#J} \frac{\mathrm{proj}(p_i)}{p_i}}{\#J} \text{ [-]} \tag{1}$$

$$PI\_2 = \sqrt{\frac{\sum_{0}^{i=\#J} (\mathrm{proj}(p_i) - p_i)^2}{\#J}} \text{ [m]} \tag{2}$$

2.3. Case study – Innsbruck

The presented method is applied to the case study Innsbruck. It is a city in the Alps with 120 thousands inhabitants. For the area of interest a significant amount of data is available. In particular, a street network data set extracted from open street map, a digital elevation map in a resolution of 10x10 meter, the total demand Q of the water supply network with 0.7 m^3/s (detailed spatial resolution, uniform distributed and normal distributed over the area of interest) and the position of all reservoirs. These data sets (Initial data set) are used as input for the model generation algorithm, which generates an EPANET2 model. The pressure distribution of that model is shown in Fig. 2.

Fig. 2. Pressure distribution within the water supply network of the case study Innsbruck.

2.4. Generation and testing procedure

The generated models are compared with an already existing and well tested EPANET2 model of the case study containing detailed information of all components (e.g. pipes, demand with spatial resolution, etc.). The network generation algorithm is tested by appending additional pipe information (location, diameter, roughness) of the detailed EPANET2 model to the initial data set. In detail, five different sets of additional pipe information are appended to the initial data set, which are all pipes of the original EPANET2 model with a diameter greater or equal than 500mm, 250mm, 150mm or 0mm, respectively. Fig. 3. shows the fraction of the total network length by choosing all pipes greater or equal a certain diameter (DN). For example the total length of all pipes with a diameter greater or equal 250mm describes 30 percent of the overall system. In the generation procedure, the rest of the pipe network is created by means of the street network and graph analysis.

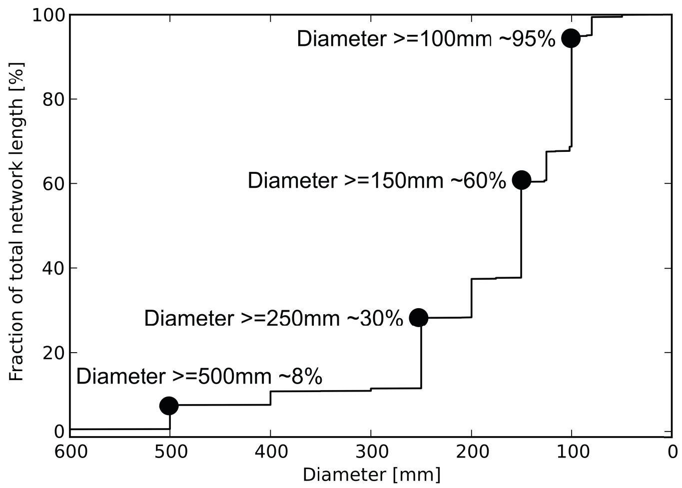

Fig. 3. Fraction of total network length of pipes equal or greater than a certain diameter.

3. Results

Fig. 4 shows the result for both performance indicators (PI_1 – left and PI_2 - right) by using the developed model generation algorithm. The x-axes for both plots are describing additional pipe information data sets which are appended to the initial data set (see also Fig. 4). The initial data set is defined by street network data set, DEM, Reservoirs/Tanks, Area of interest plus a variation of the demand data set: normal distributed, uniform distributed and exact demand.

By using the initial data set combined with additional pipe information with a diameter greater or equal to 500mm (eight percent of the total system) the generated water supply model is predicting too low pressure heads - between five and ten percent compared to the detailed model (Fig. 4. – PI_1 (>500)). This is an absolute pressure difference of between ten and seven meters (Fig. 4 – PI_2 (>500)). Compared to automatic generated models by using all pipes with a diameter greater or equal to 250mm (30 percent of total system) performance according to PI_1 is near to one, which is a pressure difference of between two and three meters according to PI_2 (Fig. 4 – PI_2 (>250)). By using all pipes of the original EPANET2 model as additional information nearly same result can be observed as compared to the pipe data set with diameters greater or equal to 250mm. The reason in this case for PI_1 not is being equal to exactly one and PI_2 not being equal to exactly zero is that the model generation algorithm uses a digital elevation map as input of a resolution of 10x10m for mapping the elevation to each junctions within the system. A fine grained DEM (e.g. 1x1 meter) would result in a lower PI_2 value.

Both diagrams are containing three different curves which is the variation of using exact distributed (original demand), uniform distributed and normal distributed demand within the initial data set. Both plots show that it is not necessary to know the exact position of all demand points within a system. For the scope and aim of this work models generated with uniform distribution of the total demand (0.7 m^3/s) within the area of interest have the same performance as knowing the real demand points of the system.

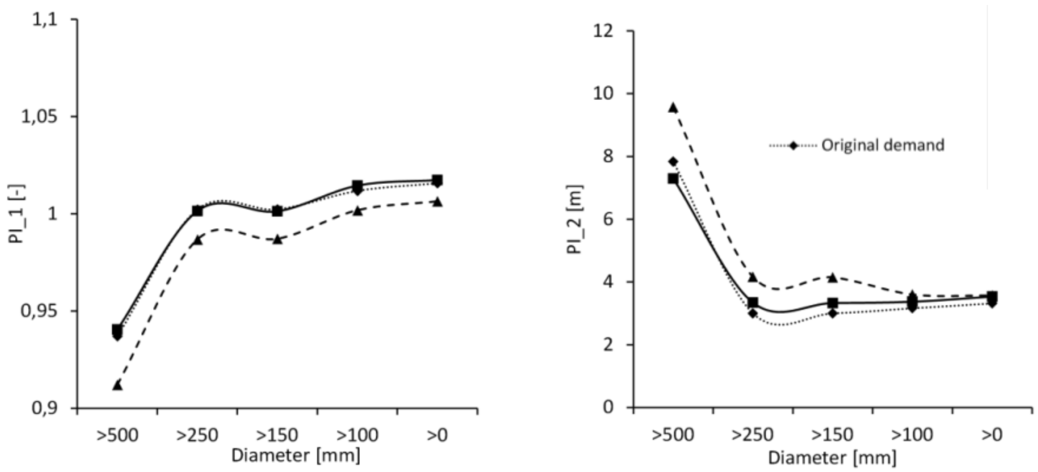

Fig. 4. Data amount for the case study Innsbruck.

4. Conclusion

In this manuscript the minimum data requirements are investigated which are required for developing a valid water supply model with the aim of generating the network layout, automatically determining pipe sizes and simulating the pressure within the system. The presented model generation algorithm uses easy accessible data sets (e.g. street network – Open Street network data) as input, combined with additional well known data of the real water supply network (e.g. data set of water supply pipes with a diameter greater or equal to a certain diameter). The simulation results of automatically generated water supply models are compared with simulation results of an exact model of the case study by using two different performance indicators. Both of them are indicating the mean pressure difference between the generated and detailed model, where the first one is a relative performance indicator and the second an absolute indicator. The whole benchmark setup was applied to the case study Innsbruck, a city within the Alps with 120 thousand inhabitants. Applying the model generation algorithm on a digital elevation map (resolution 10x10 m), reservoirs/tanks, street network, uniform distributing demand of 0.7 m^3/s and area of interest data sets plus all pipes with a diameter equal or greater than 250mm (30 percent of total length of all pipes within the real system) as additional information results in a relative performance value (PI_1) near to one and a mean absolute pressure difference of approximately three meters (PI_2).

Both performance indicators show only little difference if using exact spatial information of demand nodes as compared to uniform distributing the total demand within the area of interest for modeling the water supply system.

Acknowledgements

This work was funded by the Austrian Science Fund (**FWF**) in the project **DynaVIBe P23250-N24**. The authors gratefully acknowledge the financial support.

References

Blumensaat, F.; Wolfram, M. and Krebs, P. (2012): Sewer model development under minimum data requirements. In: Environmental Earth Sciences, Vol. 65(5), p. 1427-1437.

Chang, N.-B.; Qi, C. and Yang, Y.J. (2012): Optimal expansion of a drinking water infrastructure system with respect to carbon footprint, cost-effectiveness and water demand. In: Journal of Environmental Management, Vol. 110(0), p. 194-206.

Hellbach, C.; Moderl, M.; Sitzenfrei, R. and Rauch, W. (2011): Influence of Network Properties and Model Purpose on the Level of Skeletonization. In: Proceedings of the World Environmental & Water Resources Congress, Palm Springs, California, USA, May 22 - 26, 2011, Vol. 414, p. 10.

Kruskal, J.B. (1956): On the shortest spanning subtree of a graph and the traveling salesman problem. In: Proceedings of the American Mathematical society, Vol. 7(1), p. 48-50.

Mair, M.; Sitzenfrei, R.; Möderl, M. and Rauch, W. (2012): Identifying multi utility network similarities. In: Proceedings of the World Environmental and Water Resources Congress, Albuquerque, New Mexico, 20.05.2012-24.05.2012, p. 3147-3153.

Möderl, M.; Fetz, T. and Rauch, W. (2007): Stochastic approach for performance evaluation regarding water distribution systems. In: Water Science and Technology, Vol. 56(9), p. 29-36.

Möderl, M.; Sitzenfrei, R.; Fetz, T.; Fleischhacker, E. and Rauch, W. (2011): Systematic generation of virtual networks for water supply. In: Water Resources Research, Vol. 47.

Muranho, J.; Ferreira, A.; Sousa, J.; Gomes, A. and Marques, A.S. (2012): WaterNetGen: an EPANET extension for automatic water distribution network models generation and pipe sizing. In: Water Science & Technology: Water Supply, Vol. 12(1), p. 117-123.

Rossman, L.A. (2000): EPANET 2 user manual. National Risk Management Research Laboratory - U.S. Environmental Protection Agency.

Rossman, L.A. (2004): Storm water management model - User's manual Version 5.0. National Risk Management Research Laboratory - U.S. Environmental Protection Agency.

Sitzenfrei, R.; Fach, S.; Kleidorfer, M.; Urich, C. and Rauch, W. (2010a): Dynamic virtual infrastructure benchmarking: DynaVIBe. In: Water Science and Technology: Water Supply, Vol. 10(4), p. 600.

Sitzenfrei, R.; Moderl, M.; Mair, M. and Rauch, W. (2012): Modeling Dynamic Expansion of Water Distribution Systems for New Urban Developments. In: Proceedings of the World Environmental & Water Resources Congress, Albuquerque, New Mexico, USA, May 20 - 24, 2012.

Sitzenfrei, R.; Möderl, M. and Rauch, W. (2010b): Graph-based approach for generating virtual water distribution systems in the software ViBe. In: Water Science and Technology: Water Supply, Vol. 10(6), p. 923-932.

Sitzenfrei, R.; Möderl, M. and Rauch, W. (2013): Automatic generation of water distribution systems based on GIS data. In: Environmental Modelling & Software, Vol. 47, p. 138-147.

Torres, J. (2006): Micropolis: A Virtual City for Water Distribution System Research Applications. Texas A&M University, Zachry Department of Civil Engineering.

Trifunovic, N.; Maharhan, B. and Vairavamoorthy, K. (2012): Spatial network generation tool for water distribution network design and performance analysis. In: Water Science and Technology: Water Supply, Vol. 13(1).

Urich, C.; Burger, G.; Mair, M. and Rauch, W. (2012): DynaMind–A Softwaretool for Integrated Modelling of Urban Environments and their Infrastructure. In: Proceedings of the 10th International Conference on Hydroinformatics, Vol. 14, p. 18.

Yazdani, A.; Otoo, R.A. and Jeffrey, P. (2011): Resilience enhancing expansion strategies for water distribution systems: A network theory approach. In: Environmental Modelling & Software, Vol. 26(12), p. 1574-1582.

9

PAPER IV

GIS-based applications of sensitivity analysis for sewer models

Authors: **Michael Mair**, Robert Sitzenfrei, Manfred Kleidorfer, Michael Möderl and Wolfgang Rauch

Date: 2012
Place: Water Science & Technology
Type: Journal

10

PAPER V

GIS Based Applications of Sensitivity Analysis for Water Distribution Models

Authors: Michael Möderl, Christian Hellbach, Robert Sitzenfrei, **Michael Mair**, Lukas Aditya Alexander, Ernest Mayr, Reinhard Perfler and Wolfgang Rauch

Date: 2011

Place: World Environmental and Water Resources Congress

Type: Proceedings paper

11

PAPER VI

Cascade vulnerability for risk analysis of water infrastructure

Authors: Robert Sitzenfrei, **Michael Mair**, Michael Möderl and Wolfgang Rauch

Date: 2011
Place: Water Science & Technology
Type: Journal

12

PAPER VII

Identifying Hydropower Potential in Water Distribution Systems of Alpine Regions

Authors: Michael Möderl, Robert Sitzenfrei, **Michael Mair**, Hannes Jarosch and Wolfgang Rauch

Date: 2012

Place: World Environmental and Water Resources Congress

Type: Proceedings paper

13

PAPER VIII

Assessing Model Structure Uncertainties in Water Distribution Models

Authors: Robert Sitzenfrei, **Michael Mair**, Kegong Diao and Wolfgang Rauch
Date: 2014
Place: World Environmental and Water Resources Congress
Type: Proceedings paper

14

PAPER IX

Modeling Dynamic Expansion of Distribution Systems for New Urban Developments

Authors: Robert Sitzenfrei, Michael Möderl, **Michael Mair** and Wolfgang Rauch

Date: 2012

Place: World Environmental and Water Resources Congress

Type: Proceedings paper

15

PAPER X

Stability of Traditional Urban Water Systems — Integrated Assessment of Transitions Scenarios

Authors: Robert Sitzenfrei, **Michael Mair** and Wolfgang Rauch
Date: 2014
Place: Procedia Engineering
Type: Journal

Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 89 (2014) 727 – 733

Procedia Engineering

www.elsevier.com/locate/procedia

16th Conference on Water Distribution System Analysis, WDSA 2014

Stability of Traditional Urban Water Systems – Integrated Assessment of Transitions Scenarios

R. Sitzenfrei[a,]*, M. Mair[a], W. Rauch[a]

[a]*Unit of Environmental Engineering, Department of Engineering Science, University Innsbruck, Technikerstrasse 13, Innsbruck, Austria*

Abstract

In this work it is quantified to what extent existing water networks can cope with urban and population change developments before they fail. For that, the VIBe (Virtual Infrastructure Benchmarking) approach is used in order to create numerous city scale test cases (including urban form, water supply and drainage system) with different characteristics. Based on 81 test cases, probabilities of failure for different system characteristics are determined. For such an integrated assessment, a key issue is the weighting of the different technical performances when identifying stable ranges for the overall performance. Therefore, different weighting strategies are tested and compared.

Peer-review under responsibility of the Organizing Committee of WDSA 2014

Keywords: Integrated water system, coupled sewer and water supply modelling, Virtual Infrastructure Benchmarking - VIBe

1. Introduction

Industrialized cities typically rely on traditional urban water infrastructure. Due to the long life span of such systems and the great asset value, projections for several decades into the future have to be taken into account during the planning process. If the actual city growth still excesses the expectations, pressure deficits in supply, flooding and environmental problems can occur. In addition, if the expected urban development does not take place, problems arise as e.g. deposit problems in the sewer system or water quality issues in the water supply system. A key issue is to find flexible and sustainable solutions for urban water systems [1] which can operate on a broad range of boundary conditions respectively can easily be adapted to new ones.

* Corresponding author. Tel.: +43-512-507-62195; fax: +43-512-507-2911.
E-mail address: robert.sitzenfrei@uibk.ac.at.

Peer-review under responsibility of the Organizing Committee of WDSA 2014
doi:10.1016/j.proeng.2014.11.500

A combination of centralized and decentralized urban drainage facilities provides such a flexibility [2]. This is an interlinked problem [3] as e.g. usually most of the water consumption is directly discharged in the sewer system [4]. Therefore, in this work an integrated, coupled assessment of centralized water infrastructure is applied to investigate the stability of the existing traditional water networks [5]. Usually, such case study investigations are based on a single or few cases and therefore only very case specific results can be obtained from that. In this work, the VIBe (Virtual Infrastructure Benchmarking [6]) approach is used in order to create numerous city scale test cases (including urban form, water supply and drainage system) with different characteristics (e.g. city sizes). Based on that, statistical evaluations can be performed for the described investigations and probabilities of failure for different system characteristics can be obtained [7]. In detail, for numerous different sized test cases it is quantified to what extent an existing water networks can cope with urban and population developments but also with a transition to decentralized measures before they fail (insufficient technical performance).

The investigations show that the systems can operate sufficiently at a broad range of population growth and decline scenarios respectively, also when successively installing decentralized systems for e.g. water supply. However, a reduction further than to 40% and increase beyond about 40% pushes the traditional water systems to their limit.

2. Material and Methods

2.1. Virtual Infrastructure Benchmarking

One key method of this study is the VIBe approach [5]. With that approach, numerous city-scale test cases with different characteristics can be created automatically for case study analysis. These virtual test cases include all relevant data for the urban structure (e.g. topography, land use, population densities, impervious area, etc.) but also data for hydraulic modelling of water supply networks and urban drainage networks (Epanet2 and SWMM5 files).

The modelling procedure of the VIBe approach considers as input parameter ranges and characteristics of real world case studies. These parameter ranges can be further extended by a broader bandwidth of physically feasible parameter ranges. E.g. for the topography, different types of possible slopes can be tested or also different extents of the city. Subsequent, based on these input parameters, possible city layout are automatically generated. With the GIS-data of the city layouts as input [8], hydraulic water infrastructure models are automatically created based on design guidelines [6, 9]. With these numerous case studies, investigations and evaluations can be performed. The advantage of applying numerous test cases is that the obtained results are not case specific as usual when evaluating a single benchmark model or only a few test cases for an investigation. Further, the results can be evaluated statistically and investigated according to different characteristics (e.g. probabilities of failures depending on the size of cities).

2.2. Systematic Scenario Investigations

In this study, 81 test cases (80 virtual and 1 real world case studies) are used. The populations of the case studies vary between 70,000 and 170,000 inhabitants. For the design of the urban drainage and the water supply model, state-of-the-art technical guidelines are applied. The drainage and the supply network models are linked spatially referenced via the population. A detailed description of the generation procedure and characteristics of the test case set can be obtained in [5].

For each of the 81 test cases, reduction scenarios and increase scenarios are systematically investigated. The reduction scenarios consider a reduction in water demand respectively dry weather flow production up to -90% (in 5 steps) and the increase scenarios consider population growth (up to +100% in 5 steps) respectively with that connected increase of impervious areas. With the initial values this results in 11 change scenarios for each test case which results in total in 891 investigated city scale scenarios. These change scenarios are further referred to as different *variation factors.* A detailed description of the different scenarios and the scenario generation can be obtain in [5].

2.3. Performance Assessment

For performance assessment, three normalized performance indicators (0 indicates insufficient performance and 1 excellent performance) for urban drainage systems are used. For water supply, 2 normalized performance indicators were used (see Table 1).

One approach to assess the overall performance of a system with the help of several performance indicators can be done by weighting each individual indicator and summing up all resulting products to one system indicator. A state-of-the-art technique to systematically combine several indicators based on the described approach is the Analytic Hierarchy Process [10] by creating a prioritization of all indicators (comparing each pair of the system indicators) and finally assigning weighting values to each indicator (sum of all weights is equal to one). One shortcoming here is that the weighting values are assumptions made by experts (e.g. scientist and stakeholders) and therefore the impact of each performance indicator of a system on the overall performance indicator underlies decisions of human beings. The weighting values may not be correct or even can intentionally be chosen according to a certain aim to let the performance of system look better than it is in reality.

Table 1. Different used performance indicators.

Name of normalized performance indicator	description of performance indicator
UD CSO	urban drainage emissions assessment
UD flooding	urban drainage flooding assessment
UD shear stress:	urban drainage shear stress assessment
WDS hydraulic	pressure performance in the water distribution network
WDS quality	water ages assessment in the water distribution network

For an integrated assessment, the described five performance indicators are used as basis. Therefore, it is investigated how different weightings of the performance indicators impact the general conclusion. In total, four different weighting scenarios are tested. The sum of weights of the scenarios is kept constant. The first scenarios (I) assumes that all five performance indicators are of same interest and therefore equally weighted. In the second scenario (II), a focus on the hydraulic performance is assumed. Therefore, the *UD flooding* and *WDS hydraulic* are weighted more importantly than the others. As scenario (III), the focus is on operation of the networks. Therefore, the *UD shear stress* and *WDS quality* are weighted more importantly. The different used weightings are summarized in Table 2.

The aim of this study is to investigate how the conclusion on the overall performance is impacted by different weightings and not to quantify the overall performance. This means that the different scenarios are compared among each other. Therefore, as a fourth scenario (IV), the five normalized performance indicators are multiplied with each other, making a weighting of the different indicators obsolete.

Table 2. Different weightings of performance indicators for integrated performance

Name of performance indicator	scenario I (Fig 2a - c)	scenario II (Fig 2d - e)	scenario III (Fig. 2f – h)	scenario IV (Fig. 3 a –c)
UD CSO	1	1	0.5	(-)
UD flooding	1	1.5	1	(-)
UD shear stress:	1	0.5	1.5	(-)
WDS hydraulic	1	0.5	0.5	(-)
WDS quality	1	1.5	1.5	(-)

All investigations are composed for three different classes of city sizes. In class 1 (population of 70,000 – 100,000), there are 25% of the test cases. In class 2 (population between 100,000 and 140,000) there are 50% of the test cases and in class 3 (population between 140,000 and 170,000) there are again 25% of the test cases.

3. Results and Discussion

In Fig. 1, the raw data [5] for the 5 different performance indicators and the 81 test cases are plotted. Additional, the results for the three different size-classes are shown (class 1 - 3). The boxplots are the results obtained by the 80 automatically generated test cases and the red circulars show the results for the used real world case study (population of 121,000). On the bottom of each plot (respectively as annotation for each boxplot), the different *variation factors* for growth and decline scenarios are listed.

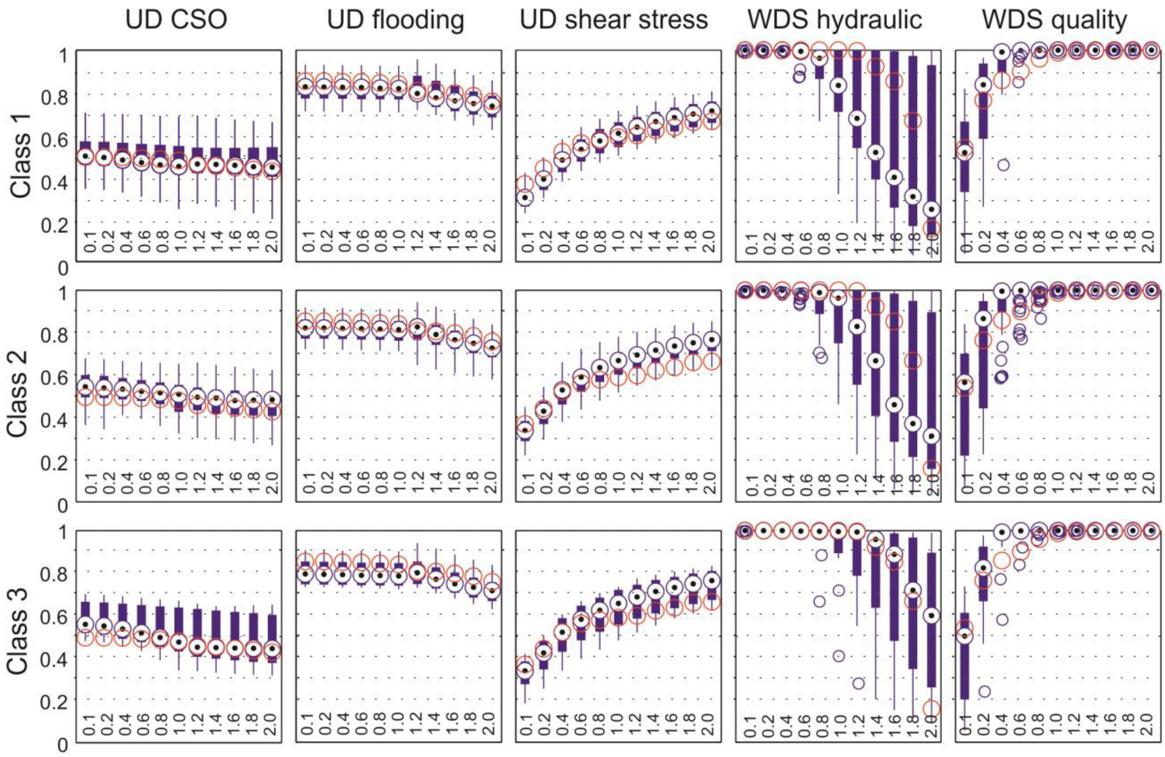

Fig. 1. Integrated results (reprinted from [5])

In Fig. 2, the results for the integrated performance for the weightings scenarios are shown. In the first row (a) – (c) the results for the weighting scenario (I) are shown for the different population classes. Likewise, in the second row for the weighting scenario (II) and in the third row of Fig. 2 for the weighting scenario (III).

Due to the used large set of test cases the analyses for the integrated performance can be done statistically. Therefore, for the different reduction and increase scenarios (*variation factors*), the 25%, 50% (red line), 75% and the 5%, 95% percentiles can be calculated. The black dashed line shows the results obtained with the single real case study. In [5] stable ranges for that single case study between 0.4 and 1.6 were assessed.

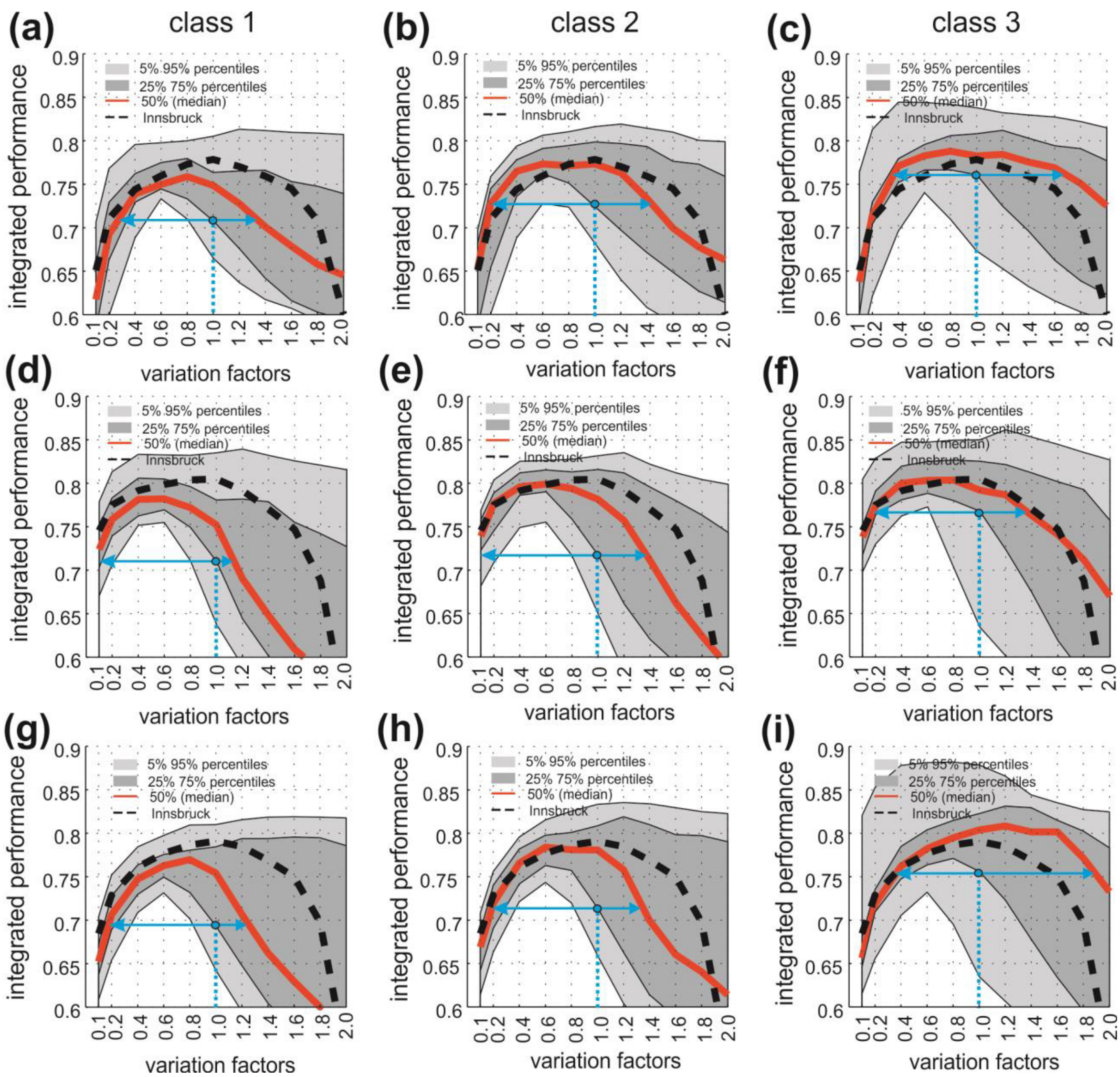

Fig. 2. Integrated results with different parameter weightings

To compare and systematically analyze these different statistical distributions among each other, a special metric is required. For that the 25 percentile for the initial size variation (*variation factor* = 1.0) is used as start value. In the generation, the systems are designed based on this design load. Therefore, the systems are expected to perform sufficient for that load. Subsequent, for that obtained integrated performance (*variation factor* = 1.0 and 25% percentile); the value range above this value for the medians (red line) is determined and denoted as "*stable range*". This evaluation process is visualized in blue color in Fig. 2 and Fig.3 and the obtained results are summarized in Table 3.

From Fig.2 it can be observed that systems in class 1 operate stabile on a more narrow range than the systems in class 2 and 3. Also the overall performance in class 3 is higher than in the others. Systems in class 1 and 3 only perform better compare to class 3, when intensely reducing the population.

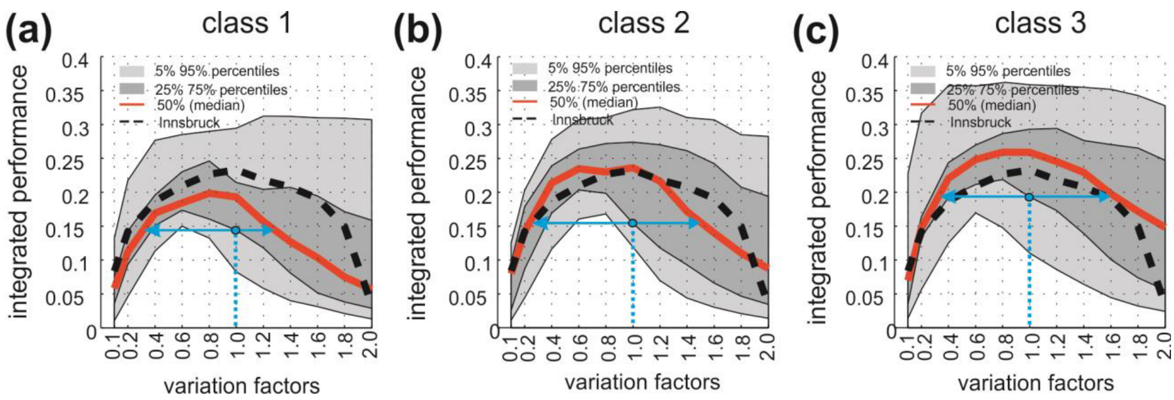

Fig. 3. Integrated results with multiplication (reprinted from [5])

In Fig. 3 the integrated results when multiplying all performance indicators are shown. The overall value of the integrated performance indicator is reduced compared to Fig. 2. The aim of the study is only to compare the systems for different variation factors among each other. Therefore, the quantitative value is in this context of no interest.

Table 3. Ranges for integrated performance for different weighting scenarios

Name of performance indicator	scenario I (Fig 2a - c)	scenario II (Fig 2d - e)	scenario III (Fig. 2f – h)	scenario IV (Fig. 3 a –c)
Integrated performance stable range class 1	0.3 – 1.3	0.1 – 1.1	0.2 – 1.3	0.3 – 1.3
Integrated performance stable range class 2	0.2 – 1.5	0.1 – 1.4	0.2 – 1.3	0.3 – 1.5
Integrated performance stable range class 3	0.3 – 1.7	0.2 – 1.4	0.3 – 1.9	0.3 – 1.6

Comparing the results summarized in Table 3 shows that the systems can operate sufficiently at a broad range of *variation factors* (i.e. population growth and decline scenarios respectively also when successively installing decentralized systems for e.g. water supply). But a reduction further than to 30% pushed the traditional water systems to their limit regardless of the size (i.e. classes). The values for the different weighting scenarios differ marginally but especially for the maximum values there are differences. In this study due to the definition of performance indicators usually no zero values occur. Further, the sensitivities of the single performance indicators due to the variation factors are all monotonically in- or decreasing. Therefore, multiplying the different single performance indicators nevertheless gives independent from weightings an objective picture of the behavior.

4. Conclusions

In this work, the VIBe (Virtual Infrastructure Benchmarking) approach is used in order to create numerous city scale test cases (including urban form, water supply and drainage system) with different characteristics. Based on that, statistical evaluations can be performed for investigations and probabilities of failure for different system characteristics can be obtained. In detail, for numerous different sized test cases it is quantified to what extent existing water networks can cope with urban and population change developments before they fail (insufficient technical performance). For an integrated assessment, a key issue is how to aggregate the different technical performances to single integrated performance. Therefore, different weighting scenarios are tested to assess the robustness of the obtained results.

The investigations show that independent from the weightings, the systems can operate sufficiently at a broad range of population growth and decline scenarios respectively also when successively installing decentralized systems for e.g. water supply. But a reduction further than to 40% and increase beyond about 40% pushed the traditional water systems to their limit. The values for the different weighting scenarios differ marginally. As in this study due to the definition of performance indicators usually no zero values occur, multiplying the different single performance indicators nevertheless gives independent from weightings an objective picture of the behavior.

Acknowledgements

This work was funded by the Austrian Science Fund (**FWF**) in the project **DynaVIBe P23250**. The authors gratefully acknowledge the financial support. With permission from Elsevier, fig. 1 and 3 were reprinted (License for [5]: http://creativecommons.org/licenses/by-nc-nd/3.0/)

References

[1] D. R. Marlow, M. Moglia, S. Cook, D. J. Beale, Towards sustainable urban water management: A critical reassessment, Water Research, 47 (2013) 7150-7161.
[2] T. Siekmann, M. Siekmann, Resilient urban drainage – Options of an optimized area-management Special Issue: Towards more flood resilient cities, Urban Water Journal, 8 (2013) 1-8.
[3] P. M. Bach, W. Rauch, P. S. Mikkelsen, D. T. McCarthy, A. Deletic, A critical review of integrated urban water modelling – Urban drainage and beyond," Environmental Modelling & Software, 54 (2014) 88-107.
[4] R. Sitzenfrei, W. Rauch, Investigating Transitions of Centralized Water Infrastructure to Decentralized Solutions – An Integrated Approach, Procedia Engineering, 70 (2014) 1549-1557.
[5] R. Sitzenfrei, M. Möderl, W. Rauch, Assessing the impact of transitions from centralised to decentralised water solutions on existing infrastructures – Integrated city-scale analysis with VIBe, Water Research, 47 (2013) 7251-7263.
[6] R. Sitzenfrei, Stochastic Generation of Urban Water Systems for Case Study Analysis, PhD-thesis (free online available at http://www.uibk.ac.at/umwelttechnik/teaching/phd/diss_sitzenfrei.pdf), Unit of Environmental Engineering, University of Innsbruck, Innsbruck, 2010.
[7] R. Sitzenfrei, S. Fach, H. Kinzel, W. Rauch, A multi-layer cellular automata approach for algorithmic generation of virtual case studies - VIBe, Water Science and Technology, 61 (2010) 37 - 45.
[8] R. Sitzenfrei, M. Möderl, W. Rauch, Automatic generation of water distribution systems based on GIS data, Environmental Modelling & Software, 47 (2013) 138-147.
[9] C. Urich, R. Sitzenfrei, M. Moderl, W. Rauch, An agent-based approach for generating virtual sewer systems, Water Sci Technol, 62 (2010) 1090-7
[10] T. L. Saaty, How to make a decision: the analytic hierarchy process, European journal of operational research, 48 (1990) 9-26.

16

PAPER XI

Stochastic performance assessment and optimization strategies of the water supply network transition of Kiruna during city relocation

Authors: Jonatan Zischg, **Michael Mair**,
Robert Sitzenfrei and Wolfgang Rauch

Date: 2015
Place: World Environmental and Water Resources Congress
Type: Proceedings paper

17

PAPER XII

DYNAMIND - A Softwaretool for Integrated Modelling of Urban Environments and their Infrastructure

`Authors:` Christian Urich, Gregor Burger, **Michael Mair** and Wolfgang Rauch
`Date:` 2012
`Place:` International Conference on hydroinformatics
`Type:` Proceedings paper

10th International Conference on Hydroinformatics
HIC 2012, Hamburg, GERMANY

DYNAMIND - A SOFTWARE TOOL FOR INTEGRATED MODELLING OF URBAN ENVIRONMENTS AND THEIR INFRASTRUCTURE

URICH CHRISTIAN(1), GREGOR BURGER(1), MICHAEL MAIR(1), WOLFGANG RAUCH (1)
(1): Unit of Environmental Engineering, University of Innsbruck, Technikerstr. 13, Innsbruck, 6020, Austria

The DynaMind Framework is based on VIBe, a tool for generating virtual case studies including the urban environment and the urban drainage system. In DynaMind the computational framework has been enhanced and generalised to enable the simulation of complex dynamic urban environments and their infrastructure. Simulations can be defined by modules (small working task within a simulation) and the data flow between them. A module has in-ports to receive data from the outside. The data are then processed in the module and sent to out-ports. As data format raster and vector data can be used. Different modules can be linked together to a simulation workflow that is executed in parallel by the DynaMind core. DynaMind comes with a set of basic modules. This includes simple modules for data import and export, complex modules (for example the generation and adaptation of combined sewer networks or the placement of ground water heat pump systems) as well as modules for data visualisation. New modules can easily be integrated in DynaMind by using C++ or Python. DynaMind is freely available as open source software using the GPL license. To show the applicability a simple integrated case study to investigate the effects of urban development on an existing combined sewer system is set up in this paper.

INTRODUCTION

To test the effectiveness of new technologies and strategies in urban water management and to tackle the problems induced by climate and urban change, integrated modelling approaches on city scale are required. Traditional modelling tools for urban water systems like SWMM [1] consider the urban networks as static and decoupled from the urban environment. To couple the urban environment with the water infrastructure SWMM has been integrated in GIS based software tools like Mike Urban. Even though, the urban water infrastructure is still considered as a static system and the evaluation of performance is possible only at a certain point in time.

To prepare a case study for integrated modelling on city scale huge amounts of data at a very fine scale are required. Existing data are often either of poor quality or simply not available and therefore the preparation of a case study is a very time consuming and frustrating task. The situation gets even worse when future scenarios are evaluated. Thus often only a small number of simplified story lines can be investigated where only the present state and the future state are evaluated (Semadeni-Davies et al. [2]).

The DynaVIBe approach (Sitzenfrei et al. [4]) tackles the problem by introducing a method for the algorithmic generation of integrated case studies that evolve dynamically in the future. The so called virtual case studies are based on parameters derived from real world case studies. This methodology enables an analysis of future scenarios on a spatio-temporal city scale. Based on the DynaVIBe approach a dynamic environment for integrated modelling – DynaMind has been developed and is presented in this paper.

Aim of DynaMind is to provide an environment where modules can be linked together to prepare data for integrated modelling or to simulate dynamic urban environments. It should be an easy to use, fast and memory efficient and flexible tool. DynaMind is freely available as open source software using the GPL license. To show the potential of DynaMind a simple simulation is set up where the urban development is linked with an infrastructure adaptation module to update the impervious area. The results are used in the SWMM module to assess the performance. This model can be used to investigate the impacts of urban development on an existing combined sewer network.

METHODS

To describe an urban environment raster and vector data are used in DynaMind. This data are modified by modules. A module is a small program in itself, it can receive data via in-ports and the results of the module are sent to its out-ports. To describe the dynamics in an urban environment, several modules are linked together to enable complex simulations. DynaMind has been designed as a fast, memory efficient and easy to use modelling environment. Also the development of new modules is straight forward as well as the exchange of modules.

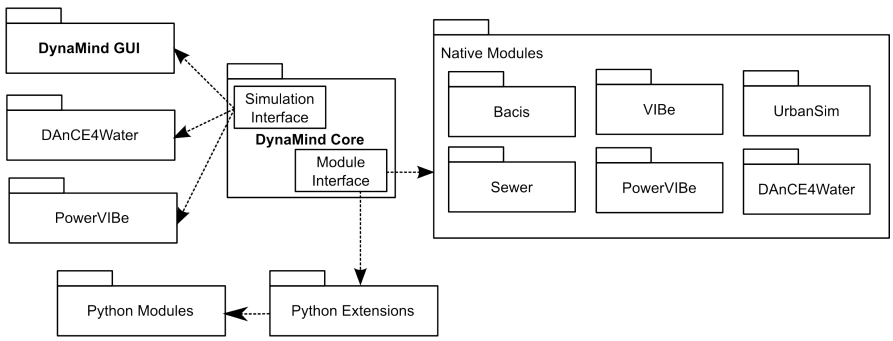

Figure 1. DynaMind Component Diagram

DynaMind is split up in several components to guarantee a strong and robust code base which is easy to adapt in the future (see figure 1). The heart of the software is the DynaMind Core that provides a fast and memory efficient environment to set up and execute simulations. DynaMind comes already with a number of modules; basic modules like import and export of shapefiles, to plot raster and vector data or a generic cellular automate; complex modules like modules to analyse networks, to auto design combined sewer systems and modules that integrate external simulation tools like SWMM [1] for the hydraulic performance assessment of drainage networks or UrbanSim [6] to simulate the complex dynamics in an urban environment. DynaMind provides easy to use interfaces to develop new modules in C++ or Python. The Modules are dynamically loaded during runtime from external libraries. To make DynaMind usable for non-programmers a graphical user interface has been developed that provides an easy way to create and execute simulations.

Modules and Groups

A Module is a small program (interface class) that is used to modify and or create raster and vector data. It consists of an initialisation and a run method (see figure 2). Within the initialisation method data (raster or vector data) in and output as well as parameters are defined. Supported parameters are boolean, int, double and strings. The DynaMind Core provides methods to modify these parameters during runtime. In the DynaMind GUI standard input dialogues are created to set module parameters at runtime (see figure 2)

```
class WhiteNoise(Module):
        def __init__(self):
            Module.__init__(self)
            self.Height = 200
            self.Width = 200
            self.CellSize = 20
            self.RasterData = RasterDataIn
            self.addParameter(self, "Height", VIBe2.LONG)
            self.addParameter(self, "Width", VIBe2.LONG)
            self.addParameter(self, "CellSize", VIBe2.DOUBLE)
            self.addData(self, "RasterData", VIBe2.RASTERDATA_OUT)

        def run(self):
            ...
```

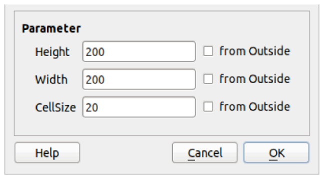

Figure 2. Module Sample Code in Python and Standard GUI

Modules can be lumped together in groups. A group can contain modules as well as groups and can be executed repeatedly - like a for-loop in a programming language. Groups can be used like modules (see figure 4).

Data management

Vector data are used to describe a variety of different objects like pipes in a sewer network or houses in a city. A vector data system has been developed that can be used for a variety of different objects that describe the geometry and/or attributes of urban environments like sewer networks, single family houses or households. The simplest object in the vector data system is a component. The component object contains a vector of attributes (double or string). Derived from components are nodes, lines and the system objects. A system consists of a vector of nodes, lines and subsystems and is used to describe complex objects.

Every object is identified by a unique identifier that is used to link objects together e.g. to assign a household (component) to a house (system).

Between modules complex data sets in form of raster or vector data describing the urban environment are exchanged. Therefore a fast and efficient data management system is required. The data are managed by the core and modules only hold references to them. This guarantees that each data set is stored only once during a simulation. The references are managed by the data management system and are updated before a module is executed. This method avoids that huge amounts of data are copied between different modules.

If a module adds or modifies an object in a data set, a new state of that data set is created. A data set state only contains references to objects that are stored in a data base. If a new object is added to the data set, it is added to the data base and the reference is added to the state. If an already existing object is modified, the reference in the new state is changed to point to the modified object. The states are representing some kind of content versioning system to analyse the development of a data set. Furthermore saving all states during a simulation can help to debug a simulation.

The data management system is connected via an interface to a database. At the moment the data is kept in the main memory but in a later step a SQL backend will be developed to handle huge data sets. The main memory will then be used for caching data stored in a SQL data base.

Simulations

Simulations consist of several linked modules. Before a module can be added to a simulation the module needs to be registered in the simulation environment. Therefore interfaces are provided by the DynaMind Core that dynamically load modules (implemented in C++ or Python) from external libraries. When all modules are linked together a simulation can be executed.

Since all modules are linked together the data flow in the simulation is defined and modules can be if possible executed in parallel. As a first step the following simple parallelization strategy has been implemented: within a group of modules the simulation looks for modules that don't need data from outside (modules with no in-port). These modules are executed in a thread. After a module has been executed, the program checks if all connections of the next downstream modules are satisfied and if so the module is executed in a new thread.

To access generated data during the runtime of the simulation, a data observer can be registered within the simulation. With the help of simulation observers a dynamic adaption of simulations during the runtime is possible. E.g. modifying parameters in downstream modules or adding new modules to the simulation.

Graphical User Interface

On top of the DynaMind core a graphical user interface has been developed. The GUI supports the user by providing an easy way to create and modify simulations. As shown in figure 3 the available modules are shown on the left hand side. By drag and drop modules or groups from the list of available modules to the right side of the GUI a simulation workflow can be defined. The modules are then represented as boxes and could be seen as an atomic working task within the simulation. Modules are linked together by drawing the line between in and out ports which represent the data flow. By clicking on a module a dialog box appears to set module parameters. The GUI also provides a basic viewer for raster and vector data. A freely available and multi-platform Python text editor called Editra (http://editra.org/) has been integrated in the DynaMind GUI. Therefore new modules can be developed within the GUI.

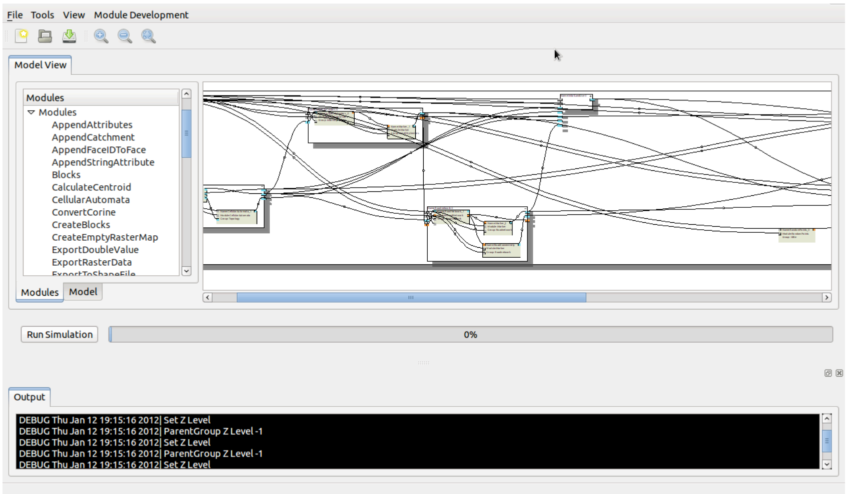

Figure 3. DynaMind GUI

RESULTS AND DISCUSSION

To demonstrate how DynaMind can be used a simple example of an integrated model is shown in figure 4. Aim of this model is to link an urban environment with a drainage system to investigate the effects of urban development on an existing combined sewer network. For this example only basic modules that come with the DynaMind tool are used.

Before the urban system (urban environment and its water infrastructure) can be evolved the initial system which represents a particular point in time needs to be set up. In this example the urban environment is based on raster maps for land use and population. For the sewer network and the catchments (linked with inlet points of the sewer network) vector data are used. For the initial system we import the data by using the basic modules *Import Shapefile* and *Import Rasterdata* grouped together in the *Initial Environment*.

The *Dynamic Environment* evolves the urban system one year in the future. It is split up into three major parts, *Urban Development*, *Infrastructure Adaptation* and *SWMM*. For the *Urban Development* we use a cellular automata model as described in Sitzenfrei et al.

[3]. To set up the model in DynaMind several *Cellular Automata (CA)* modules are linked together. Next we adopt our urban drainage system. In this example we update the impervious area of the catchment. Based on the rules presented in Sitzenfrei et al. [3] a *CA* is used to generate the impervious area based on land use and population. To update the information in the catchment we use the *Append Data (AD)* module. This module intersects the catchment shape with the raster data of the impervious area, calculates the median value and appends/modifies the attributes stored in the catchment. The performance of the combined sewer system is assessed with the *SWMM* module. This module exports the combined sewer system to the hydraulic solver SWMM [1] and appends the results of the hydraulic simulation to the combined sewer system.

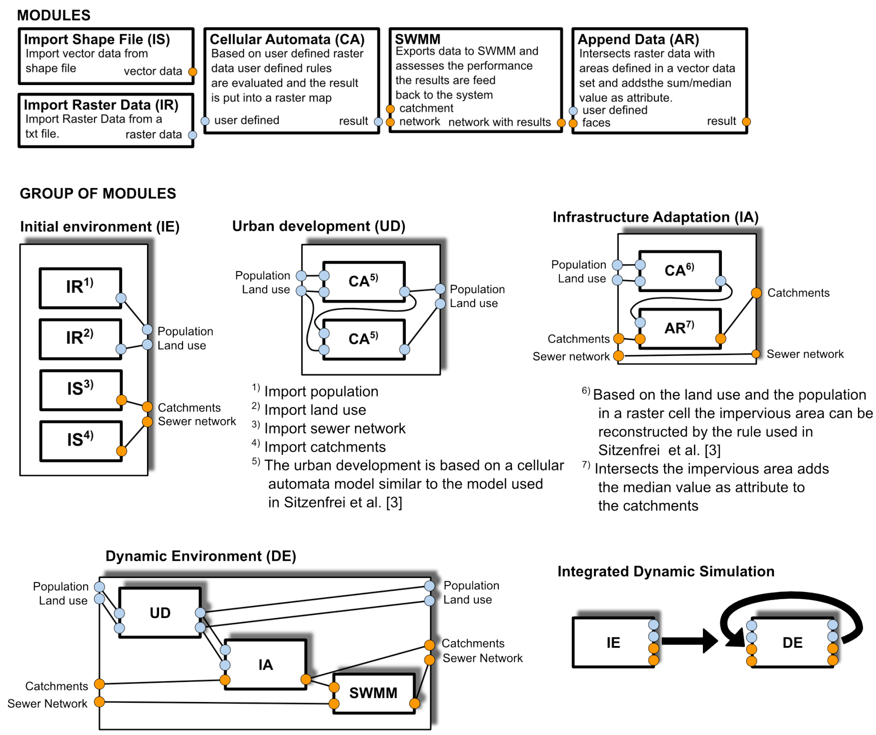

Figure 4. Example Simulation of an Integrated Urban Environment

The three modules evolve the urban system one year. To evolve the urban environment another year the results of the current simulation are used as input for the *Dynamic Environment*. (link out port – in port). To evolve the urban system 20 years in the future the *Dynamic Environment* is repeated 20 times.

In Urich et al. [5] DynaMind has been successfully applied to set up a complex integrated environment to test the applicability of infiltration trench system as adaptation strategy for urban drainage networks. Therefore UrbanSim (Waddell et al. [6]) has been integrated as *Urban Development* model. UrbanSim simulates the complex interactions within an urban environment on household level. The *Infrastructure Adaptation (IA)* module has been enhanced by algorithms that extend the existing network to connect new build up areas and to place infiltration trench systems. As shown in figure 5 the simulation is based on the same structure as the simple example.

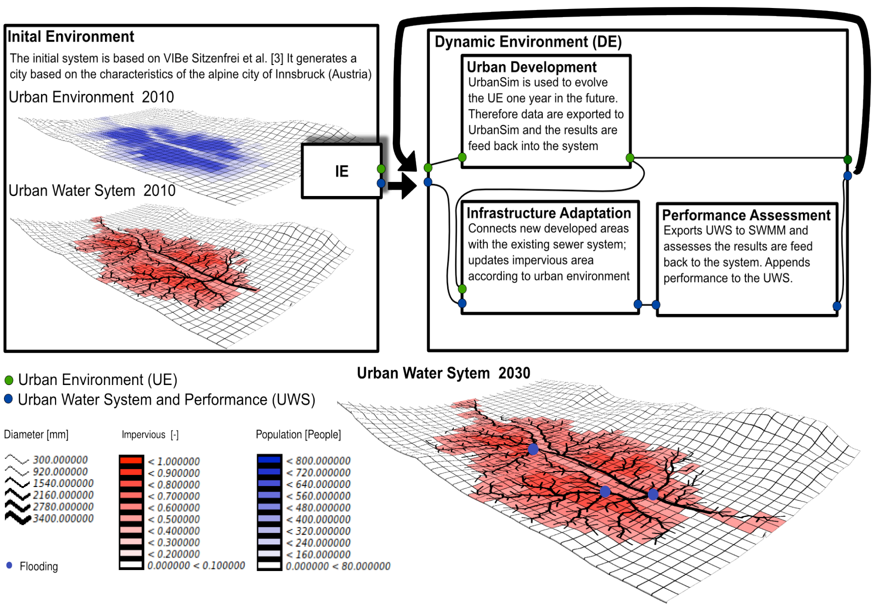

Figure 5. Complex Integrated Model (Urich et al. [5]) and Results

The applicability of DynaMind has also been successfully demonstrated in several scientific projects that enable the simulation of dynamic integrated urban environments like PowerVIBe, DynaVIBe and DAnCE4Water.

CONCLUSION AND OUTLOOK

This paper presents the design of the DynaMind framework and graphical user interface for integrated modelling of urban environment their infrastructure. Aim of DynaMind is to provide a software tool where modules, modifying the urban environment, are linked together to a simulation workflow to simulate dynamic urban environments.

DynaMind is split up in several components. A small, fast and efficient core provides interfaces to dynamically load modules from external libraries (implemented in C++ or Python) and to set up and execute simulations. The data to describe the urban environment (raster and vector data) are managed by the core. Based on the core a graphical user interface has been developed to provide the user with a simple tool to create and modify simulations.

To show the potential of DynaMind a simulation setup for a simple integrated environment is shown that can be evolved 20 years in the future. The case study can be used to investigate the effects of urban development on an existing combined sewer system. In this simulation the urban development is linked with a simple infrastructure adaptation module to update the impervious area. This is used in the SWMM model to assess the performance.

The applicability of DynaMind has also been successfully demonstrated in several scientific projects that enable the simulation of dynamic integrated urban environments like PowerVIBe, DynaVIBe and DAnCE4Water.

ACKNOWLEDGMENTS

This work was funded by Austrian Climate and Energy fund within the Energy 2020 program (project PowerVIBe). The authors gratefully acknowledge the financial support.

REFERENCES

[1] L. A. Rossman, "Storm water management model User's manual version 5.0," *Water Supply and Water Resources Division National Risk Management Research Laboratory Cincinnati*, 2008.

[2] Semadeni-Davies, A., Hernebring, C., Svensson, G., & Gustafsson, L. (2008). "The impacts of climate change and urbanisation on drainage in Helsingborg, Sweden": *Suburban stormwater. Journal of Hydrology*, 350(1-2), 114-125.

[3] R. Sitzenfrei, S. Fach, H. Kinzel, and W. Rauch, "A multi-layer cellular automata approach for algorithmic generation of virtual case studies: VIBe," *Water Science & Technology*, vol. 61, no. 1, pp. 37-45, Jan. 2010.

[4] R. Sitzenfrei, S. Fach, M. Kleidorfer, C. Urich, and W. Rauch, "Dynamic virtual infrastructure benchmarking: DynaVIBe," *Water Science & Technology: water supply*, vol. 10, no. 4, pp. 600–609, Sep. 2010.

[5] C. Urich, P. Bach, and C. Hellbach, "Dynamics of cities and water infrastructure in the DAnCE4Water model," in *Proceedings of the 12th International Conference on Urban Drainage, Porto Alegre/Brazil, 10-15 September 2011*, 2011.

[6] P. Waddell, "UrbanSim: Modelling urban development for land use, transportation, and environmental planning," *Journal of the American Planning Association*, vol. 68, no. 3, pp. 297–314, 2002.

18

PAPER XIII

Performance improvement with parallel numerical model simulations in the field of urban water management

Authors: **Michael Mair**, Robert Sitzenfrei, Manfred Kleidorfer and Wolfgang Rauch
Date: 2014
Place: Journal of Hydroinformatics
Type: Journal

477 © IWA Publishing 2014 Journal of Hydroinformatics | 16.2 | 2014

Performance improvement with parallel numerical model simulations in the field of urban water management

Michael Mair, Robert Sitzenfrei, Manfred Kleidorfer and Wolfgang Rauch

ABSTRACT

Numerical models are used to enhance the understanding of the behavior of real world systems. With increasing complexity of numerical models and their applications, there is a need of more computational power. State of the art processors contain many cores on one single chip. As such, new programming techniques are required if all these cores are to be utilized during model simulation runs. This manuscript reviews the runtime and speedup behavior of parallel model analysis software (e.g. Calimero and Achilles) applied to simulation tools for urban water management (e.g. CityDrain3, EPANET2, SWMM5, par-SWMM). The potential of using a parallel programming environment for 'coordinating' tasks of multiple runs of commonly used modeling software is analyzed. This is especially interesting as the modeling software itself can be implanted sequentially or parallel. Performance tests are performed on a set of real-world case studies. Additionally, a benchmark set of 2,280 virtual case studies is used to investigate performance improvement in relation to the size of the system. It was found that speedup depends on the system size and the time spent in critical code sections with increasing number of used cores. Applying parallelism only at the level of the model analysis software performs best.

Key words | Achilles, Calimero, nested parallelism, runtime, speedup, virtual test set application

Michael Mair (corresponding author)
Robert Sitzenfrei
Manfred Kleidorfer
Wolfgang Rauch
Unit of Environmental Engineering,
Institute of Infrastructure,
University of Innsbruck,
Technikerstr. 13,
6020 Innsbruck,
Austria
E-mail: *michael.mair@uibk.ac.at*

INTRODUCTION

Numerical models are applied in many research fields to better understand the behavior of real world systems. In the field of urban water management there are a number of software products frequently applied (e.g. EPANET, SWMM5, CityDrain3 (Burger *et al.* 2010)). Increasing complexity of models (i.e. increasing number of parameters), as well as modern analysis methods applied to urban water management, require a multitude of model simulation runs (e.g. in sensitivity-, scenario- or uncertainty analysis). As a result, more computational power is needed. In the last decades, the ever increasing demand for computational power was satisfied simply by increasing the clock frequency of Central Processing Units (CPUs). However, a point has been reached where this method of improving hardware performance is no longer efficient. Adding more CPUs on a single chip and leaving the clock frequency nearly unchanged was deemed a better alternative (Olukotun *et al.* 1996). This decision has consequently changed conventional programming techniques that developers are used to (Sutter 2005). The era of parallel programming has now reached all fields of software development as the necessary hardware has become available on inexpensive desktop machines (Hill & Marty 2008). This new trend implies several new programming paradigms and frameworks for concurrent programming on different levels. For example, fine-grained parallelism can be realized with the help of GPGPUs (general-purpose graphics processing units) or using SIMD (single instruction multiple data) registers of CPUs by exploiting data parallelism. Medium-grained parallelism is, for example, using programming paradigms like OpenMP (Dagum & Menon 1998) on shared memory systems and MPI (Message Passing Interface) (Gabriel *et al.* 2004) on distributed memory systems. On this level, the parallelization is mostly realized on functional concurrency. The third and last level of parallelism is called

doi: 10.2166/hydro.2013.287

coarse-grained parallelism where parallelism is on the level of the workflow within a scientific domain.

Numerical models in the field of urban water management are often implemented sequentially, e.g. SWMM5 (Rossman 2010) and EPANET2 (Rossman *et al.* 2000). Such algorithms can only utilize one single CPU within a simulation run. Recently developed or optimized numerical models, however, also contain parallel code e.g. CityDrain3 (Burger *et al.* 2010), par-SWMM (Burger & Rauch 2012) a parallel version of SWMM5. Moreover, work was carried out on exemplifying the potential to solve hydraulic network equations on GPGPUs (Crous *et al.* 2012), development of a parallel demand driven hydraulic solver by exploiting SIMD registers (Guidolin *et al.* 2013) and executing integrated flood models on clusters (Moya *et al.* 2013). The parallelization of certain algorithms within one software product (e.g. CityDrain3 and parSWMM) represents one of many solutions for decreasing runtime of model runs. For many analyzing techniques, numerous independent and dependent model simulation runs are required. This is the case, for example, during the model development process where one major step is to calibrate and validate model simulation results with measured data by adapting model parameters. The adaptation of the parameters is influenced by the deviation of previous results of model simulations and real world data (Kleidorfer *et al.* 2009). Alternatively, sensitivity analysis can be used to understand the behavior of a real world system. Mair *et al.* (2012a) or Möderl *et al.* (2011a) performed such analyses (spatial sensitivity analyses) to assess the vulnerability of water supply and sewer networks, respectively. By varying model parameters (e.g. simulating a conduit collapse or pipe burst) the consequences can be analyzed by observing the change of infrastructure-specific hydraulic performance. To get a complete analysis, all conduits or pipes have to be tested. Each model modification defines a new model setup, which has to be simulated. This also results in a huge amount of independent simulation runs.

The tasks described above (calibration, sensitivity analysis, uncertainty analysis, etc. of numerical models) can either be performed manually (by starting a sequence of model runs) or by a software in itself that coordinates these tasks. This manuscript shows the potential of using a parallel programming environment in such a 'coordinating' software for multiple runs of parallel and sequentially implemented modeling software in the field of urban water management. Usually, for such tests, a limited number of case studies can be used due restricted availability of case study data. However, evaluating the speedup behavior with one or more case studies results in case specific results which cannot (or at least are difficult to) be generalized or transferred to boundary conditions (e.g. memory usage). Thus, the aim of this study is to generate case unspecific results for the question at hand. Therefore, the impact of different model sizes on the overall runtime is analyzed with a benchmark set of 2,280 synthetic model setups (Möderl *et al.* 2011b) and additionally with three real model setups. With the presented approach, this knowledge gap can be addressed. Additional to Mair *et al.* (2012b) one more test case is analyzed where nested-parallelism is applied by using a parallel version of SWMM5 (Burger & Rauch 2012).

MATERIAL AND METHODS

The analysis of urban water management approaches and techniques requires a large number of model simulations – in essence parameter variations within an initial model setup. Two main levels can be identified where a parallelization of programming code may decrease the overall runtime of the analysis.

1. The first level is to parallelize the programming code at the level of the modeling software (Figure 1, blank box). This has been done, for example, in the modeling software CityDrain3 (CD3) (Burger *et al.* 2010) or parSWMM (Burger & Rauch 2012) (medium- or fine-grained parallelism). Consequently the runtime of a single model simulation is decreased with the help of several parallelization strategies. We define this level as MS-1 level (**M**odeling **S**oftware at level **1**).
2. The second level parallelization is done in the model analysis software (i.e. the 'coordinating' software, Figure 1, gray colored box) by executing model simulation processes in parallel. If we define the execution of a model simulation as a 'Task', this can be denoted as task level parallelism (coarse-grained parallelism). Consequently, there is no speedup for a single model

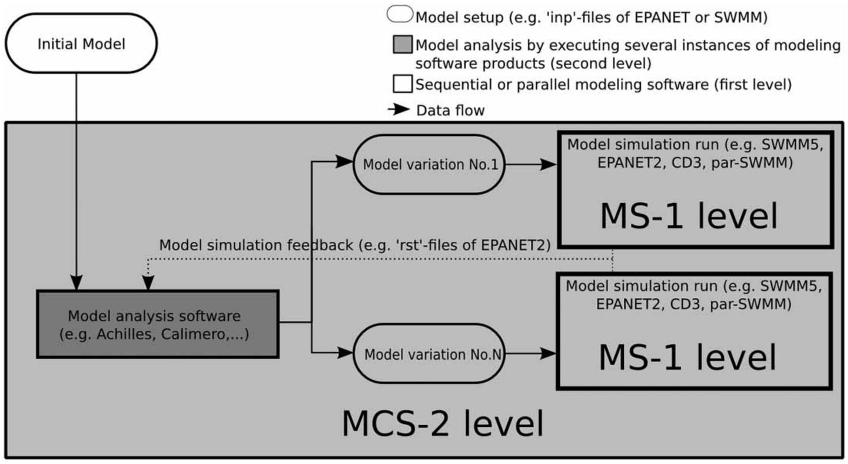

Figure 1 | Parallelization strategy at the MCS-2 level.

run but there is a speedup if multiple simulation runs are required. We define this level as MCS-2 level (**M**odel **C**oordinating **S**oftware at level **2**).

In this investigation, we put emphasis on decreasing the overall runtime by parallelizing programming code on the MCS-2 level. The code base of the MS-1 level is hence left unchanged. This is the typical scenario for a software user who does not have access to the source code or the ability to implement parallel algorithms in the modeling software. A speedup should be achieved by parallel execution of different model simulation runs, regardless of whether the employed modeling software includes parallel or sequential code. Parallel code on both levels is denoted as nested parallelism. This is of special interest in this study as a mutual interference is to be expected for such cases. Figure 1 shows the general parallelization strategy used in this work. The input is an initial model setup (e.g. input-file of SWMM5). The model analysis software performs several analyses by altering the initial setup and simulating the new model setups in parallel (e.g. SWMM5, EPANET2) or nested parallel (e.g. CD3 and par-SWMM).

Description of case studies

According to the previously defined parallelization strategy, two types of MCS-2 level tools (Achilles and Calimero, see below) are analyzed with regard to their runtime and speedup behavior with increasing number of available cores on a multi core system with shared memory. The speedup is calculated by dividing the measured runtime of the sequential version by the runtime of the parallelized version. While there are many other examples of MSC-2 level tools known and applied, these two are chosen as they cover two distinct features of analysis methods. The most generic MCS-2 level tool would be, for example, a simple Matlab script which sequentially or parallel executes several MS-1 level instances (any model simulation tool).

Achilles

The first model analysis software and example for a MCS-2 level parallelization is called Achilles (Möderl *et al.* 2010). Achilles is a module for the open source environment SAGA GIS and implemented in C++. The module emulates hazardous events on urban water infrastructure (operational failures, sabotage, land-use change, climate change, earthquakes, debris flows, fluvial flooding, etc.) and performs a spatial sensitivity analysis. According to the type of hazard, defined model adaptations are first performed automatically, then simulated (e.g. sequentially setting increased fire flow demand on each demand junction in a water supply network) and lastly, assessed in terms of hydraulic performance. Each model adaption represents a new scenario, which could be a potential candidate to become representative for a hazardous event with a negative impact on the urban water infrastructure. To analyze the whole system, this step has to be repeated for each infrastructure element (e.g. for each pipe or junction in a water supply network model or for each conduit or node in a combined sewer system model). This results in a huge amount of independent model variations and simulations, which can be carried out in parallel. The number of required model simulation runs is even higher if Achilles analyzes cascading effects of hazardous events within an infrastructure network (Sitzenfrei *et al.* 2011). For this work, Achilles was used to analyze hazardous events on water supply networks (Case study one – CS1) and combined sewer systems (Case study two – CS2) by adapting and simulating EPANET2 and SWMM5 models. The communication between Achilles (MCS-2 level) and the modeling software (MS-1 level) is made true by reading and writing output and input files of the modeling software, respectively.

In CS1.1, component outages within a water supply network (e.g. pipe burst) are simulated and analyzed. The parallelization is only realized on the MSC-2 level (i.e. only in Achilles itself). Each outage of a component is simulated with the help of EPANET2. With global performance indicators taking into account, for example the hydraulic performance, the impact of the component failure is evaluated. All simulations can be performed simultaneously. To point out model size-dependent effects on the speedup, all computational performance tests are based on synthetic water supply networks of different sizes. Therefore, a set of 2,280 water supply models (Möderl *et al.* 2011b) ranging from 26 up to approximately 4,000 nodes were used.

In CS1.2 component failures within a combined sewer system (e.g. conduit collapse) are simulated and analyzed. Here, a combined sewer system that is located in an alpine city with approximately 120,000 inhabitants (model size of 5,400 nodes) is used as an example. The reason for testing the computational performance in such a test setup with only one model (no artificial case studies) is the comparatively high runtime of SWMM5 models. Hydrodynamic storm water simulation generally results in much higher computational effort as compared to EPANET2 simulations, which prevents us here from testing a large set of models. Parallelism in this case study is realized on both levels (nested parallelism) in which each component outage is represented as different SWMM5 input file and simulated with par-SWMM (parallel version of SWMM5) simultaneously. The runtime and speedup tests are realized with increasing the number of used cores of Achilles and par-SWMM (e.g. two cores used by Achilles and three cores used by par-SWMM resulting in six used cores in total).

Calimero

The second model analysis software and example for a MCS-2 level parallelization is called Calimero (Kleidorfer *et al.* 2009). Calimero can calibrate any model as long as its software is controllable over text-based input and output files or the source code of the model software is available for embedding the model in Calimero. In the second case communication between Calimero (MCS-2 level) and the modeling software (MS-1 level) can be realized without reading and writing files. In this work, communication between Calimero and the modeling software is performed with reading and writing output and input files of the modeling software, respectively. The required number of model alterations and simulations can be independent or dependent on the adopted calibration algorithm. The used optimization algorithm in this work is a parallel particle swarm algorithm (PSO) (Eberhart & Kennedy 1995). Each particle within a swarm, represented by one simulation of the modified initial model setup, is independent of each other particle. In this case, one movement step of the swarm can be calculated in parallel by simultaneously executing several instances of the modeling software. The parallelization strategy is exactly the same as in Achilles.

For this work, the runtime and speedup of a parallel PSO is tested by calibrating EPANET2 (CS2.1), SWMM5 (CS2.2) and CD3 (CS2.3) models alongside a varying number of used threads on a multi core system. In CS2.1 and CS2.2 parallelism is only on the MCS-2 level. In CS2.3 parallelism is applied on both levels and thus serves as an example for nested parallelism. The used models are all from the same region which is a water supply model, a hydrodynamic combined sewer system model and a conceptual combined sewer system model of the alpine city that also is used in case study CS1.2.

Test environment (TE)

The performance tests (runtime and speedup) of the two case studies were performed on the following test environment: Intel® Xeon® Processor X5650 (TE). The system has two Intel® Xeon® ProcessorX5650 @ 2.67 GHz and 24 GB of DDR3 ram. Each processor has 12 MB L2 cache and six cores running 12 hardware threads in Hyper-threading mode. The installed operating system is Arch Linux using the Linux kernel version 2.6.39-ARCH.

RESULTS

Runtime and speedup tests of all case studies are presented below along with an interpretation and discussion of their value.

Achilles (CS1.1)

Figure 2 shows the runtime of Achilles while analyzing component failures of 2,280 EPANET2 models. For all three subfigures the *x*- and *y*-axes represent the number of total threads used by Achilles and the runtime, respectively. Based on the runtime of a sequential execution of Achilles (using one thread only), three categories are defined containing small, medium and large sized models. The median runtime of the small, medium and large sets are approximately 3.4, 14 and 200 s by using one thread. Comparing these values with the related model sizes, it was observed that the small category contains 114 models (26–70 nodes), the medium category contains 342 models (70–210 nodes) and the large category contains 1,824 models (210–4,000 nodes).

Figure 3 shows the speedup of all three categories. The maximum median speedup in the small and large categories

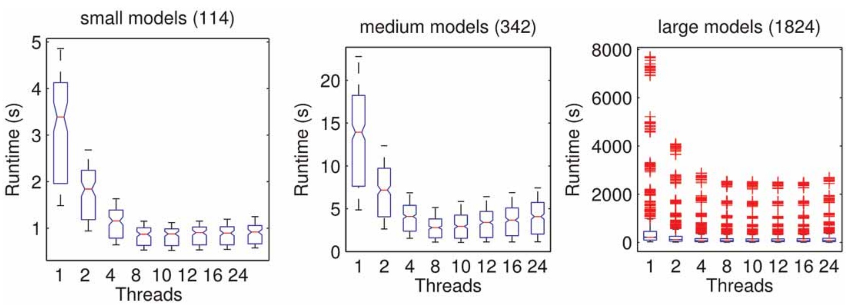

Figure 2 | Runtime of Achilles on TE testing 2,280 models of the set of synthetic generated water supply networks (CS1.1).

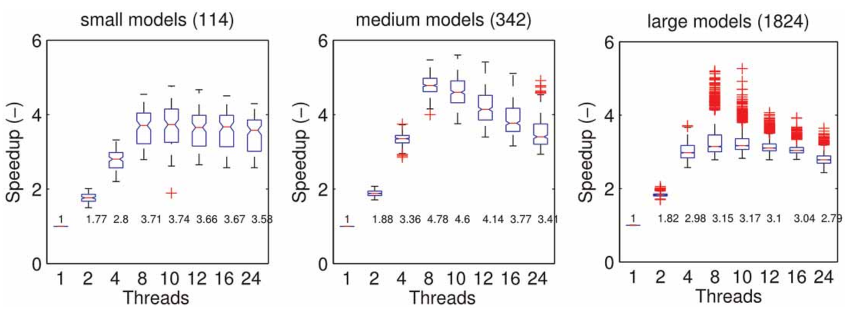

Figure 3 | Speedup of Achilles on TE testing 2,280 models of the set of synthetic generated water supply networks (CS1.1).

is at 3.74 and 3.17 by using 10 threads. In the medium category the median speedup is even better with 4.78 by using eight threads.

Assuming that the used parallelization strategy is embarrassingly parallel there should be a linear speedup up to 12 used threads (number of real physical cores on the test environment). Between 12 and 24 threads there should also be a speedup increase but not as high as is expected between one and 12 cores because of the hyper threading mode. However, the box plots in Figure 3 show that there must be another boundary for the speedup maximum because the peak is at eight and 10 used threads instead of 12.

$$S(N) = \frac{1}{(1-P) + P/N} \qquad (1)$$

This behavior can be explained with the help of Amdahl's law (Equation (1)), which predicts the theoretical maximum speedup $S(N)$ by knowing the number of cores N on a multi core system and P the fraction of parallel code in the program. According to the medium category in Figure 3, the median highest speedup is 4.78 by using eight threads. Upon solving the Amdahl's law equation for the unknown P, it is determined that the Achilles program contains approximately 90% of parallel executed code for this example. Figure 4 shows the percentage of parallel executed code of all models by solving the Amdahl's law equation for the unknown P for all three categories. We can see that the percentage of parallel executed code is dependent on the number of used threads and is not constant which means the program contains both a constant overhead and a thread depending overhead.

The problem is that an increase of the EPANET2 model size results in an increase of the fraction of parallel code in Achilles. This subsequently increases the data communication between EPANET2 and Achilles. However, in the current implementation, data communication is sequential if more threads are trying to enter the corresponding code section. Further investigations revealed that the bottleneck is not reading and writing files of EPANET2 result and input files. Instead the bottleneck is a critical code section within Achilles, which converts the results of the model simulation into the internal SAGA GIS data structure. Exactly this conversion is not thread safe and hence cannot be parallelized. The corresponding code can only be entered one thread at a time. Also, with increasing model size the time spent in this code section increases. With increasing number of used threads by Achilles the probability increases that more than one thread is trying to enter the critical code section. This results in a sequential execution of the code and in a decrease of parallel executed percentage of code.

Achilles with nested parallelism (CS1.2)

In this test case parallelism is applied on the MS-1 and MCS-2 level. The model analysis software applies a SWMM5 model by parallel executing par-SWMM instances. In Figure 5 the x-axis in both diagrams show the total number of used threads by Achilles and par-SWMM software. Different numbers of threads used by par-SWMM are indicated with different markers. If par-SWMM uses one thread for each instance, the numbers on the x-axes are equal to the numbers of used threads by Achilles. If par-SWMM uses four threads for each instance, the number of threads used by Achilles is the

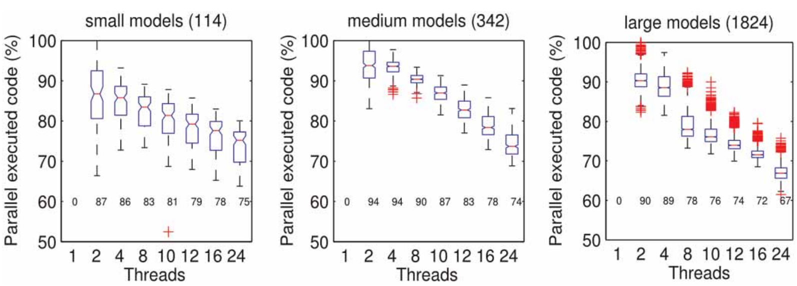

Figure 4 | Parallel executed code of Achilles by solving the Amdahl's law equation with known speedup and number of used threads (CS1.1).

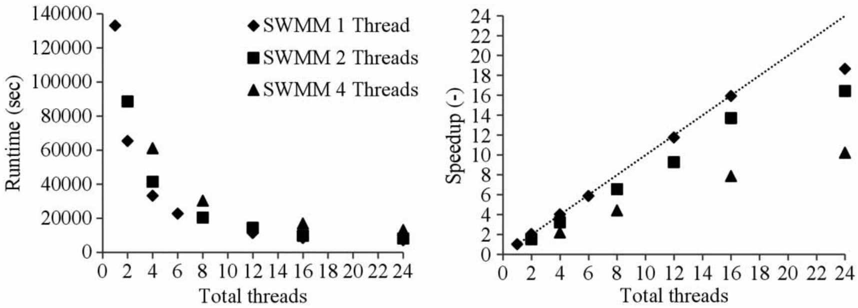

Figure 5 | Speedup of Achilles analyzing a SWMM5 model by executing par-SWMM instances in parallel on TE (CS1.2) (nested parallelism).

number of total threads divided by four (Figure 5, triangle marker – par-SWMM 4 threads). There is still an increase in speedup when applying nested parallelism (Figure 5 – par-SWMM 4 threads). Compared to the version with parallelism only on the MCS-2 level, the speedup behavior is better (Figure 5, diamond marker – par-SWMM 1 thread). The reason being that the efficiency in terms of processor load of the single threaded par-SWMM version is higher than the multi-threaded version and therefore parallelism only on the MCS-2 level can better utilize all available cores on a multi-core system. Comparing the resulting speedup with CS1.1 we can see that Achilles using par-SWMM with one thread for par-SWMM and 1–24 threads for Achilles is scaling slightly better. The reason for this is because of the smaller critical code section, which converts the SWMM5 results into the Achilles data structure. In this example the runtime spent within MS-1 level is much higher than in CS1.1 and therefore data communication has a minor impact.

Calimero

Figures 6 and 7 show the runtime and speedup test of Calimero (MCS-2 level) calibrating EPANET2 (CS2.1) and SWMM5 (CS2.2) models, respectively. Since most calibration algorithms contain heuristic mechanisms to determine the set of possible solution candidates, the required number of iterations must vary. Therefore, the runtime is measured by samples (number of iterations or number of tested possible solution candidates) per second. The maximum speedup obtained by calibrating an EPANET2 model with Calimero is approximately eight when using 24 threads (Figure 6). One kink occurs at 12 used threads with a speedup of seven. There is nearly a linear speedup up to 12 threads, which is the number of physical cores on the test environment. By using more threads than available cores (between 12 and 24 threads), speedup once again increases but not that high. This is because of the hyper threading mode in which resources

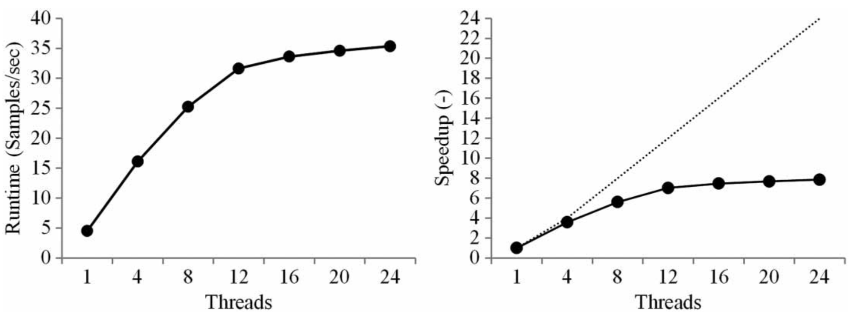

Figure 6 | Runtime (samples/sec) and speedup of Calimero calibrating an EPANET2 model on TE (CS2.1).

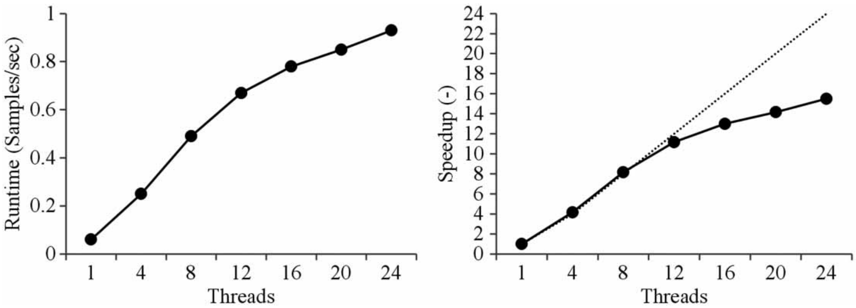

Figure 7 | Runtime (samples/sec) and speedup of Calimero calibrating a SWMM5 model on TE (CS2.2).

are shared (e.g. floating point unit – FPU). However, results are showing that an increase in speedup is still possible.

The reason for the better performance of CS1.1 compared to CS2.1 is that there is no critical code section within Calimero which is not thread safe. Additionally, only parts of the EPANET2 result files are read. Here, data communication between MCS-2 level and MS-1 level is less time consuming as compared to the Achilles case study. Solving the Amdahl's law equation again shows that the percentage of parallel executed code is nearly constant in the range of real physical cores, which is approximately 92%.

Figure 7 shows runtime and speedup tests for Calimero using SWMM5 models. The layout of the speedup curve is similar to the speedup curve in CS2.1. Again, the highest speedup is at 24 used threads with a factor of 15. A kink at 12 used cores occurs with a speedup value of nearly 12. Up to 12 used threads the speedup is nearly linear. Between 12 and 24 used threads the speedup increases but not that high. We can conclude Calimero calibrating a SWMM5 model is embarrassingly parallel and has nearly 100% of parallel executed code. This is due to the fact that the runtime of SWMM5 models is so high compared to the communication overhead that the latter is of minor importance.

Calimero with nested parallelism (CS2.3)

In this test case parallelism is applied on the MCS-2 and MS-1 level. The model analysis software calibrates a CD3 model by parallel executing CD3 instances (Figure 8). The *x*-axis in both diagrams shows the total number of used threads by Calimero and CD3. Again, applying nested parallelism does not increase the speedup (Figure 8, triangle marker – CD3 12 threads). Applying parallelism only on the second level utilizes all available cores better and therefore has a better speedup behavior (Figure 8, diamond marker – CD3 1 thread).

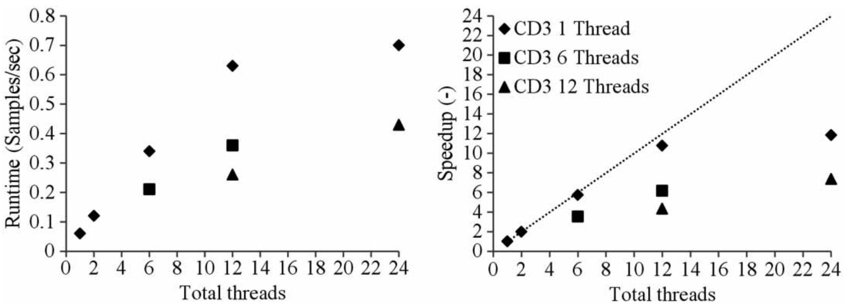

Figure 8 | Speedup of Calimero calibrating a CD3 model on TE (CS2.3) (nested parallelism).

CONCLUSION

In this manuscript the potential of a simple parallelization strategy is outlined for analysis software (**M**odel **C**ontrolling **S**oftware at level 2 – MCS-2 level), which inherently needs multiple simulations. The parallelization strategy executes several modeling software instances in parallel (the software itself implemented either sequentially or parallel (**M**odeling **S**oftware at level **1** – MS-1 level). This strategy is shown and analyzed based on runtime and speedup using two different analysis software products denoted as Achilles and Calimero. The investigations show that the speedup of model analysis software with parallelism on the MCS-2 level is increased significantly.

For all case studies data communication between MCS-2 and MS-1 level was realized with reading and writing output and input files of the modeling software. Case studies with Achilles showed that a performance improvement is possible by applying parallelism at the MCS-2 level. Speedup tests were performed on a set of 2,280 different sized water supply systems. Therewith, statistical evaluations and bandwidths of results were obtained. Also size dependent characteristics were analyzed by splitting the data set into three categories. It was shown that Achilles has a non-constant overhead depending on the number of used cores. This is because of a critical code section within Achilles, which can only be executed by one thread at a time. Also, with increasing number of used cores the probability increases that more than one thread tries to enter this code section at a time. This results in a decrease of parallel executed code by increasing the number of used threads. In contrast to that our investigations proved that Calimero performs better. The percentage of parallel executed code is constant in the range of real physical cores used, resulting in a nearly linear speedup.

In cases in which parallelism was applied on both levels (nested parallelism) no performance improvement was seen as compared to the parallel version only on the MCS-2 level. All used modeling software (par-SWMM and CD3) are not fully utilizing all given resources at runtime if they are using more threads. The best performance improvement can be obtained by applying parallelism only on the MCS-2 level, which is an important message for practical applications.

In conclusion, performance improvement of numerical model simulations in the field of urban water management can be obtained by parallelizing model simulation runs regardless of whether the modeling software itself is implemented sequentially or parallel. Even more, best speedup results can be obtained by applying parallelism only on the level of the model analyzing software (MCS-2 level). The more the data communication can be decreased, and especially critical code sections eliminated, the higher speedup values are obtained.

ACKNOWLEDGEMENTS

This work is part of the DynaVIBe project; funded by the Austrian Science Fund (FWF): P 23250-N24.

REFERENCES

Burger, G., Fach, S., Kinzel, H. & Rauch, W. 2010 Parallel computing in conceptual sewer simulations. *Water Sci. Technol.* **61** (2), 283–291.

Burger, G. & Rauch, W. 2012 Parallel computing in urban drainage modeling: A parallel version of EPA SWMM. *Proceedings of the Ninth International Conference on Urban Drainage Modelling*, Belgrade, Serbia, 4–6 September 2012. Faculty of Civil Engineering, University of Belgrade, Belgrade.

Crous, P., Van Zyl, J. E. & Roodt, Y. 2012 The potential of graphical processing units to solve hydraulic network equations. *J. Hydroinform.* **14** (3), 603–612.

Dagum, L. & Menon, R. 1998 OpenMP: an industry standard API for shared-memory programming. *Comput. Sci. Eng. IEEE* **5** (1), 46–55.

Eberhart, R. & Kennedy, J. 1995 A new optimizer using particle swarm theory. *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*. IEEE, Nagoya Municipal Industrial Research Institute, October 4–6, 1995, pp. 39–43.

Gabriel, E., Fagg, G., Bosilca, G., Angskun, T., Dongarra, J., Squyres, J., Sahay, V., Kambadur, P., Barrett, B. & Lumsdaine, A. 2004 Open MPI: Goals, concept, and design of a next generation MPI implementation. In: *Recent Advances in Parallel Virtual Machine and Message Passing Interface*. Springer, Berlin, Heidelberg, pp. 353–377.

Guidolin, M., Kapelan, Z. & Savic, D. 2013 Using high performance techniques to accelerate demand driven hydraulic solvers. *J. Hydroinform.* **15** (1), 38–54.

Hill, M. D. & Marty, M. R. 2008 Amdahl's law in the multicore era. *Computer* **41** (7), 33–38.

Kleidorfer, M., Leonhardt, G., McCarthy, D. T., Kinzel, H. & Rauch, W. 2009 CALIMERO – A model independent and generalized tool for auto calibration. *Proceedings of the 8th International Conference on Urban Drainage Modelling*, Tokyo, Japan, September, 2009. IWA Publishing, London.

Mair, M., Sitzenfrei, R., Kleidorfer, M., Moderl, M. & Rauch, W. 2012a GIS-based applications of sensitivity analysis for sewer models. *Water Sci. Technol.* **65** (7), 1215–1222.

Mair, M., Sitzenfrei, R., Urich, C., Kleidorfer, M., Möderl, M. & Rauch, W. 2012b Performance improvement with parallel numerical model simulations in the field of urban water management. *Proceedings of the HIC 2012 – 10th International Conference on Hydroinformatics*, Hamburg, Germany, 14–18 July 2012. TuTech Verlag, Hamburg.

Möderl, M., Hellbach, C., Sitzenfrei, R., Mair, M., Lukas, A., Mayr, E., Perfler, R. & Rauch, W. 2011a GIS based applications of sensitivity analysis for water distribution models. *Proceedings of the 2011 World Environmental and Water Resources Congress*, May 22–26, Palm Springs, California. ASCE, American Society of Civil Engineering, Reston.

Möderl, M., Sitzenfrei, R., Fetz, T., Fleischhacker, E. & Rauch, W. 2011b Systematic generation of virtual networks for water supply. *Water Resour. Res.* **47** (2), W02502.

Möderl, M., Sitzenfrei, R. & Rauch, W. 2010 Achilles approach to identify vulnerabilities in urban water infrastructure for operation and emergency management. *Proceedings of the IWA World Water Congress and Exhibition*, Montreal, Canada, 19–24 September 2010. IWA Publishing, London.

Moya, Q. V., Popescu, I., Solomatine, D. & Bociort, L. 2013 Cloud and cluster computing in uncertainty analysis of integrated flood models. *J. Hydroinform.* **15** (1), 55–70.

Olukotun, K., Nayfeh, B. A., Hammond, L., Wilson, K. & Chang, K. 1996 The case for a single-chip multiprocessor. *Proceedings of the ACM Sigplan Notices*, Vol. 31. Association for Computing Machinery, New York, pp. 2–11.

Rossman, L. A. 2010 *Storm Water Management Model User's Manual, Version 5.0*. National Risk Management Research Laboratory, Office of Research and Development, US Environmental Protection Agency, Athens, GA, USA.

Rossman, L. A., Information, C. f. E. R. & Laboratory, N. R. M. R. 2000 *EPANET 2: Users Manual*. Lewis A. Rossman, Water Supply and Water Resources Division, Athens, GA, USA.

Sitzenfrei, R., Mair, M., Möderl, M. & Rauch, W. 2011 Cascade vulnerability for risk analysis of water infrastructure. *Water Sci. Technol.* **64** (9), 1885–1891.

Sutter, H. 2005 The free lunch is over: a fundamental turn toward concurrency in software. *Dr. Dobb's J.* **30** (3), 202–210.

First received 30 October 2012; accepted in revised form 19 April 2013. Available online 30 May 2013

19

PAPER XIV

The application of Web-geographic information system for improving urban water cycle modelling

Authors: **Michael Mair**, Christian Mikovits, Markus Sengthaler, Martin Schöpf, Heiko Kinzel, Christian Urich, Manfred Kleidorfer, Robert Sitzenfrei and Wolfgang Rauch

Date: 2014
Place: Water Science & Technology
Type: Journal

1838 © IWA Publishing 2014 Water Science & Technology | 70.11 | 2014

The application of a Web-geographic information system for improving urban water cycle modelling

M. Mair, C. Mikovits, M. Sengthaler, M. Schöpf, H. Kinzel, C. Urich, M. Kleidorfer, R. Sitzenfrei and W. Rauch

ABSTRACT

Research in urban water management has experienced a transition from traditional model applications to modelling water cycles as an integrated part of urban areas. This includes the interlinking of models of many research areas (e.g. urban development, socio-economy, urban water management). The integration and simulation is realized in newly developed frameworks (e.g. DynaMind and OpenMI) and often assumes a high knowledge in programming. This work presents a Web based urban water management modelling platform which simplifies the setup and usage of complex integrated models. The platform is demonstrated with a small application example on a case study within the Alpine region. The used model is a DynaMind model benchmarking the impact of newly connected catchments on the flooding behaviour of an existing combined sewer system. As a result the workflow of the user within a Web browser is demonstrated and benchmark results are shown. The presented platform hides implementation specific aspects behind Web services based technologies such that the user can focus on his main aim, which is urban water management modelling and benchmarking. Moreover, this platform offers a centralized data management, automatic software updates and access to high performance computers accessible with desktop computers and mobile devices.

Key words | hydrodynamic sewer system simulations, online scenario simulation, urban water management modelling, Web processing services

M. Mair (corresponding author)
C. Mikovits
M. Sengthaler
M. Kleidorfer
R. Sitzenfrei
W. Rauch
Unit of Environmental Engineering, University of Innsbruck,
Technikerstr. 13, 6020 Innsbruck,
Austria
E-mail: *Michael.Mair@uibk.ac.at*

M. Schöpf
H. Kinzel
hydro-IT GmbH,
Technikerstr. 13, 6020 Innsbruck,
Austria

C. Urich
Monash University Water for Liveability Centre,
Cooperative Research Centre for Water Sensitive Cities,
Department of Civil Engineering,
Monash University,
Clayton VIC 3800,
Australia

INTRODUCTION

In the last few years, research in urban water management has experienced a transition from traditional model applications to modelling water cycles as an integrated part of urban areas. The focus has been set on scenario based simulations for modelling alternative infrastructure systems (e.g. centralized vs. decentralized systems (Zhou *et al.* 2012; Sitzenfrei *et al.* 2013)), interactions with urban development (Willuweit & O'Sullivan 2013), transitions and socio-economic aspects (Bos *et al.* 2013; Ferguson *et al.* 2013). Mainly, population change and urban development, especially the connection of newly developed areas, put pressure on the urban water infrastructure (Semadeni-Davies *et al.* 2008; Mikovits *et al.* in press).

One of the projects dealing with these research questions is DAnCE for Water (Dynamic Adaptation for eNabling City Evolution for Water). The main aim of this project is research on adaptation of urban water infrastructures in response to a changing environment (Urich *et al.* 2013). Due to the high complexity of newly developed models and its interactions, a free available (licensed under the terms of the GNU general public licence) scientific workflow engine implemented in C++ called 'DynaMind' has been developed. It provides a platform for researchers and planners to combine urban water specific performance models, geographic information system (GIS) functionality, efficient data management and data visualization (Urich *et al.* 2012). The main strength of this framework is the ability to handle huge data sets in space and time. However setting up the framework is still a challenging task, especially for engineers who are not trained in programming. Going one step further, the next challenge is to provide end users (such as decision-makers, etc.) and

doi: 10.2166/wst.2014.327

model developers with an easy to use Web based service to apply and test their approaches and models.

For a comprehensive assessment of such complex and dynamic systems, conventional planning methods are successively replaced by dynamic models. Further, to increase their usability, these models are moved towards Web based applications (Dubois *et al.* 2013; Evangelidis *et al.* 2014). This enables a quick development of distributed, integrated, interoperating and complex models which need computer power beyond the capacity of standard desktop or laptop computers. Moving urban water management modelling towards Web service based technologies means also bringing high performance computing into the user's offices. Complex integrated urban water management modelling and simple user interactions can be brought together. User interaction, data storage, model setup and computational processing/model simulation does not happen necessarily on the same computing device. This allows a set up for customers, stakeholders and decision-makers to use complex models which are expected to interact with each other. Consequently the tool enables better evaluation of options leading to reduced costs and risks.

This paper presents the basic ideas and concepts of such a Web based urban water management modelling platform by using Web service based technologies. At the same time, this platform allows the cooperation of many user groups (e.g. model developer and decision-makers) within the same data set. A first approach of interlinking models and data handling by using Web service based technologies will be investigated and demonstrated with an application example applied on a case study within the Alpine region. As an example for a complex model, an existing DynaMind model (impact of new catchments on the flooding behaviour of a combined sewer system) is remotely executed via a Web browser or GIS software (e.g. ArcGis or QGis).

MATERIAL AND METHODS

Figure 1 shows the general concepts and ideas of the Web service based DynaMind platform. It consist of three main physical hardware components, which are: (1) client hardware (Figure 1 – lower left box), (2) a data server (Figure 1 – top left box), and (3) a Web processing service (WPS) server (Figure 1 – lower right box). The latter two may also be running on the same physical machine.

Data server

The data server (Figure 1 – top left box) is a GIS server which stores spatial referenced data. Many different

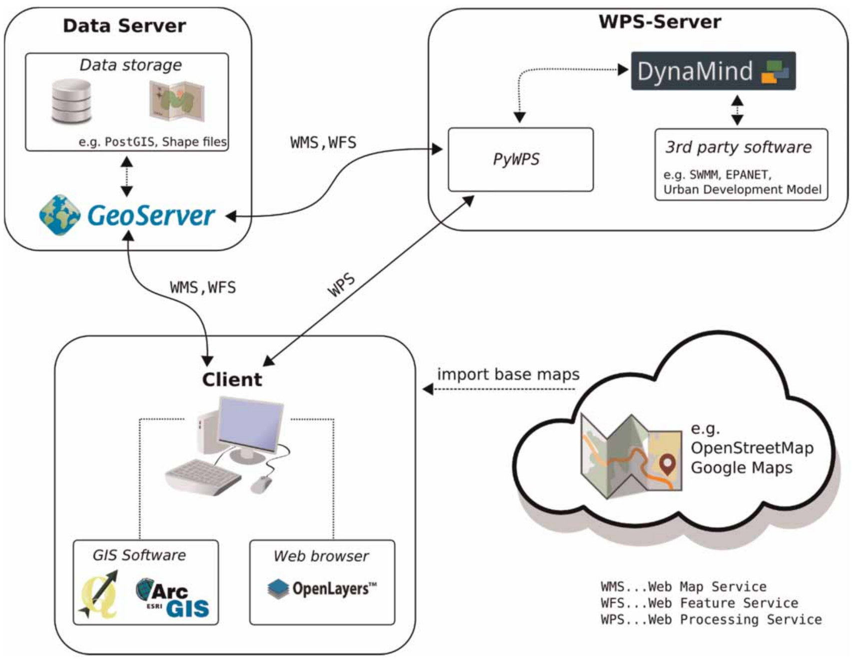

Figure 1 | General concept of the Web service based DynaMind platform.

techniques and storage formats are nowadays used e.g. PostGIS (data are stored in an object-relational database management system, such as PostgreSQL), Shapefiles (data are stored in a set of different files) or GeoTIFF files (an extension of the popular TIFF image format for storing georeferenced raster data). All these different storage formats can be seen as different data sources with the common goal to store spatial referenced vector and raster data. The development of all these formats happened over many years and also the development focus at each of them was slightly different. Consequently each format has its own access and modification routines. Hence, the development of applications which access many different data sources is time consuming. To overcome this problem, spatial data management environments were developed (e.g. GeoServer (Deoliveira 2008) or MapServer (Kropla 2005)) which are able to handle a huge set of different data sources and at the same time offering standardized protocols for accessing the stored raster and vector data. Such protocols are, e.g. Web map service (WMS) for raster data exchange and Web feature service (WFS) for vector data exchange. Data exchange between all three hardware components in the presented platform is realized with the standardized protocols developed by the Open Geospatial Consortium (Wenjue *et al.* 2004).

WPS server

The data stored at the data server and accessible with the standardized protocols can be modified by using WPS (Foerster & Stoter 2006). Which means the workflow for data manipulation can be defined by calling a single service or by cascading several services. The deployment of the services is realized with PyWPS (Schut & Whiteside 2007), which is a realization of the WPS 1.0.0 standard in the popular programming language Python. Data manipulation can now be implemented directly within this framework or by calling external programs. Under the scope of this work, data processing is realized by executing DynaMind models within PyWPS. Each model which is defined within the DynaMind framework can be seen as own WPS (Figure 1 – top right box).

Client

Due to the standardized exchange and manipulation of georeferenced data stored at the data server and manipulated at the WPS server respectively, the client can access the presented platform in two different ways: (1) by using a desktop GIS or (2) by using WPS, WMS, WFS capable frameworks (e.g. OpenLayers (Hazzard 2011)) within a Web browser. The latter technique is also known as Web-GIS where the client can be seen as data viewer and the major part of data manipulation is done at the WPS-server. This enables the usage of the whole platform for nearly any computing device that can run a Web browser.

Application example and proof of concept

The used application example in this work is a DynaMind model (Figure 3), which is callable via WPS using the PyWPS framework. The model simulates the impact of new catchments on the flooding behaviour of a combined sewer system by varying the catchment imperviousness of a new catchment and design storm events over the whole combined sewer system model. Therefore, the needed input is a polygon defining the new catchment area. This polygon/catchment can be manually drawn/specified by the user within a Web browser by using the OpenLayers framework and sent as argument within the WPS call to the DynaMind framework. Results of the model simulation are stored at the data server and can be accessed by the Web browser via the WFS.

Dynamind model

Figures 2 and 3 show the workflow of the used DynaMind model. The inputs are the user defined catchment area (e.g. new urban development area) represented as polygons and an existing or user defined combined sewer system model (e.g. stored at the WPS server or at a Web space which is accessible by the WPS server) which will be affected by the new catchments (Figure 3 – top white boxes).

In the first step of the model (Figure 3 – Step 1) the user defines a polygon (representing a new urban development area and therefore a new catchment – Figure 2(a)). This catchment is divided into several smaller polygons/subcatchments by normal distributing new inlet nodes within the polygon (Figure 2(b)) and creating a Voronoi diagram (Aurenhammer 1991) on the bases of these nodes (Figure 2(c)). Afterwards these new catchments are connected to the existing combined sewer system model (Figure 2(d)). Due to the splitting of the new urban development area into several subcatchments the runoff is distributed over several inlet nodes to the existing system. This prevents unrealistic flooding at single nodes due to a too coarse catchment representation. The result of these two steps is a new combined sewer system model which can further be

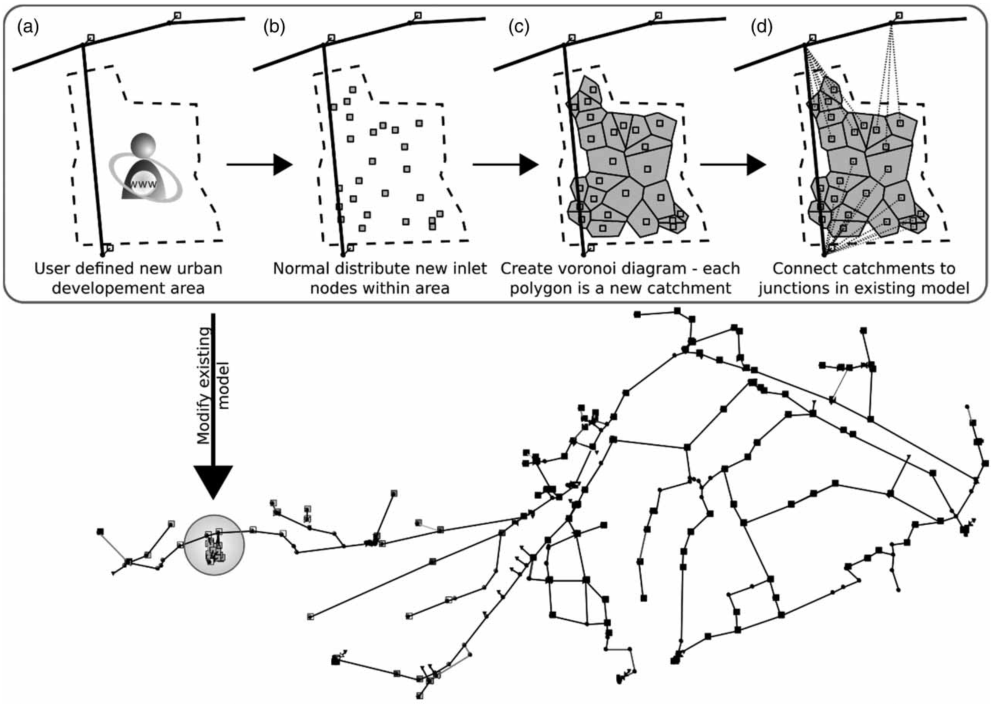

Figure 2 | Creating and connecting new catchments to an existing sewer system model.

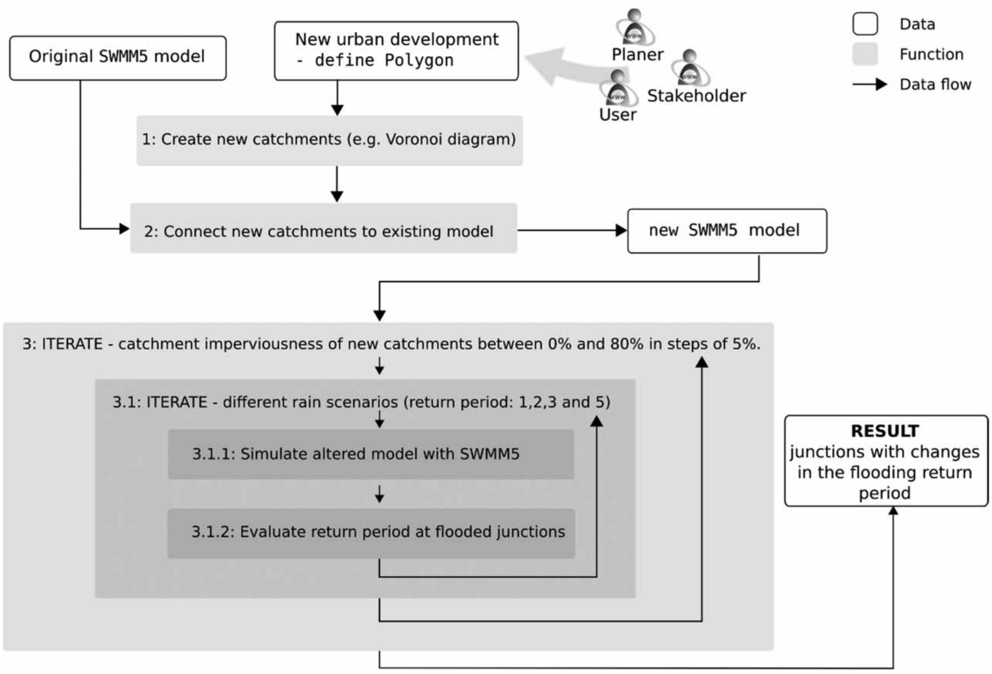

Figure 3 | DynaMind model of the application example.

analysed by varying model parameters and evaluating the results. For hydraulic performance assessment a parallelized version of the storm water management model (SWMM) software (Burger *et al.* 2014) is used. As benchmark procedure a one at a time parameter variation is used (Mair *et al.* 2012) in which the catchment imperviousness of the new catchments is varied between zero and 80 percent in steps of 5 percent (Figure 3 – Loop 3). As rainfall input and scenario variation, four different EULER II design storm events with a return period of one, two, three and five years are used (Kleidorfer *et al.* 2009) (Figure 3 – Loop 3.1). Each parameter variation defines a new combined sewer system model which will be simulated with all storm event scenarios. The simulation results are evaluated by investigating the flooding behaviour at each model junction (Figure 3 – Box 3.1.2).

First, all junctions where flooding occurs during the simulation period are extracted and afterwards the return period of the used design storm event within this simulation is mapped to that junction. If the catchment imperviousness for the new catchments is zero (First iteration of loop 3 – Figure 3), the resulting design storm events for the flooded nodes can be seen as initial value. All ongoing results are compared to these values and only stored if the return period for which flooding occurs decreases compared to the initial system. The result of the whole DynaMind simulation is a set of junctions showing a change in the flooding return period. Each junction has result parameters which are: old flooding return period, new flooding return period and the minimum catchment imperviousness of the new catchments which triggered the change in the flooding return period. This is a simple example to demonstrate the capabilities of the Web service based DynaMind platform. It can be used, e.g. by urban planners to assess the impact of a new development area on the sewer system performance. The advantage of this is that these user groups are not made responsible for data management, model building and maintenance. To get sound propositions about the system behavior a more detailed and accurate model (e.g. a coupled 1D/2D model for modelling conduit and surface flow, respectively) may be used in the future.

Case study

The whole platform is demonstrated on a small Alpine city with approximately 120,000 inhabitants. Storm water and sewage is discharged in a mainly gravity-driven and combined sewer system. The total catchment is 2,076 ha of which 774 ha are effective impervious area. For the whole case study a coarse combined sewer system model (SWMM5 model) is used. This model consists of 250 nodes 290 links and 182 subcatchments.

RESULTS

The developed platform is presented using a small case study in the Alpine region. Figure 4 shows a whole use case and workflow of the presented application example. The user/urban developer enters the presented platform via a Web browser (e.g. Firefox, Chrome or Internet Explorer) (Figure 4 – top screen shot). The platform includes an open street map as base map and therefore any investigation area can be presented on this platform. However, to benchmark an existing sewer system with new urban development areas the underlying sewer system model must exist. Currently, the platform supports one specific area which is the sewer system service area of the described case study, but it is easy to integrate new areas assuming all needed data are available. First, the user has to zoom to the area of interest (Figure 4(a)) where a change in land use is planned (e.g. new urban development area). As a next step (Figure 4(b)) a polygon can be drawn indicating this area. Now the benchmark can be started (Figure 4(c)). The data (user defined polygon) are sent to the WPS server and the DynaMind model starts. The subcatchments are automatically generated, connected to the existing combined sewer system model and simulated with the one at a time parameter variation method combined with all storm water event scenarios. Once all SWMM5 simulations are finished the results are sent back to the Web browser (client). The results in the demonstrated application example are geometric data of all conduits of the existing combined sewer system model, geometric data of the newly created catchments and a data set including all junctions where a change in the flooding behaviour occurred (Figure 4 – bottom screen shot). By moving the mouse pointer over a junction the results at this junction are shown in a message box. In this example the results show that the selected junction gets flooded when using a design storm event with a return period of three years (Figure 4 – result box – Old return period) during the simulation and no new catchments are connected to the system at the same time (catchment imperviousness is set to zero for all new catchments). This demonstrates the flooding behaviour of the initial system with no change in the input data. When connecting a new catchment (with fraction imperviousness 10%), the return period when flooding

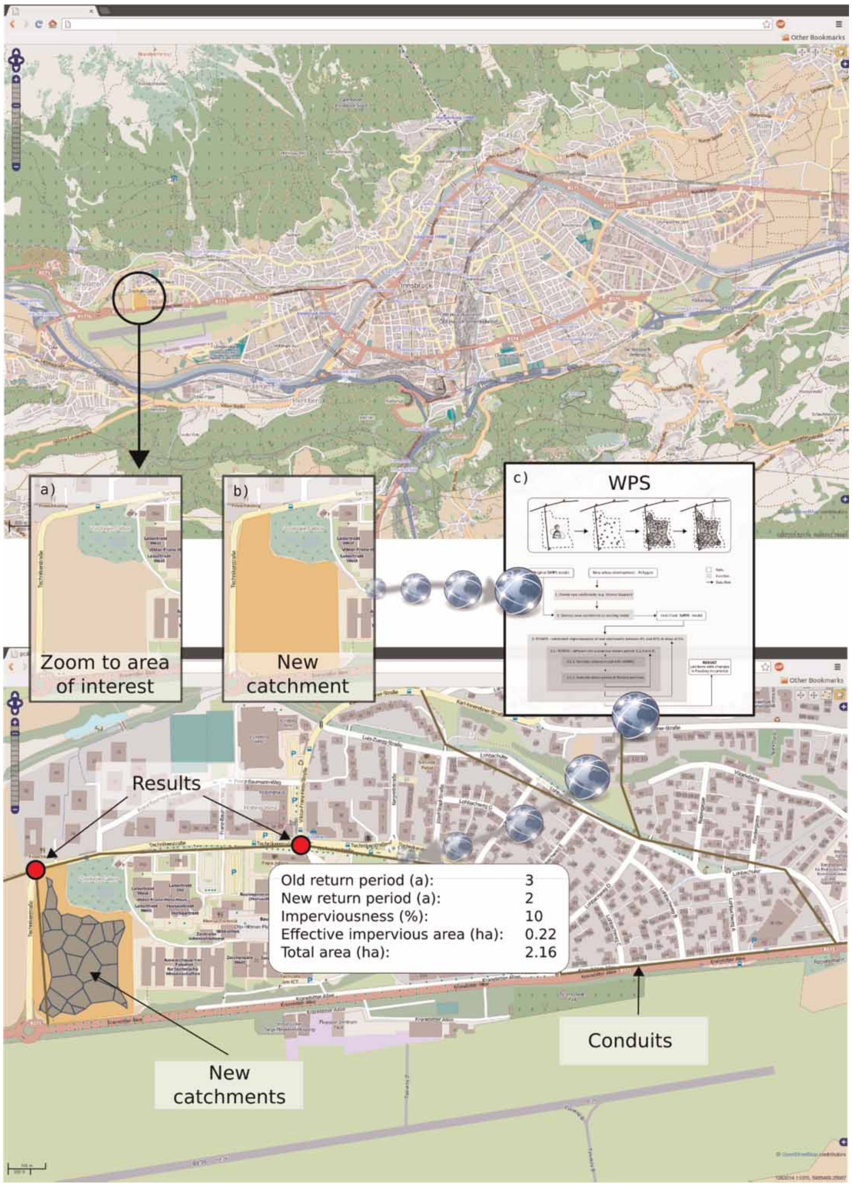

Figure 4 | Screen shots of the application example by executing a DynaMind model using WPS.

occurs changes from 3 to 2. This means that the development with a fraction imperviousness lower than 10% does not affect the flooding behaviour of the model. For higher values adaptation measures are required. In the same way, the impact of disconnection of impervious areas (e.g. when decentralized storm water management strategies are implemented) can be tested.

Figure 5 demonstrates a use-case for decentralized urban water management like water sensitive urban design (WSUD or low impact development (Bach *et al.* 2013a, b). Using the same platform stakeholders, urban planners or property owners can design, e.g. infiltration measures via the Web. Therefore, they can mark the surface, which generates runoff (e.g. roof area), and an area available for the

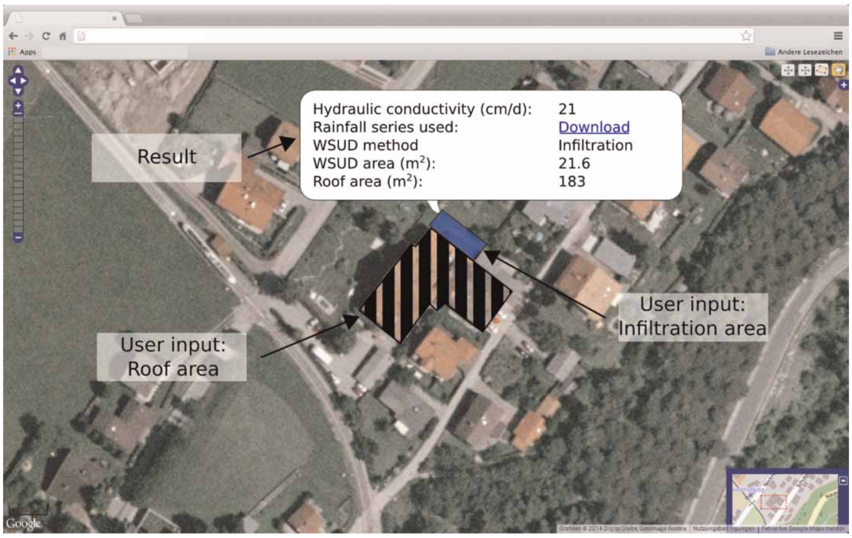

Figure 5 | Water sensitive urban design – Web based WSUD planning tool.

treatment facility (e.g. infiltration area) in their Web browser. In the background, the model identifies the nearest point for which rainfall data are available (stored on the data-server), evaluates soil characteristics (also available on the data-server) and designs the infiltration facility, e.g. the infiltration trench by determining minimum required storage volume.

DISCUSSION AND CONCLUSION

In this manuscript a prototype of a new Web based urban water management modelling platform is presented. Due to the high complexity of newly developed urban water management models and therefore the need of programming skills, this platform demonstrates an approach to simplifying the usage of these models by using Web GIS technologies. At the same time, implementation specific aspects are hidden and kept away from the user, so that they can focus on their main aim. The basic components of the platform are a client (e.g. Desktop GIS or Web browser), a data server and a Web processing service server (WPS server). Complex models (e.g. DynaMind models) are hidden behind the WPS server.

To demonstrate the platform, an application example was set up as proof of concept. The application includes a DynaMind model which represents a benchmark for the flooding behaviour of existing sewer systems when connecting new catchments to the systems. The application was demonstrated on a case study within the Alpine region, where a coarse combined sewer system model (SWMM5 model) exists. The user has to define a polygon (new catchment area/new urban development area) within a Web browser. Once the benchmark has finished, junctions with a change in the flooding return period are presented within the Web browser.

The current state of this platform is a prototype for demonstrating the idea and to prove the general concept of a Web based urban water management modelling platform. Currently, many limitations exist due to the fact that the platform is a prototype (e.g. user access control, simultaneous model simulations, calibration and validation of models) and therefore a huge amount of development and implementation work has to be done. However, we think the investigated concept does not change when implementing new features. For example access control for various users within different domains can be implemented within the GeoServer, simultaneous model simulations or long time simulations of complex models can be sourced out to well-known grid/cloud computing platforms such as Amazon cloud where the WPS server acts as a master node to guarantee the scalability of the whole platform for many users. However, one key question in integrated urban water management modelling (e.g. integrated model containing a sewer, water supply and socio-economic model) is: how can we calibrate and validate such complex models (Voinov & Shugart 2013; Bach *et al.* 2014)? Research on this topic is evolving in new directions, such as exploratory modeling and robust decision-making (e.g. Urich & Rauch (submitted)). The virtue of the presented platform is that these novel approaches are supported as well as traditional concepts.

Concluding, a Web based urban water management modelling platform can simplify the usage of complex models. This enables a group of users (e.g. stakeholders or decision-makers) with minor programming skills (needed for setting up and creating such models) to benefit from complex and interlinked model simulations.

ACKNOWLEDGEMENT

This work was funded by the *Austrian Science Fund* (FWF) in the project *DynaVIBe* P23250, by the Austrian Climate and Energy Fund as part of the Austrian Climate Research Program in the project *DynAlp* KR11AC0K00206 and by the Australian Government's 'Cooperative Research Centre for Water Sensitive Cities'. As part of this program, it will contribute to the development of an explorative computational strategic planning tool, DAnCE4Water (Dynamic Adaptation for eNabling City Evolution for Water). The authors gratefully acknowledge the financial support.

REFERENCES

Aurenhammer, F. 1991 Voronoi diagrams – a survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)* **23** (3), 345–405.

Bach, P., Deletic, A., Urich, C., Sitzenfrei, R., Kleidorfer, M., Rauch, W. & McCarthy, D. 2013a Modelling interactions between lot-scale decentralised water infrastructure and urban form – a case study on infiltration systems. *Water Resources Management* **27** (14), 4845–4863.

Bach, P. M., McCarthy, D. T., Urich, C., Sitzenfrei, R., Kleidorfer, M., Rauch, W. & Deletic, A. 2013b A planning algorithm for quantifying decentralised water management opportunities in urban environments. *Water Science and Technology* **68** (8), 1857–1865.

Bach, P. M., Rauch, W., Mikkelsen, P. S., McCarthy, D. T. & Deletic, A. 2014 A critical review of integrated urban water modelling – Urban drainage and beyond. *Environmental Modelling and Software* **54**, 88–107.

Bos, J., Brown, R., Farrelly, M. & de Haan, F. 2013 Enabling sustainable urban water management through governance experimentation. *Water Science and Technology* **67** (8), 1708–1717.

Burger, G., Sitzenfrei, R., Kleidorfer, M. & Rauch, W. 2014 Parallel flow routing in SWMM 5. *Environmental Modelling and Software* **53**, 27–34.

Deoliveira, J. 2008 GeoServer: uniting the GeoWeb and spatial data infrastructures. In: *Proceedings of the 10th International Conference for Spatial Data Infrastructure*, St. Augustine, Trinidad.

Dubois, G., Schulz, M., Skøien, J., Bastin, L. & Peedell, S. 2013 eHabitat, a multi-purpose Web Processing Service for ecological modeling. *Environmental Modelling and Software* **41**, 123–133.

Evangelidis, K., Ntouros, K., Makridis, S. & Papatheodorou, C. 2014 Geospatial services in the Cloud. *Computers and Geosciences* **63**, 116–122.

Ferguson, B. C., Brown, R. R., Frantzeskaki, N., de Haan, F. J. & Deletic, A. 2013 The enabling institutional context for integrated water management: Lessons from Melbourne. *Water Research* **47** (20), 7300–7314.

Foerster, T. & Stoter, J. 2006 Establishing an OGC Web Processing Service for generalization processes. In: *ICA Workshop on Generalization and Multiple Representation*, Portland, USA.

Hazzard, E. 2011 *OpenLayers 2.10 Beginner's Guide*. Packt Publishing Ltd, Birmingham.

Kleidorfer, M., Möderl, M., Sitzenfrei, R., Urich, C. & Rauch, W. 2009 A case independent approach on the impact of climate change effects on combined sewer system performance. *Water Science and Technology* **60** (6), 1555–1564.

Kropla, B. 2005 *Beginning MapServer:* Open Source *GIS* Development. Apress.

Mair, M., Sitzenfrei, R., Kleidorfer, M., Möderl, M. & Rauch, W. 2012 GIS-based applications of sensitivity analysis for sewer models. *Water Science and Technology* **65** (7), 1215–1222.

Mikovits, C., Rauch, W. & Kleidorfer, M. in press A dynamic urban development model designed for purposes in the field of Urban Water Management. *Journal of Hydroinformatics*.

Semadeni-Davies, A., Hernebring, C., Svensson, G. & Gustafsson, L.-G. 2008 The impacts of climate change and urbanisation on drainage in Helsingborg, Sweden: Combined sewer system. *Journal of Hydrology* **350** (1), 100–113.

Schut, P. & Whiteside, A. 2007 OpenGIS Web processing service. OpenGIS Standard, Version 1.0. 0, OGC 05-007r7. Available at: http://portal.opengeospatial.org/files/?artifact_id=24151.

Sitzenfrei, R., Möderl, M. & Rauch, W. 2013 Assessing the impact of transitions from centralised to decentralised water solutions on existing infrastructures-integrated city-scale analysis with VIBe. *Water Research* **47** (20), 7251–7263.

Urich, C. & Rauch, W. 2014 Exploring critical pathways for urban water management to identify robust strategies under deep uncertainties. *Water Research* **66**, 374–389.

Urich, C., Burger, G., Mair, M. & Rauch, W. 2012 DynaMind – A Software Tool for Integrated Modelling of Urban Environments and their Infrastructure. In: *10th International Conference on Hydroinformatics*. Hamburg, Germany, p. 18.

Urich, C., Bach, P. M., Sitzenfrei, R., Kleidorfer, M., McCarthy, D. T., Deletic, A. & Rauch, W. 2013 Modelling cities and water infrastructure dynamics. *Proceedings of the ICE-Engineering Sustainability* 166 (5), 301–308.

Voinov, A. & Shugart, H. H. 2013 'Integronsters', integral and integrated modeling. *Environmental Modelling and Software* **39**, 149–158.

Wenjue, J., Yumin, C. & Jianya, G. 2004 Implementation of OGC web map service based on web service. *Geo-spatial Information Science* **7** (2), 148–152.

Willuweit, L. & O'Sullivan, J. J. 2013 A decision support tool for sustainable planning of urban water systems: Presenting the Dynamic Urban Water Simulation Model. *Water Research* **47** (20), 7206–7220.

Zhou, Q., Halsnæs, K. & Arnbjerg-Nielsen, K. 2012 Economic assessment of climate adaptation options for urban drainage design in Odense, Denmark. *Water Science and Technology* **66** (8), 1812–1820.

First received 28 February 2014; accepted in revised form 8 July 2014. Available online 4 August 2014

Part C

APPENDIX

20

CURRICULUM VITAE

Michael Mair

Msc, Bakk.techn.

General

Name **Michael Mair**.
Birthday **11.01.1983**.

Education

1990-1994 **Volkschule**, *Rum*, Austria.
1994-1998 **Bundesrealgymnasium**, *Adolf-Pichler-Platz 1*, Innsbruck.
1998-2003 **HTL, Key course element: Structural engineering**, *Trenkwalderstraße 2*, Innsbruck.
2003-2007 **Bakk.techn.**, *Computer Science*, University Innsbruck.
2007-2011 **M.Sc.**, *Computer Science*, University Innsbruck.
since 2011 **ongoing Phd**, *Unit of Environmental Engineering*, faculty of Civil Engineering, University Innsbruck.

Experience

Vocational

2006–2007 **Project collaborator**, *Distributed and parallel system group, Faculty of MIP*, University Innsbruck.

since 2008 **Project collaborator**, *Unit of Environmental Engineering, Faculty of Civil Engineering*, University Innsbruck.

\- **Agriculturist**, *Farm at Garneid 1 6063 Rum.*

Miscellaneous

2000 **Alpine Mayreder Bau GmbH**, *1 month.*
2001 **Alpine Mayreder Bau GmbH**, *1 month.*

Languages

German **native speaker**
English **good**

Garneid 1 – 6063 Rum
0664/1943880 • ☎ *0512/205259* • ✉ *michael.mair@uibk.ac.at*

Computer skills

Programming techniques: Extensive knowledge of programming techniques in parallel computing. (Cluster and Grid)

OS: Linux and Windows

Programming skills: C, C++, Java and Python

Master thesis

title: *A model independent framework for parallel calibration*

supervisors: Dr. Hans Moritsch

description: In many areas numerical models are getting more and more important for analysing and predicting the behaviour of real world systems.Calibrating a new model is one essential part during model development to guarantee a certain accuracy of model simulation output. The aim is to minimize the deviation between model prediction and measured data of the real world system by adapting model parameters. The deviation is represented by one or several objective functions. In mathematical sense this represents an optimisation problem.
This thesis describes the general concepts of numerical model calibration with focus on developing and implementing a model independent and generalized framework for parallel calibration, called CALIMERO.Model calibration is a computational intensive task. With CALIMERO the runtime of the model calibration process can be decreased by using all available cores on state-of-the-art multi core systems.The framework is benchmarked on calibrating urban water management models by using different objective functions, optimisation algorithms and programming languages (C++ and Python).
The benchmark results show that a speedup of the calibration process can be reached in all test cases. Depending on the runtime and implementation (sequential or parallel) of used modelsthe speedup varies. By increasing the runtime of the numerical model simulation the speedup of parallel calibration increases and vice versa. Moreover it showed that even with nested parallelism (parallel optimisation algorithm and parallelmodel simulation) an increased speedup is recognized.

Bachelor thesis

title: *Internal Workflow Representation / Workflow Conversion and Workflow Prediction in ASKALON*

supervisors: Univ.-Prof. Dr. Thomas Fahringer

Garneid 1 – 6063 Rum
0664/1943880 • ☎ 0512/205259 • ✉ michael.mair@uibk.ac.at

description For the application developer, it is usually hard to use the whole power of Grid computing. The ASKALON Grid application development and computing environ- ment was developed in order to provide the user with a powerful Grid composition and execution environment which can facilitate the process of creation and execu- tion of workflow application on the Grid. ASKALON uses the workflow representa- tion model based on a high-level XML-based language called AGWL. AGWL cannot express the runtime information needed for keeping track of the correct execution state, and for actual execution of workflow tasks. In order to be capable of workflow processing on real resources in an effective and flexible way, a low-level workflow representation must be applied. Therefore the Internal Workflow Representation of ASKALON (IWR) was developed, which fulfills the above stated requirements for a dynamic workflow representation.We present the workflow converter and the workflow predictor implemented in ASKALON Grid environment. The workflow converter is responsible for converting workflows represented in AGWL into equiv- alent workflows represented in IWR. The workflow predictor supports the dynamic and flexible workflow execution model based on a full-graph scheduling approach, by implementing several dynamic conversion techniques for IWR performed in the presence of incomplete information. Among other techniques, the workflow predic- tor unrolls sequential and parallel loops, and replaces conditional constructs (if, switch) with their single variants (branches). As both of these cases may require some explicit assumptions to be done (e.g., regarding the evaluation of some condi- tions), this conversion process is referred to by us as workflow prediction.

Publications

[1] Manfred Kleidorfer, Guenther Leonhardt, Michael Mair, David McCarthy, Heiko Kinzel, Wolfgang Rauch, et al. Calimero-a model independent and generalised tool for autocalibration. In *Proceedings of the 8th Int. Conf. on Urban Drainage Modelling (8UDM) and 2nd Int. Conf. on Rainwater Harvesting and Management*, pages 1–9, 2009.

[2] M Mair, W Rauch, and R Sitzenfrei. Spanning tree based algorithm for generating water distribution network sets by using street network data sets. *World Environmental & Water Resources Congress EWRI*, pages 1–5, 2014.

[3] M Mair, R Sitzenfrei, M Kleidorfer, M Moderl, and W Rauch. Gis-based applications of sensitivity analysis for sewer models. *Water science and technology*, 65(7):1215, 2012.

[4] M Mair, R Sitzenfrei, M Möderl, and W Rauch. Identifying multi utility network similarities. *Proceedings of the World Environmental and Water Resources Congress, Albuquerque, New Mexico*, 20:2012–24, 2012.

[5] Michael Mair, Jun Qin, Marek Wieczorek, and Thomas Fahringer. Workflow conversion and processing in the askalon grid environment. *2nd Austrian Grid Symposium*, pages 67–80, 2007.

[6] Michael Mair, Wolfgang Rauch, and Robert Sitzenfrei. Improving incomplete water distribution system data. *Procedia Engineering*, 70:1055–1062, 2014.

[7] Michael Mair, Robert Sitzenfrei, Manfred Kleidorfer, and Wolfgang Rauch. Performance improvement with parallel numerical model simulations in the field of urban water management. *Journal of Hydroinformatics*, 16(2):477–486, 2014.

Garneid 1 – 6063 Rum
0664/1943880 • ☎ *0512/205259* • ✉ *michael.mair@uibk.ac.at*

[8] M Möderl, C Hellbach, R Sitzenfrei, M Mair, A Lukas, E Mayr, R Perfler, and W Rauch. Gis based applications of sensitivity analysis for water distribution models. *World Environmental and Water Resources Congress*, pages 129–136, 2011.

[9] M Möderl, R Sitzenfrei, M Mair, H Jarosch, and W Rauch. Identifying hydropower potential in water distribution systems of alpine regions. *Bridges*, 10:9780784412312–314, 2014.

[10] R Sitzenfrei, M Kleidorfer, M Meister, G Burger, C Urich, M Mair, and W Rauch. Scientific computing in urban water management. In *Computational Engineering*, pages 173–193. Springer International Publishing, 2014.

[11] R Sitzenfrei, M Mair, K Diao, and W Rauch. Assessing model structure uncertainties in water distribution models. *World Environmental & Water Resources Congress*, 2014.

[12] R Sitzenfrei, M Mair, M Moderl, and W Rauch. Cascade vulnerability for risk analysis of water infrastructure. *Water Science & Technology*, 64(9):1885–1891, 2011.

[13] R Sitzenfrei, M Möderl, M Mair, and W Rauch. Modeling dynamic expansion of water distribution systems for new urban developments. *Bridges*, 10:9780784412312–320, 2014.

[14] Christian Urich, Gregor Burger, Michael Mair, and Wolfgang Rauch. Dynamind–a softwaretool for integrated modelling of urban environments and their infrastructure. *Proceedings of the 10th International Conference on Hydroinformatics HIC*, 2012.

Garneid 1 – 6063 Rum
0664/1943880 • ☎ 0512/205259 • ✉ michael.mair@uibk.ac.at

21

PRESENTATION LIST

Presentation title	Event	Place	Date
DynaVIBe-Web: A web based water distribution network generator	EWRI	Austin, Texas	20.05.2015
Modelling the impact of Green/Blue Infrastructure	GBC	Obergurgl, Austria	30.03.2015
Spanning Tree Based Algorithm for Generating Water distribution Network Sets by using Street Network Data Sets	EWRI	Portland, Oregon	03.06.2014
Web-GIS Anwendungen in der Siedlungswasserwirtschaft	Aqua Urbanica	Innsbruck, Austria	23.10.2014
Improving incomplete water distribution system data	CCWI	Perugia, Italy	04.09.2013
Identifying multi utility network similarities	EWRI	Albuquerque, New Mexico	21.05.2012
Performance of auto-calibration algorithms in the field of urban drainage modelling	UDM	Belgrad, Serbia	04.09.2012
Performance improvement with parallel numerical model simulations in the field of urban water management	HIC	Hamburg, Germany	15.07.2012
GIS Based Applications of Sensitivity Analysis for Sewer Models	ICUD	Porto Alegre, Brasil	15.09.2011

Table 21.1: Presentation list — *Presented at event:* GBC = Mini-Symposium on Green-Blue-Cities; EWRI = World Environmental and Water Resources Congress; CCWI = International Conference on Computing and Control for the Water Industry UDM = International Conference on Urban Drainage Modelling HIC = International Conference on Hydroinformatics ICUD = International Conference on Urban Drainage

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

`http://code.google.com/p/classicthesis/`

Final Version as of March 21, 2016 (`classicthesis` Version 0.99).