Journal of Water Supply: Research and Technology—AQUA | 64.5 | 2015

Water distribution model for decision-making with updatable data links

Raido Puust, Margus Koor and Anatoli Vassiljev

ABSTRACT

From February 2011 to December 2012, the Tallinn Water Company, AS Tallinna Vesi (ASTV), aimed at improving the previous hydraulic model creation procedures so that the model can be operatively updated through the available geographic information system, client information system (Navision), supervisory control and data acquisition system (SCADA). The goal of the project was to analyse various model building alternatives so that data reload can be easily managed by ASTV. The model rebuild is necessary because of continuous changes in the water network (pipe renewals, new developments, demands, etc.). The project was divided into two phases. Phase 1 was mainly about hydraulic network model build-up to the pre-calibrated state where all possible data sources were created and data are loaded into hydraulic modelling software. Phase 1 was finished in May 2012. Phase 2 started in summer 2012, and included pressure measurement point selections, data validation and model calibration. The final production model provides the possibility of using the model for real decision-making.

Key words | calibration, genetic algorithm, geographic information system (GIS), hydraulic network model, supervisory control and data acquisition (SCADA)

Raido Puust (corresponding author)
Anatoli Vassiljev
Tallinn University of Technology,
Ehitajate tee 5,
Tallinn 19086,
Estonia
E-mail: *raido.puust@ttu.ee*

Margus Koor
AS Tallinna Vesi,
Ädala 10,
Tallinn 10614,
Estonia

INTRODUCTION

This paper is an extension of the one presented at the 12th International Conference on 'Computing and Control for the Water Industry – CCWI2013' (Koor *et al.* 2014). Hydraulic model update is quite a time-consuming task because it involves several data sources that should be carefully checked against possible faults before any model can be built. Depending on the available tools, hydraulic network model (HNM) can be built up by directly pulling various data sources into hydraulic modelling software (HMS) or incorporating hydraulic modelling capabilities into geographic information system (GIS) application (Martinez *et al.* 2005; Speight *et al.* 2011). City-scale water network modelling has been recently described in several research papers (Crozier *et al.* 2012; Loubser *et al.* 2012). Depending on the available tools, techniques and also considering whether the Water Company is willing to pay for additional software developments and/or licences, various routes can be taken. Although today's computers and database management software can easily generate 'all pipe' models for any city or municipality, the uncertainty and complexity of the water distribution system makes it difficult to predict its performance (Perelman & Ostfeld 2011). Quite often, some simplifications are made to overcome the complexities like real time changes in network data (open/closed valves, new pipe data, etc.). Those simplification routes are made to predict failure scenarios, detection of sources of contamination intrusions, sensor placement locations and surge analysis (Jung *et al.* 2007; Almeida & Ramos 2010). Using various algorithms, the simplification can be made with HNM (when a skeletonization module is available) or at GIS database level. The advantage of doing it in HNM is that various control strategies can be used to validate exactness of the reduced model, including hydraulic as well as energy analysis (Paluszczyszyn *et al.* 2011). Although simplified models can be more easily managed, it should be noted that oversimplified models can lead to 'too good' solutions

doi: 10.2166/aqua.2014.038

as with water quality models where skeletonization usually decrease the average water age in pipes and therefore tend to give incorrect results (Bahadur *et al.* 2006). Several approaches have been published how to simplify network topology that inherits a wide range of operating conditions from the parent, 'all pipe' model (Ulanicki *et al.* 1996; Grayman & Rhee 2000).

The current study started in 2011 when AS Tallinna Vesi (ASTV) was interested in updating their water network model with the aim that it can be managed (updated) by their own personnel in the future for easier model build-ups and calibration. Phase 1, that lasted to May 2012, mostly dealt with database connections, model skeletonization and demand aggregation (Koor *et al.* 2012). Phase 2 started in summer 2012 and focused on pressure measurement point selections, data validation and model calibration (Koor *et al.* 2014). Tallinn University of Technology was involved in ASTV full-scale city model creation and calibration also from 2001 to 2004 when for model building, Bentley WaterCAD (then Haestad Solutions) was used; calibration was purely done in EPANET with custom calibration tools (Ainola *et al.* 2000; Vassiljev *et al.* 2005; Koppel & Vassiljev 2009, 2012). For the current project, Bentley WaterGEMS V8i (SELECTseries 3) was used from start to finish, including calibration with WaterGEMS built-in tool Darwin Calibrator (Bentley Systems 2011). The genetic algorithm has had many success stories since its first appearance in water network-related problems (Simpson *et al.* 1994; Savic & Walters 1995, 1997) including real, city-scale calibration studies (Randall-Smith *et al.* 2006). Although the genetic algorithm takes a considerably longer time to calculate and needs a great deal of computational power compared to other methods (Vassiljev & Koppel 2012), the main advantage for ASTV has been its out-of-the-box solution inside their already own hydraulic modelling package as well as its easier set-up routines for model calibration.

This paper details city-scale water network model creation procedures, emphasizing model calibration topics and its representation with close-to-real-life model components so that it can be used for daily decision-making tasks. During the case study, various database/software-related challenges arose. Custom routines to overcome those issues are discussed in the current paper.

METHODOLOGY/PROCESS

Database-driven model creation procedure is based on a real water network data changing with time. For model recreation it is unwise to waste time on routine steps or activities that can be automated at database level. Losing a valuable time factor is especially inefficient in large water networks where thousands of pipes and consumer accounts exist. Tallinn's pipe infrastructure is the biggest in Estonia, serving almost 392,000 people (that is about one-third of the Estonian population). Although database-based information handling is nothing new in companies of that scale, the development of a well-established hydraulic model creation procedure is quite often the opposite. At least once model creation stage has happened in most moderate-to-large scale water companies. Every new update, when done manually, is time-consuming and may introduce new errors that are difficult to perceive and to eliminate because of the scale of the model/database. Near real-time rehabilitation and optimization tasks need an updated model from the current timeframe and therefore optimized hydraulic model update/rebuild procedure plays an important role. Tallinn's water network has 85,058 pipes (950 km), 35 pump stations (including 11 pressure booster stations) and 11 main pressure zones. There is one surface water resource (Lake Ülemiste) and several groundwater resources to satisfy customers' water needs (Figure 1).

Different types of data are kept in various databases (infrastructure, customer demand data and four different supervisory control and data acquisition system (SCADA) systems). Several approaches exist of how to combine these databases (Koor *et al.* 2012). In current studies, GIS and HNM are kept separately and simplifications are done already at GIS database level. The simplification workflow is illustrated in Figure 2.

To ensure that the simplification does not involve large discrepancies in the sense of hydraulic and energy analysis, three main types of simplifications are done: (a) dead end demands are carried over to the next node; (b) dead end links are removed (no demand); and (c) similar pipes are merged that have a common diameter, installation year and material. Model update is regularly needed because of timely basis pipe renewals, new developments as well as

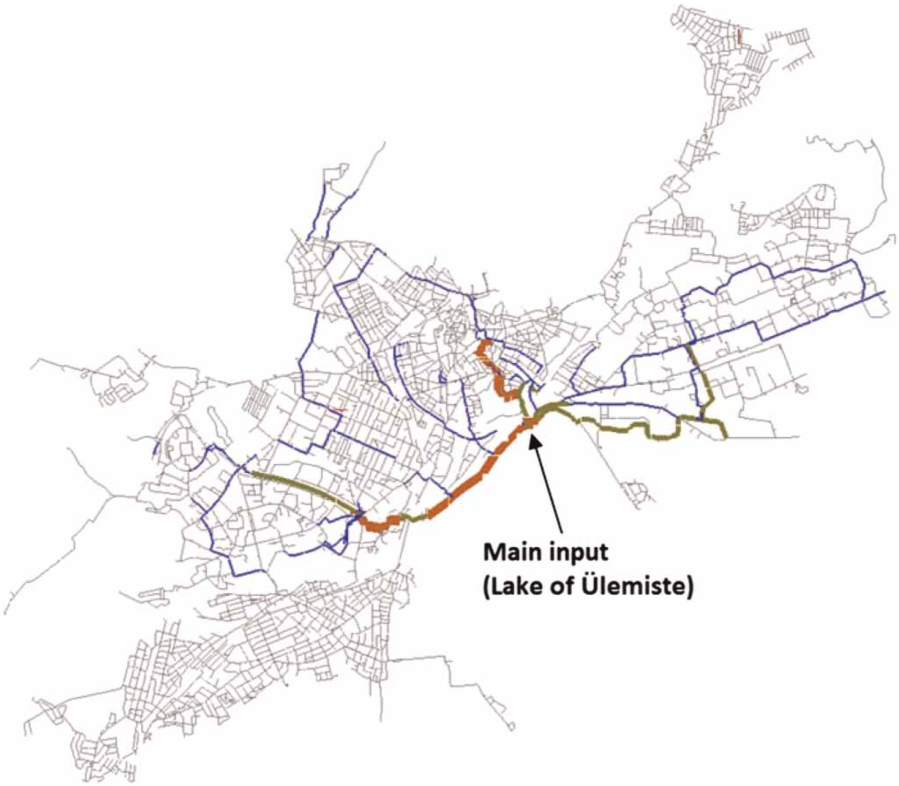

Figure 1 | Hydraulic water network model with highlighted transmission lines.

changes in demands. For example, in Estonia the average yearly water demand has dropped about three times within the last 10 years. Meanwhile, the average leakage has dropped about twice (Figure 3).

Also, as people tend to change the place they live, the demand pattern can change over time (Koppel & Vassiljev 2009, 2012). The simplification was divided into three stages (Level 0–Level 3) and the difference in the number of main elements is shown in Table 1.

ASTV uses Bentley's WaterGEMS v8i HMS (Bentley Systems 2011). Different approaches were discussed about which functionalities (connecting directly with GIS database, using sub-model feature or simple EPANET *.inp file format) can be used for effective model build-up. During the project the best possible solution for network topology import that arose was EPANET *.inp file format (Rossman 2002) which was generated with GIS SQL database tools and custom queries. The main reason to use GIS generated EPANET file format was that the company's GIS platform was not able in 2012 to produce an industry standard shapefile format. In 2014, the GIS platform was updated and it includes shapefile export/import functionality that should improve the overall ease of model updatability, however, it needs further testing. Before EPANET network was imported into the WaterGEMS software package, all component libraries were added (11 material definitions, 6 24-h demand patterns with daily and monthly coefficients).

ADDITIONAL DATA/MODEL MANAGEMENT

After model import (pipes, nodes, hydrants, pumping stations and valves) into WaterGEMS software, additional information was included with pipe/node elements based on pressure zones. In addition to zone information, all boundary data (flow inputs) were included into the model. All sub-models were test-run to ensure the integrity of the model. During the build-up of the model, various pumping station representations

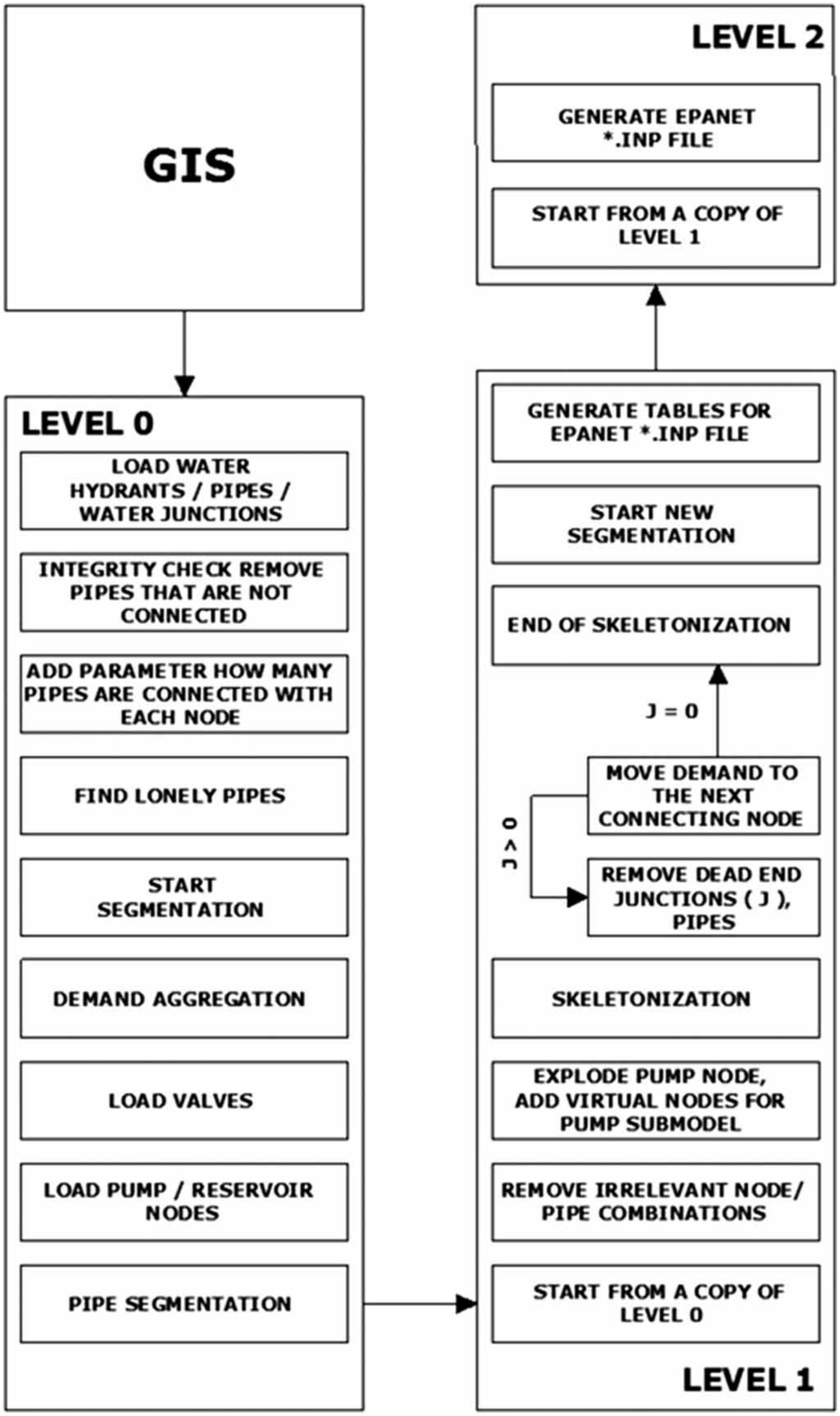

Figure 2 | Main simplification procedures at SQL-based database level.

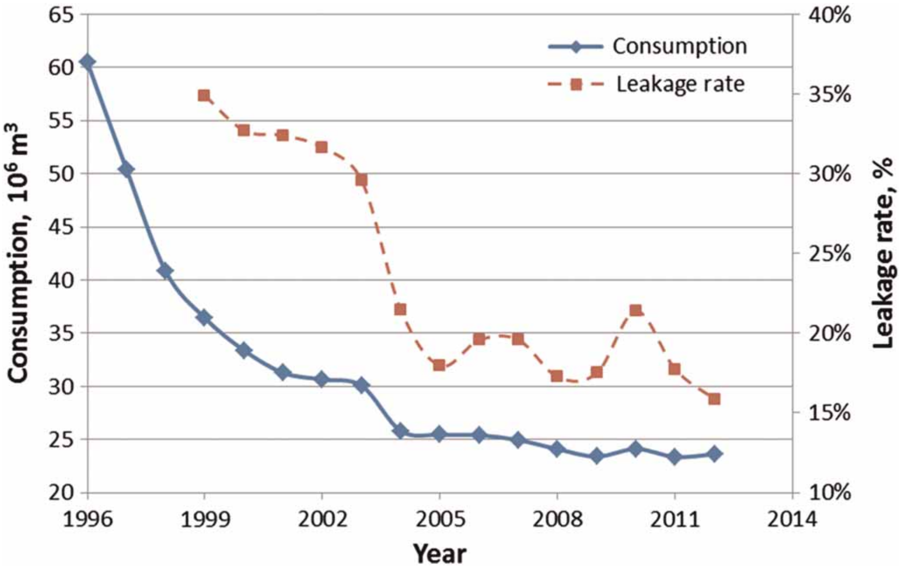

Figure 3 | Average yearly water consumption and leakage rates in Tallinn from 1996 to 2012.

Table 1 | Number of main elements at various database levels

	Level 0	Level 1	Level 2
Number of segments	933	8	8
Number of pump connections (input/output)	84	170[a]	170[a]
Number of hydrants	5,345	5,096	4,761
Client connections	22,239	18,566[b]	17,576[b]
Number of pipe segments	85,058	31,731	29,971
Number of isolation valves	36,837	8,265	7,740
Time to generate	~1 h	~15 min	~1 min

[a]The number of pump connections increases because of pump node expansion into separate input/output nodes in pump station.
[b]Some clients do not belong to the main Tallinn water system and therefore this number is reduced.

were considered, as can be seen in Figure 4, including also groundwater pumping stage if present at source or pressure booster stations.

The detail level of pumping stations models may have serious drawbacks for calculation (calibration) speeds and/or for the convergence of the simulation results. Option 1 (Figure 4) might be the true representation of the pumping station but it will tremendously increase the calculation times and the software may not find the solution at some times steps at all (Koor *et al.* 2012). In normal situations, simpler pumping station representations are used. For example, for calibration, Option 3 (Figure 4) is used where the pumping station is represented by a simple reservoir element with a measured head pattern (derived from pumping station measurements). Option 2 would be much more preferable for calibration but again, during the calibration, the preliminary hydraulic calculation might not converge and might cause a backflow through the pump station. It was especially problematic in Zone 11 where a large number of pumping stations are present (see Figure 5). Therefore, Option 3 was considered for calibration studies because of the easier water balance calculations, and later on, Option 2 was used with a calibrated model (including pump relative speed patterns or fixed pump head behaviour).

Several mathematical models exist to predict the initial pipe roughness values based on pipe age, material and soil type (Duchesne *et al.* 2013). Because of ASTV flushing strategies it was decided that it is better to leave the estimation of pipe roughness in the calibration stage, where pipe roughness lower and upper limits are defined considering pipe materials as well as their age. One of the reasons was that ASTV carries out regular pipe flushing and therefore a direct link between pipe age and roughness was not assumed before the calibration.

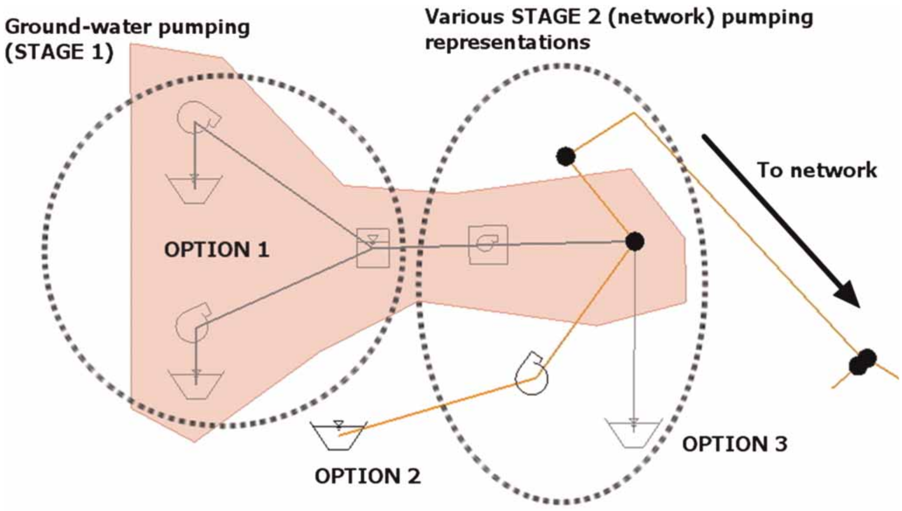

Figure 4 | Various representations of pumping stations.

Figure 5 | Eleven pressure zones at Tallinn City water network.

SELECTION OF PRESSURE MEASUREMENT POINTS

Any calibration study needs good quality measurement data. The Tallinn City network has 11 main pressure zones that were measured independently by additional pressure loggers, including inflows/outflows (Figure 5). Each pressure zone was measured for about 1 week which included all the weekdays at least once. The size of the pressure zones varied quite a lot, ranging from 110 pipes (smallest) to 11,200 pipes (largest) but the number of additional pressure loggers was limited. In addition to zone inflows/outflows and pressure measurement stations (SCADA-based), temporary pressure loggers were used in every zone for about 1 full week to record pressures with a 1-minute time step. Before the pressure measurements, simple, model-based sensitivity analyses were carried out at zone level to find out the best measurement point locations. As measurements were carried out at hydrants, also the sensitivity analysis was carried out at hydrant level and so-called sensitivity coefficient was calculated for every hydrant. Sensitivity coefficient was assumed to be a combined value of two different approaches. First, roughness sensitivity was researched at maximum demand hour at hydrants (08:30 am). Second, fire flow test (model based) was carried out with all hydrants to find the most sensitive areas for possible measurement points. Sensitivity maps were created in conjunction with WaterGEMS User Defined parameter values for nodes. Sample results can be seen in Figures 6 and 7.

It can be easily seen that when comparing two different sensitivity analyses (Figures 6 and 7), two different network maps are derived. For example, the most sensitive points with fire flow test are mostly shown at dead-ends. For final pressure logger locations there are various, endless number of combinations to choose from, especially when the amount of loggers is limited (15 in our case) and large zones should be measured. Calibration results are greatly affected by the number of pressure measurements per pipe kilometre or simply per number of pipes/nodes/hydrants. Therefore, simple sensitivity analysis helps us to draw attention to some particular sub-areas in the network and pressure logger locations can be chosen more calibration-safely. The same procedure as described here was carried out with each zone (11 altogether) and a network map as well as final hydrant selection (Zone 11, Figure 8) was shared with ASTV team members.

Measurements were carried out during 2012 (spring to autumn) and the measurement period varied from 7 to 20 days (Table 2). The overall number of different measurement periods is highly dependent on pressure zone size (in terms of nodes/pipes) and the number of available pressure loggers. It should be also noted that with each period a considerable amount of man-hours are needed to carry out the relocation of loggers. Although 7 days was planned for each zone, longer periods were gathered at some zones

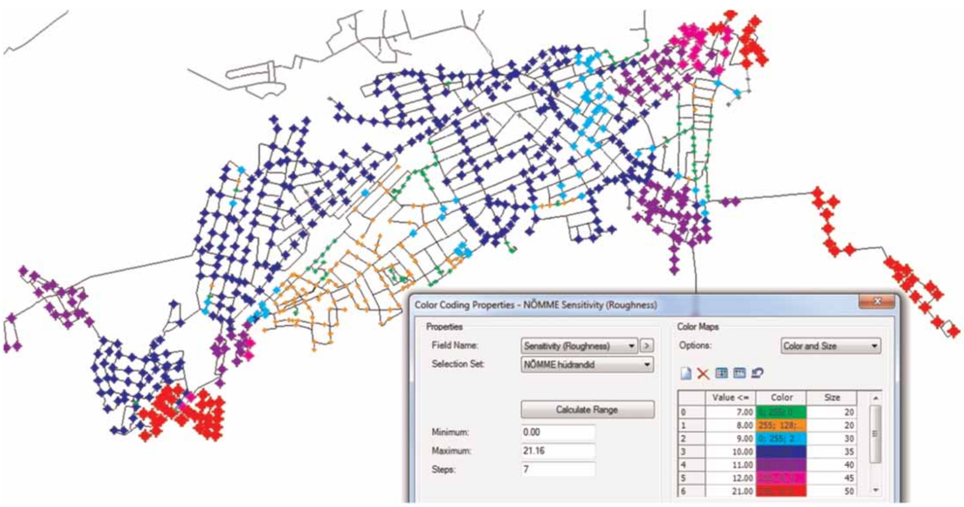

Figure 6 | Roughness sensitivity map for Zone 11. Larger node sizes indicates more sensitivity.

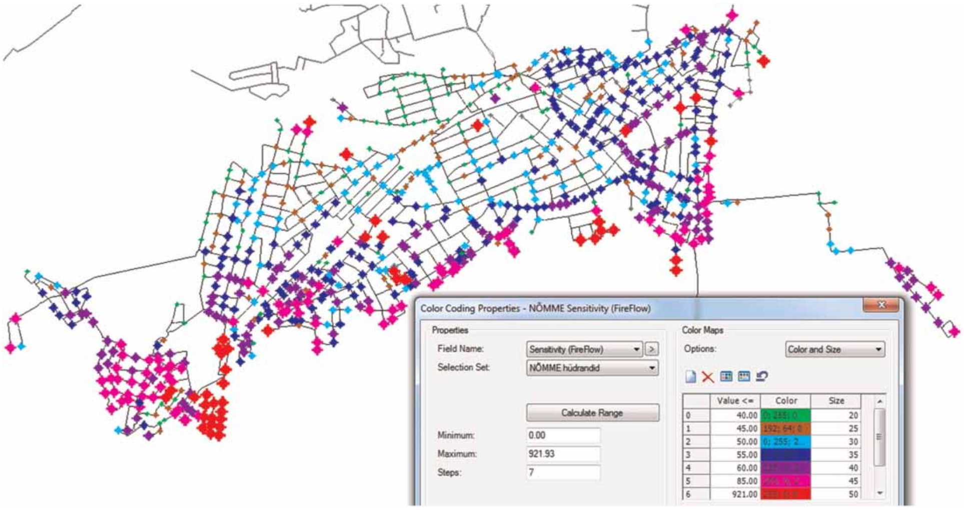

Figure 7 | Fire flow sensitivity map for Zone 11. Larger node sizes indicates more sensitivity.

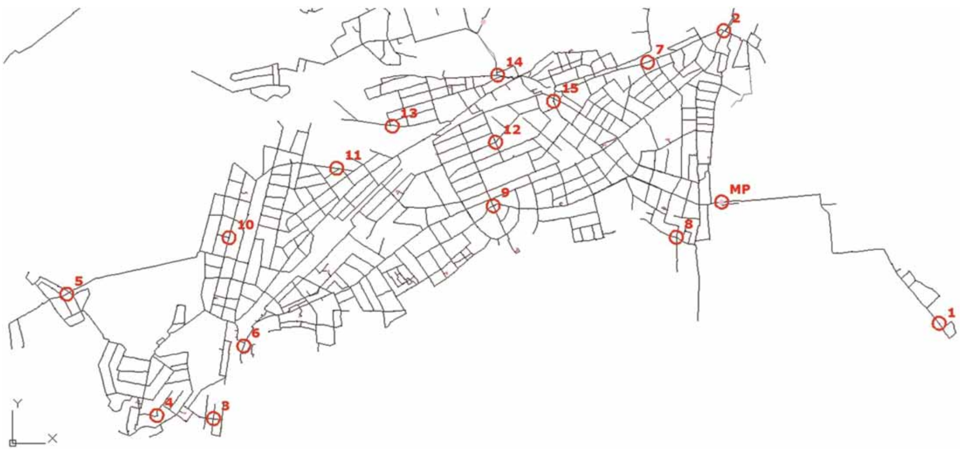

Figure 8 | Pressure measurement points for Zone 11.

Table 2 | Measurement periods

Zone	Zone name	Measurement period	Full days
1	Merivälja	19.06.2012–08.07.2012	20
2	Pirita	07.06.2012–16.06.2012	10
3	Kose	12.07.2012–18.07.2012	7
4	Lasnamäe-III	26.05.2012–03.06.2012	9
5	Lasnamäe-II	04.05.2012–17.05.2012	14
6	Linna-III	25.08.2012–02.09.2012	9
7	Toompea	06.09.2012–12.09.2012	7
8	Linna-II	21.04.2012–01.05.2012	11
9	Taela	Not measured	n/a
10	Mustamäe-Õismäe	11.04.2012–18.04.2012	8
11	Nõmme	18.09.2012–27.09.2012	10

due to vacation seasons. Therefore, the optimum measurement period, assuming 2-day relocation period was 99 days but it took altogether 105 days. The measurement period can be shortened with more loggers but in this case the relocation time may go up because more man-hours are needed to maintain and relocate the loggers. Zone 9 was not measured because of the different type of hydrants in that region and additional pressure loggers were not used there. That zone was also excluded from later calibration studies.

All measurements were recorded with 1-minute time step. Before importing that data into the calibration module, the basic filtering and averaging was carried out. Final measurement data were with 1-hour time step. If some anomalies were discovered within measurements, those were not accounted for in the final data selection. For example, there was one pressure logger that showed most of the time higher pressures in the network than at the zone inflow point (pumping station). The cause of the problem could be a faulty sensor as well as wrong elevation (placement) height recording. As final measurement data should be imported directly into WaterGEMS calibration module, ModelBuilder was used to connect with the filtered measurement data source. For online loggers (pump flows/pressures, pressure measurement stations) two different import alternatives were evaluated. At first, WaterGEMS module SCADAConnect was considered and tested as the most obvious choice to carry over the historical data based on zone measurement period. It turned out that the capabilities of that module were not suitable for the current workflow: (a) there was no possibility of saving the signal settings for future updates, for example, when a model is rebuilt by ModelBuilder links into a new file, there is no possible way to import predefined signal definitions; and (b) there was no possibility of combining various time steps into one calibration study. The latter issue was most problematic for planned workflow as can be seen in later sections where the overview of calibration module is given. As an alternative, all online data (historical SCADA data) were imported into the calibration module in the same way as with off-line measurements. For that purpose, it was necessary to prepare an additional data table that included flows, pressures from the pumping stations and from the fixed measurement stations. ModelBuilder was used to connect with that data table and, as such, all the necessary measurement data were gathered into WaterGEMS calibration module called Darwin Calibrator.

MODEL CALIBRATION

At the very first stage of the project, it was decided that the main goal for the current study is to use out-of-the-box tools that are readily available for ASTV. Previously it has been shown that the Levenberg–Marquardt optimization algorithm works much faster than genetic algorithm (Koppel & Vassiljev 2009; Vassiljev & Koppel 2012, 2013) but as the Darwin Calibrator calibration module resides inside the WaterGEMS environment, it was the most obvious choice.

In addition to measurement data, various other settings should be defined before calibration can be started. Calibration was carried out for pipe roughness values (Darcy–Weisbach) as well as for emitter coefficient values (leakage representation). Therefore pipe and node groupings were created. For pipes, plastic and metal pipe groupings were created separately considering only the diameter aspect with a 50 mm increment. It was found that a smaller increment did not add value for the calibration procedure considering calibration speed and there is really no great difference if the pipe belongs in the 100 or 110 mm group. Maximum roughness value for metal pipes was up to 75% of their diameter with an increment of 10% of their maximum value. Plastic pipe groups were separated from the metal pipes because of much lower maximum possible roughness (up to 52 mm with increment of 12 mm) and it was assumed that the maximum value does not depend on pipe diameter. The same principles were used in all zones to create pipe groupings.

Leakage node groupings were purely based on measurement locations. Due to the fact that pipes are mostly oversized in the whole city area and because of that very low flows exist in the system there is no point to search a leakage far away from the measurement point. Therefore the groupings were created around measurement point. The closest node to the measured hydrant location was considered as a leakage candidate. Although each node group has one single leakage candidate the results of leakage calibration should be expanded over all other nodes that are in that region. Various methods of how to calibrate leakages have been reviewed by Puust *et al.* (2010). Leakage calibration in the current study was defined as a search problem of optimal emitter coefficient for each node grouping. The ranges for emitter coefficient were assumed so that any one maximum coefficient value can cause about 10% of additional outflow from the zone.

In addition to pipe/node groupings that help to keep the calibration times within a reasonable timeframe, the increment values that drive all possible roughness/emitter coefficient values, play an important part. Roughness increment has been chosen so that a maximum 15 different values (metal pipes) are considered in one particular pipe group. Emitter coefficient increment has been selected so that 30 different values fall into the demand group. The number of all possible groupings and possible values greatly affects the speed of calibration. In the current study, pipe roughness and emitter coefficients were calibrated in a separate calculation but multiple times with different initial values. The last thing to consider before the calibration run is to check the parameters of the calibration algorithm. Most values were kept at their default values suggested by the software and only slight changes were made such as increasing the maximum trials (5,000,000) and population size (200).

After setting up all the necessary parameters the sensitivity analysis of measurements was carried out. There are various research studies of how to pick the most sensitive data for final calibration (Bowden *et al*. 2002). In the current research, a different approach was used. Considering the fact that no real fire hydrant tests were carried out to get additional pressure data, the whole measurement cycle was imported into the calibration module and so-called all-data-calibration was carried out at zone level. For example, if a particular zone was measured for 10 days, 10 days $\times$ 24 hours $=$ 240 time steps were fed into the calibration procedure. Of course, such calibration has a dramatic effect on calculation speed and our main purpose was to test the robustness of the calibration procedure itself and not to optimize the calculation time. Virtual Machine (64-bit, 4GB memory) was used to carry out calculations. The so-called sensitivity calculation took about 1 hour (small zones) to 2 days (large zones). After calculation, the general model agreement with the measurements was drawn by WaterGEMS tools.

Based on the measurement sensitivity results the error of observed and simulated HGL values are ordered and divided into smaller groups. Additional pre-calibration studies are carried out to find out how the number of good measurement snapshots affects the final result (in the sense of error and calculation time). It was concluded that six to 12 different measurement snapshots are enough to carry out the final calibration study. In addition to pipe roughness evaluation over all the measurement data, the analogous analysis was carried out with leakage studies to get the overall feeling which data are good enough to use in the final calibration. The final calibration was divided into three main stages. Table 3 shows the calibration procedure that was applied for each separate zone.

Table 3 | Three-stage calibration procedure

Calibration stage	Calibration study	Method	Note
Stage 1	Roughness calibration	Pick a value between the boundaries	Roughness of the new pipes is assumed as starting point
Stage 2	Leakage calibration	Pick a value between the boundaries	Results from Stage 1 where used as a starting point
Stage 3	Roughness calibration	Multiply the roughness value in between 0.1–2.0	Results from Stage 2 where used as a starting point

In general, the error in between measured and simulated values was reduced at every stage. As up to 12 measurement snapshots were used in every calibration run, the calculation time was reasonable, ranging from a few minutes to half an hour depending on the zone size. Table 4 shows an overview of calibration results in the sense of maximum errors in between measured and simulated nodal pressures.

In the current project three-stage calibration is used with a separate portion of measurement data for each calibration stage. Because of the sensitivities within pipe roughness grouping might be different than with leak node grouping (emitter coefficients); different data portions for those sub-calibrations can be used.

In general, it can be clearly seen that calibration results are better for smaller zones (more measurements per overall unit of pipes). Maximum errors in Table 4 are caused by some particular pressure logger. Attention should be drawn to the fact that as this logger was showing large error over all measurement snapshots (times) it may indicate that either the logger was faulty or some errors were made during data analysis (including the logger elevation data at that particular hydrant). It has been shown that using a method described in Vassiljev *et al*. (2007) can eliminate the elevation error, but in WaterGEMS it was impossible to apply that approach.

The most questionable calibration results are in Zones 4 and 10. Both zones are pressure booster zones where

Table 4 | Main results of the calibration

Zone	Zone name	Mean square error (m)	Maximum error (m) (absolute)	Minimum error (m) (absolute)	Number of pipes in zone	Number of pressure measurements	Number of flow measurements	Pipe groups	Leak node groups
1	Merivälja	0.03	0.47	0	819	11	2	2 Plastic, 5 Metal	11
2	Pirita	0.05	0.65	0	1,431	13	3	5 Plastic, 5 Metal	12
3	Kose	0.06	0.74	0	775	11	2	3 Plastic, 3 Metal	11
4	Lasnamäe-III	1.42	4.25	0	1,332	22	2	6 Plastic, 10 Metal	22
5	Lasnamäe-II	0.02	0.53	0	1,641	14	3	6 Plastic, 13 Metal	14
6	Linna-III	0.23	1.13	0.16	338	8	1	4 Plastic, 6 Metal	8
7	Toompea	0.01	0.17	0	112	5	1	1 Plastic, 4 Metal	5
8	Linna-II	0.58	1.62	0.03	11,274	34	3	8 Plastic, 13 Metal	34
9	Taela	n/a	n/a	n/a	477	n/a	n/a	n/a	n/a
10	Mustamäe-Õismäe	1	3.19	0.01	2,436	18	3	8 Plastic, 9 Metal	18
11	Nõmme	0.13	1.39	0	8,316	28	15	5 Plastic, 7 Metal	28

various hydraulic head targets are kept throughout the day in pumping stations. As much larger errors were noticed at times when there was a change in pressure regime, it may indicate a false interpretation of pump SCADA data. Additional analyses are therefore needed in the future to better understand the problematic side of those calibration results. In general, the achieved calibration result satisfied the ASTV. Pipe roughness and emitter coefficient values were exported into the final model to create a calibrated model with proper pump components.

CALIBRATED MODEL

After model calibration, the whole network model was built. The key in this step was to analyse what model components (e.g., reservoir with a hydraulic grade pattern, pumps in parallel with on-off timing controls, pumps with variable speed patterns or variable speed pump batteries) to use to mimic a real network operational model as closely as possible. It turned out that as the complexity of the model (sub-model) increased the choice of available pumping station representations decreased.

Although the full city network has 11 major pressure zones not all are independent from each other on a daily basis. Therefore, zone flow inputs/outputs that were used at calibration stage are now exchanged for proper, real-life network components, like pumps and valves. Depending on the source of water (surface or groundwater) and zones' connectivity with each other, the whole network can be divided into four areas, as shown in Figure 9. While Area 2 and Area 3 are mostly on surface water source, Area 1 and Area 4 are on a groundwater source. All areas can be run separately through WaterGEMS Scenarios Manager.

Most of the areas can be easily represented with real network components, including variable speed pumps and variable speed pump batteries, fixed head pumps and pressure reducing valves. The most challenging is to represent a variable speed pump fixed head setting that changes throughout the day and/or weekday. As previous studies have indicated, a true model component does not exist for those situations and some alternative way should be used (Koor *et al.* 2012). For example, multiple pump elements can be used with timely controls. Although it replicates the true situation quite nicely, it affects the network hydraulic calculations (due to heavy amount of time controls for 1 week) and such representation cannot be used for pump optimization studies.

On the other hand, the available speed pattern-based pump representation is good for that time moment when the model was created but it might not be a correct solution in future studies. Demands are changing, leakages appear and are fixed again – those situations cause a change in pump working pattern. At this stage, the future pump optimization routine was more important and therefore pump speed patterns were

Figure 9 | Four main network areas that can be run separately.

Figure 10 | Pumping stations in Zone 11.

derived from the SCADA data (no direct speed coefficients were available from SCADA directly and therefore those were calculated based on pump head measurements).

As mentioned before, some of the zones are mostly pressure booster zones (Area 2 and Area 3), but Area 4 is purely based on groundwater source having 15 pump stations (Figure 10). From the SCADA system it was impossible to pick out the pump settings directly (speed coefficient). Therefore, those control rules were developed from the measured flow/head values.

Due to the number of pumps, it was the most complicated task in that zone. Although all pumps were able to pump into the same network, basically three main sub-areas were recognized that helped to stabilize the system (lower left, centre and right hand side sub-area) towards the whole working area. All four main areas were successfully modelled to the stage where real pump station elements for that moment were used, and the model is used on a daily basis for decision-making.

CONCLUSION

Methods for an updatable, large-scale, calibrated water network model that can be used for various modelling tasks were researched based on the tools that are readily available for ASTV (Bentley WaterGEMS) without spending additional resources for software development. Model creation tasks can be divided between database level and HMS package. Finding the best split between those tasks depends on available knowledge, tools and a wish to reduce routine tasks in the future model updates. It was found that Level 1 (elimination of sections that are not connected to main water source) and Level 2 (elimination of dead-end pipes with simple demand shifting) simplifications can be already made at GIS database by the company's personnel. Out-of-the-box tools were used mainly for two reasons: (1) readily available tools to connect to GIS database and therefore to be able to update the model on a regular basis; and (2) workflow repeatability by ASTV personnel in the future that is not possible with custom-built research tools due to lack of knowledge and/or easily manageable user interface/workflows. There were several project outcomes in a positive manner that were delivered as suggestions to ASTV for future workflow improvements, including the shapefile capability consideration at GIS database level. The final calibrated model was used for decision-making on more than 240 occasions in 2013, including for example, 201 cases where maximum fire-flow capacity analysis was needed (considering minimum pressure at demand nodes) and 23 times where pipe rehabilitation studies were carried out to analyse the effect of using smaller diameter pipe. Although the current study involved one specific water company, it is believed that the researched areas of how the model can be built up and what problems may occur are useful for a broader range of engineers/modellers. There are some suggestions that are more specific for Bentley WaterGEMS modelling package users (e.g., data connection schemes, pump station interpretations, user parameter definition for easier visualization possibilities, etc.), but also more general guidelines that can be applied to any software platform (e.g., simplification routines at GIS database level, measurement data splitting, etc.).

ACKNOWLEDGEMENTS

This research was supported by European Social Fund's Doctoral Studies and Internationalisation Programme DoRa, which is carried out by Foundation Archimedes. Financial support by the Institutional Research Funding IUT19-17 at Tallinn University of Technology is greatly appreciated.

REFERENCES

Ainola, L., Koppel, T., Tiiter, K. & Vassiljev, A. 2000 Water network model calibration based on grouping pipes with similar leakage and roughness estimates. In *Proceedings of the ASCE's Joint Conference on Water Resources Engineering and Water Resources Planning & Management* (R. H. Hotchkiss & M. Glade, eds). ASCE, CD-ROM.

Almeida, A. B. & Ramos, H. M. 2010 Water supply operation: diagnosis and reliability analysis in a Lisbon pumping system. *J. Water Supply Res. Technol.–AQUA* **59** (1), 66–78.

Bahadur, R., Johnson, J., Janke, R. & Samuels, W. B. 2006 Impact of model skeletonization on water distribution model parameters as related to water quality and contaminant consequences assessment. In *Proceedings of the 8th Annual Water Distribution Systems Analysis Symposium (CD-ROM)*, August 27–30, Cincinnati, OH.

Bentley Systems 2011 Bentley WaterGEMS V8i Help. Software Manual. http://www.bentley.com.

Bowden, G. J., Maier, H. R. & Dandy, G. C. 2002 Optimal division of data or neural network models in water resources applications. *Water Resour. Res.* **38** (2), 2.1–2.11.

Crozier, R., Foster, J. & Fumex, A. 2012 Re-designing a modern city's water supply system. In *Proceedings of the 14th Water Distribution Systems Analysis Conference* (A. C. T. Barton, ed.). September 24–27, Adelaide, Australia, pp. 133–147.

Duchesne, S., Chahid, N., Bouzida, N. & Toumbou, B. 2013 Probabilistic modelling of cast iron water distribution pipe corrosion. *J. Water Supply Res. Technol.–AQUA* **62** (5), 279–287.

Grayman, W. M. & Rhee, H. 2000 Assessment of skeletonization in network models. In *Joint Conference on Water Resources Engineering and Water Resources Planning & Management*, Minneapolis, MI, pp. 1–10.

Jung, B. S., Boulos, P. F. & Wood, D. J. 2007 Pitfalls of water distribution model skeletonization for surge analysis. *J. AWWA* **99**, 87–98.

Koor, M., Puust, R., Vassiljev, A. & Koppel, T. 2012 Database driven hydraulic model creation procedure. In *Proceedings of the 14th Water Distribution Systems Analysis Conference* (A. C. T.-Barton, ed.). September 24–27, Adelaide, Australia, pp. 407–414.

Koor, M., Puust, R. & Vassiljev, A. 2014 Database driven updatable hydraulic model for decision making. In *Proceedings of the 12th International Conference on Computing and Control for the Water Industry – CCWI2013*. Procedia Engineering, Elsevier, Perugia, 70, pp. 959–968.

Koppel, T. & Vassiljev, A. 2009 Calibration of a model of an operational water distribution system containing pipes of different age. *Adv. Eng. Softw.* **40** (8), 659–664.

Koppel, T. & Vassiljev, A. 2012 Use of modelling error dynamics for the calibration of water distribution systems. *Adv. Eng. Softw.* **45** (1), 188–196.

Loubser, B. F., Fair, K. A. & Sherrif, F. I. 2012 Applications and management of a large water distribution system model for the city of Tshwane, South Africa. In *Proceedings of the 14th Water Distribution Systems Analysis Conference* (A. C. T. Barton, ed.). September 24–27, Adelaide, Australia, pp. 179–191.

Martinez, F., Hernandez, V., Bartolin, H., Bou, V., Alvarruiz, F. & Alonso, J. M. 2005 CALNET project: Building and updating water distribution models from GIS + CIS+ O&M +SCADA. In *Proceedings of the 8th International Conference on Computing and Control for the Water Industry (CCWI)*, Vol. 1, Exeter, UK, pp. 209–214.

Paluszczyszyn, D., Skworcow, P. & Ulanicki, B. 2011 Online simplification of water distribution network models. In *Proceedings of the 11th International Conference on Computing and Control for the Water Industry (CCWI)*, Vol. 3, Exeter, pp. 749–754.

Perelman, L. & Ostfeld, A. 2011 Toplogical clustering for water distribution systems analysis. *Environ. Modell. Softw.* **26**, 969–972.

Puust, R., Kapelan, Z., Savic, D. A. & Koppel, T. 2010 A review of methods for leakage management in pipe networks. *Urban Water J.* **7** (1), 25–45.

Randall-Smith, M., Rogers, C., Keedwell, E., Diduch, R. & Kapelan, Z. 2006 Optimized design of the city of Ottawa water network: a genetic algorithm case study. In *Proceedings of the Eighth Water Distribution Systems Analysis Conference* (S. G. Buchberger, R. M. Clark, W. M. Grayman & J. G. Uber, eds). American Society of Civil Engineers, August 27–30, Cincinnati, OH, pp. 1–20.

Rossman, L. A. 2002 *EPANET 2 User's Manual, Water Supply and Water Resources Division*. National Risk Management Research Laboratory, Cincinnati, OH.

Savic, D. A. & Walters, G. A. 1995 *Genetic Algorithm Techniques for Calibrating Network Models*. Report No. 95/12, Center for Systems and Control Engineering, University of Exeter, UK.

Savic, D. A. & Walters, G. A. 1997 Genetic algorithms for least-cost design of water distribution networks. *J. Water Resour. Plann. Manage. ASCE* **123** (2), 67–77.

Simpson, A. R., Dandy, G. C. & Murphy, L. J. 1994 Genetic algorithms compared to other techniques for pipe optimization. *J. Water Resour. Plann. Manage. ASCE* **120** (4), 423–443.

Speight, V., Betanzo, E. & Porro, J. 2011 Data integration to support hydraulic modelling at a large water utility. In *Proceedings of the 11th International Conference on Computing and Control for the Water Industry (CCWI)*, Vol. 1, Exeter, UK, pp. 9–14.

Ulanicki, B., Zehnpfund, A. & Martinez, F. 1996 Simplification of water distribution network models. In *Proceedings of the 2nd International Conference on Hydroinformatics*, Zürich, Switzerland, pp. 493–500.

Vassiljev, A. & Koppel, T. 2012 Estimation of real-time demands on the basis of pressure measurements. In *Proceedings of the Eighth International Conference on Engineering Computational Technology* (B. H. V. Topping, ed.). Civil-Comp Press, Stirlingshire, Scotland, UK, paper 54. doi:10.4203/ccp.100.54.

Vassiljev, A. & Koppel, T. 2013 Use of the real-time demands for calibration of water distribution systems. In *Proceedings of the 14th International Conference on Civil, Structural and Environmental Engineering Computing* (B. H. V. Topping & P. Iványi, eds). Civil-Comp Press, Stirlingshire, Scotland, UK, p. 233.

Vassiljev, A., Koppel, T. & Puust, R. 2005 Calibration of the model of an operational water distribution system. In *Proceedings of the Eighth International Conference on Computing and Control for the Water Industry* (D. Savic, G. Walters, R. King & S.-T. Khu, eds). University of Exeter, Exeter, UK, pp. 155–159.

Vassiljev, A., Minguell Font, L. & Puust, R. 2007 Use of pressure differentials for calibration of the operational water distribution system. In *Proceedings of the Ninth International Conference on Computing and Control for the Water Industry* (B. Ulanicki, K. Vairavamoorthy, D. Butler, L. M. Bounds & F. Ali Memon, eds). Taylor & Francis, pp. 205–209.

First received 20 March 2014; accepted in revised form 5 June 2014. Available online 7 August 2014