

Theme: A. Technologies for water management and monitoring

Subtheme: A5. Optimization techniques and their application

Special session: N.A.

Practical application of optimization techniques to drinking water distribution problems

Peter van Thienen¹, Ina Vertommen¹, Karel van Laarhoven¹

¹ KWR Watercycle Research Institute, Groningenhaven 7, 3433 PE Nieuwegein, The Netherlands

Corresponding author: peter.van.thienen@kwrwater.nl

Abstract. Numerical tools for the optimization of several aspects of drinking water distribution networks have been around for some time now and are widely discussed in the scientific literature. However, their successful practical application remains a challenge, especially when combining multiple objectives and operational boundary conditions. In this contribution, we describe a number of optimization cases, including the optimization approaches applied. We discuss the technical and practical challenges that are faced when applying numerical optimization techniques to real world problems of water utilities.

Keywords: Drinking water distribution networks; numerical optimization; practical issues.

1 Introduction

Numerical tools for the optimization of several aspects of drinking water distribution networks have been around for some time now and are widely discussed in the scientific literature [1,2]. Some methods have found their way into commercial software, and can be applied for practical problems (e.g. pipe dimensioning). However, their successful application in practice remains a challenge, as will be described below, especially when combining multiple objectives and operational boundary conditions.

In order to make advanced optimization techniques available to the Dutch water utilities, KWR has developed a generic software platform for the optimization of Drinking Water Distribution Systems (DWDS), called Gondwana [3]. The platform is designed to support its user in 1) defining and solving optimization problems with respect to many aspects of design, operations and maintenance, and 2) answering research and design questions in the context of scenario studies. Over the past two years, this platform has grown towards maturity and has been applied in a number of research and consultancy projects. Further development towards new objectives and applications continues.

In this contribution, we describe a number of cases with the technical approaches applied. We then discuss the technical and practical challenges that we face when applying numerical optimization techniques to real world problems of water utilities.

2 Utility perception of optimization

Discussions with utility employees involved in hydraulic modelling and design of networks have shed light on a twofold perception of optimization techniques within the utilities. For a class of problems which are deemed too complex for optimal design by humans (such as designing for optimal performance in a multitude of scenarios), numerical techniques are perceived to be very promising. However, utilities do not appear to be 'ready' to test them. A second class of problems (such as the optimal dimensioning of pipe systems) can be solved by humans with the right tools and some investment of time. However, this does not necessarily (or likely) result in an optimal solution, though it does result in a feasible/workable one. For this class of problems, utility

employees wonder what the benefits of the application of numerical techniques are, and are not convinced that they will lead to significantly better designs and/or perform as required in practice, and time savings, but are willing to do case studies.

3 Problem types and optimization approaches

3.1 Optimization method

Gondwana combines evolutionary algorithms implemented in the *Inspired* library [4] with the EPANET2 hydraulic computation engine [5]. Rather than encoding decision variables in a binary string to which simple mutation and crossover variators can be applied, Gondwana works with problem specific genomes (and combinations of these) with their own specific variators (crossover and multiple mutators). Even though the former approach is simpler, the latter may greatly increase the rate of convergence in optimization problems by: 1) allowing problem specific information to be used to prevent too many ‘nonsense solutions’ to be formed, thereby significantly reducing the searchable solution space; 2) allowing for the implementation of heuristic variators which using problem specific knowledge to search the solution space ‘intelligently’; and 3) result in optimal designs that are more easily acceptable and/or interpretable for water utilities.

3.2 Problem types and specific variators

Table 1 gives an overview of different problem types to which Gondwana has been and is being applied. All problems concern optimizations (generally minimizations) of a certain parameter given a set of boundary conditions. Also included are lists of problem specific variators which are used in the optimization process.

Table 1 Problem types and problem specific variators.

| Problem type | description | genome type | variators* |
|-------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|---------------------------|
| 1 Pipe dimensioning | select the appropriate diameter for pipes | diameter per pipe | NPC, SeM, FM, CM, LPM, Co |
| 2 Network blueprint optimization | as 1, taking into account failure of system components | as 1 | NPC, SeM, FM, CM, LPM, Co |
| 3 DMA subdivision | subdivide a network into sectors for maximum sensitivity to leakage signals with the minimum amount of flow meters | per node 1 of n subsections, with the requirement that subsections are fully connected geometrically | SuGS, SuS, SuM, SuC |
| 4 Network transition/replacement strategy | generate an optimal phasing for pipe replacements from a current network to a network blueprint in a predefined number of steps | per pipe 1 of n subsections, with no connection restrictions | RM, CM, SuC |

* RM random mutation; NPC n point crossover; SeM selection mutation; FM flatiron mutation; CM copycat mutation; LPM list proximity mutation; SuGS subdivision growth/shrinkage; SuM subdivision merge; SuC subdivision crossover; Co covariator; see [6,7] for more details.

4 Challenges

4.1 Occurrence of issues in practical situations

Application of Gondwana to practical design problems for a number of Dutch utilities has resulted in valuable insights into the hurdles that need to be taken for such an application to be successful. The issues can be subdivided into four categories:

- problems in the construction, completion or validation of network models;
- difficulties in obtaining additional data required for the optimization;
- the requirement for additional functionality in the optimization software to solve a customer specific problem;
- technical issues, including performance and solution stability.

Table 2 gives an overview of the issues and their occurrence in our practical experiences in five cases with Dutch utilities from optimization projects on networks segments of 150-800 km length. The following sections provide a more detailed discussion of these concrete issues and more abstract considerations. We stress that many of the issues encountered provide valuable lessons for new applications. Some of these are easily implemented and lead to a smoother process in the next application. This is illustrated by cases D and E in Table 2, which benefited from the experiences in cases A-C. Others, discussed in more detail below, are a bigger challenge. This also seems to be related to the operational orientation of water utilities, in which research and development and the (data) requirements for this are not the first priority.

4.2 Problem definition

A numerical optimization of a network design or operation requires a quantitative definition of the objective(s) of the optimization [8]. The most obvious objective for the optimization of a network dimensioning is a minimization of costs. For new networks, this is trivial. However, when optimizing an existing network, the remaining monetary value of the pipes in the ground is also a part of the cost optimization problem, raising a number of additional questions w.r.t. time horizon, depreciation and additional (external) motivations for rehabilitation.

Objectives and boundary conditions are in general conceptually very similar. We can distinguish three categories in terms of quantifiability:

1. absolute boundary conditions for e.g. pressure (the pressure on node set X should never drop below value Y) or flow velocity;
2. "soft" pressure boundary conditions (we prefer the pressure on node set X to always be above value Y, but it should never drop below value Z), applying a penalization in the value range between Y and Z;
3. 'free' objectives, which have no a priori desired value, other than 'minimized or maximized'.

Note that the second category introduces the issue of selecting an appropriate penalty value.

Many parameters may be aggregated or abstract, such as uniformity of flow (direction), water age, etc. The difficulty with these is not in the formulation of a quantification as such, but in finding one that lines up with and does justice to the idea that the utility has on the parameter on a more abstract level. For all three types, in many cases, a large number of data is subjected to a single condition. Often, some form of aggregation of the data is desired, such as taking the mean, minimum or maximum (or a certain percentile) value of the set. Whereas the choice of aggregation function might be obvious for some parameters (e.g. pressure boundary conditions), it is less obvious for others, such as the example of the uniformity of flow.

Table 2: Overview of issues encountered in the application of optimization techniques to practical design problems.

| category | issue | case | | | | |
|----------------------|-------------------------------------------------------------------|------|---|---|---|---|
| | | A | B | C | D | E |
| problem definition | discussion on appropriate formulation of the optimization problem | × | × | | × | × |
| network model | construction of network model | × | | | | |
| | skeletonization of network model | × | | | | |
| | errors in the hydraulic model | × | × | × | | |
| | missing links in the hydraulic model | × | × | | | |
| | pressure issues in hydraulic model | × | | × | | |
| additional data | availability and quality of data | | × | × | × | × |
| | delays in communication | × | × | | | |
| | delays in supply of data | | × | × | | |
| | relevant valves for relevant sectioning | × | × | | | |
| added features | pressure control on reservoirs | × | | | | |
| | limitations on inflow/outflow of reservoirs/treatment plants | × | | | | |
| | network manipulations for section failure scenarios | × | × | | | |
| technical challenges | technical issues due to model size | | × | | × | × |
| | solution stability issues on utility platform | × | | | | |

4.3 Network model

As with all modelling exercises, garbage in is garbage out. This means that optimizations only result in meaningful and useful designs if the network model that is being optimized is sufficiently representative of the real world situation.

It starts with the network itself. This is generally (at least by the Dutch water utilities) made in either Infoworks or Synergi. In order to use this model in an optimization approach which is built around EPANET, a conversion is necessary. In some cases, the conversion is incomplete (e.g.

complex controls). And after an optimization, its result(s) need to be converted back. Both conversions may be a source of differences or errors.

In some cases, skeletonization of the network model is required. This may result in essential features of a network being lost. In one of our pilots, a crucial link lost in the skeletonization process was actually reintroduced by the optimizer. This serves to illustrate both the need for careful scrutiny of skeletonized models and the power of the optimization process.

In our experience, current network models do not always fulfill the boundary conditions that are prescribed for the optimization problem applied to this model. For example, minimum pressure conditions may not be met under exceptional conditions (section failure) or even under nominal conditions. We have identified two approaches to resolving these situations:

- first fix the current network model to meet the boundary conditions, then optimize;
 - have the optimizer fix the fulfilling of the boundary conditions as part of the optimization process.
- The former approach requires more work beforehand, but leads to faster convergence in the optimization process, because the starting point for the optimization will be closer to the optimum.

Solutions that are acceptable in theory may not always be acceptable in practice. A good example is the optimization of (water quality) sensor placement. Certain junctions in a network model may provide an excellent location in terms of hydraulics, but may not offer a suitable space for placing the necessary sensor peripherals (e.g. a busy crossroads), or may not have power available. A more elaborate description and discussion of such cases is provided in [9].

4.4 Obtaining additional data

Representativeness of the network model also includes other factors like the inclusion of leakages and supplies to connecting networks. But the issue goes beyond representative network models. For more complex optimizations, information on the operation of a network is required which is not necessarily formalized and written down, but may be present only in the heads of the most experienced operators. An example of this (from the network blueprint optimization in Table 1) is the actions that a water utility takes when a component of the production/distribution system fails, and other parts need to provide a backup. The question of which valves to open/close, which pumps to ramp up or down, etc., is often present in the head of experienced operators, but not directly available to hydraulic modellers trying to solve an optimization problem. Steps to be taken, specifically by utilities, include completion of the digitization of all assets, uniform storage accessibility of data, data and model quality assurance, and formalization of procedures during nominal and crisis situations.

4.5 Adding features

A careful definition of the optimization problem, as discussed in section 4.2, sometimes leads to the conclusion that the current version of the optimization tool, in our case Gondwana, is not fully equipped to address the problem, and some extension is required. In two of the three cases listed in Table 2, the requirement for network robustness (in terms of the network continuing to function when a single segment of the network fails) to be a part of the optimization problem was recognized at an early stage.

An approach was selected and implemented, using valve sections as a starting point. Assuming perfect valve reliability (which is a good approximation for well-maintained valves in the transport structure), each section was subsequently shut down, checking that the supply in all the remaining sections met the boundary conditions. This was done as an integral step of each design evaluation in

the optimization process. During the evaluation of each candidate network design, failure of each of these sections is considered, during which the pressure boundary condition at all nodes outside a failing section is either enforced or deviations are penalized. This results in the lowest cost robust network design which is capable of supplying this minimum pressure in all failure scenarios, with somewhat larger pipe diameters than a design optimized without the failure scenarios. However, as this approach was applied, both the researchers and the water utility experts realized that a range of possible actions taken by the utility when a section fails (such as the opening and closing of valves, the change of pump settings, etc.) also needed to be implemented. These were added and tested in an iterative approach.

In this way, the platform and its range of potential applications grow.

4.6 Technical challenges

Modern optimization codes can rapidly address optimization problems with a reasonable number of degrees of freedom (order of hundreds). Larger problems, with larger network models, put up higher demands in terms of memory and storage (because of the multiplication of storage requirements with candidate populations in evolutionary algorithms), and calculation time. Moreover, the nature of the optimization problem may influence the practical usability of the optimization platform. One example is the problem type 3 of DMA subdivision, which is essentially a version of the np-hard minimal k-cut problem [10]. While genetic algorithms are known to be suitable for this type of problem due to their ability to locate solutions in the vicinity of global optima, one of their limitations is the unreasonable amount of function evaluations required to converge completely [10].

Currently, problems with around 2500 degrees of freedom in network models of the order of 10,000 nodes and pipes show good convergence in acceptable calculation times on recent (but not state of the art) computers using our optimization platform Gondwana. It is our inference that the use of specific variators such as those listed in Table 1 help to stretch the boundary of applicability towards larger or more complex problems. We do not know yet how far we can go. Investing in powerful computing facilities (hardware or cloud) can further stretch the limits. But even if computational requirements prevent full convergence of an optimization problem, results which are not optimal but better than current designs provide a step forwards for water utilities.

5 Conclusions and outlook

Numerical optimization of drinking water networks and their operation remains a promising technique. Overcoming a number of issues in its practical application remains a challenge, however, both on the technical side and on the side of the utilities (perception, problem formulation, availability of information and formalization of procedures). A number of concrete steps regarding digitization of assets, data storage and availability, etc. are being taken by several utilities. This, together with advancing technological developments, will bring a wide application of optimization techniques in the drinking water sector within reach.

Our current efforts focus on robust optimization of networks, initially with respect to uncertainty in demand. Important questions, addressed before by others but without a final solution, include how to define robustness and a suitably broad range of scenarios, as well as the computational challenges involved.

An important consideration for water utilities to apply optimization techniques will be the balance between the effort to provide all necessary information for an optimization on the one hand and the

effort to generate a solution 'by hand' on the other. For now, the perception of some Dutch water utilities that have participated in pilot projects appears to be that the significant effort in gathering all necessary data may not be worth the while. Nevertheless, a strong interest in these techniques remains in the Dutch drinking water sector. Moreover, the workforce of the utilities ages and knowledge that is not formalized disappears with retiring employees. Therefore, the formalization and digitization of all knowledge and assets becomes an urgent issue that serves more purposes than just feeding optimization models.

References

- [1] P. Bieupoude, Y. Azoumah, P. Neveu, Optimization of drinking water distribution networks: Computer-based methods and constructal design, *Computers, Environment and Urban Systems* 36 (2012), 434–444.
- [2] A. De Corte, K. Sørensen, Optimisation of gravity-fed water distribution network design: a critical review, *European Journal of Operational Research* 228 (2013), 1–10.
- [3] P. van Thienen, I. Vertommen, Gondwana: a generic optimization tool for drinking water distribution systems design and operation, *Procedia engineering* 119 (2015), 1212-1220.
- [4] A.L. Garrett, Inspyred 1.0 documentation, <https://pythonhosted.org/inspyred/overview.html>, Jan. 2015.
- [5] L.A. Rossman, EPANET 2 Users Manual, U.S. Environmental Protection Agency report EPA/600/R-00/057 (2000).
- [6] K. van Laarhoven, P. van Thienen, I. Vertommen, Problem specific variators in numerical optimization of drinking water networks using genetic algorithms. in preparation (2018).
- [7] I. Vertommen, K. van Laarhoven, P. van Thienen, C. Agudelo-Vera, T. Haaijer, R. Diemel, Optimal design of and transition towards water distribution network blueprints, submitted to EWAS International
- [8] P. van Thienen, I. Vertommen, G. Mesman, Advanced modelling and optimization in drinking water distribution systems - technical requirements and steps for water utilities, *Water Solutions* (2017).
- [9] P. van Thienen, B. de Graaf, J. Hoogterp, J. van Summeren and A. Vogelaar (2018). Bounds on water quality sensor network performance from design choices and practical considerations. *Water Practice and Technology* (submitted).
- [10] J. Kim, I. Hwang, Y. Kim, B. Moon. Genetic approaches for graph partitioning: A survey. *Proceedings of GECCO'11*, Dublin, Ireland, 2011