



Water Distribution System Optimization using Metamodels

Darren Ross Broad

B.E. Civil/Env (Hons), B.Sc. (Ma. & Comp. Sc.)

Thesis submitted for the degree of

Doctor of Philosophy

The University of Adelaide

School of Civil, Environmental and Mining Engineering

September 2014

Contents

<i>Abstract</i>	<i>vii</i>
<i>Declaration</i>	<i>x</i>
<i>List of Figures</i>	<i>xi</i>
<i>List of Tables</i>	<i>xiv</i>
<i>List of Publications</i>	<i>xviii</i>
<i>Acknowledgments</i>	<i>xxi</i>
Chapter 1 <i>Background and Publications Overview</i>	1
1.1 Introduction	2
1.2 Publications	4
1.2.1 Contributions to WDS Optimization	4
1.2.2 Contributions to the Development of a Metamodelling Framework for WDS Optimization.....	5
1.2.3 Summary.....	10

Chapter 2	<i>Publication 1: Water Quality</i>	11
2.1	Introduction	14
2.2	Literature Review	14
2.3	Objectives	16
2.4	Optimization Approach	17
2.5	Metamodel Development	18
2.6	Optimization with Metamodels	21
2.6.1	Checking Solutions with the Simulation Model.....	22
2.6.2	Constraint Adjustment.....	26
2.7	Case Study	29
2.7.1	Analysis Conducted	29
2.7.2	Metamodel Performance.....	32
2.7.3	Optimization Results – EPANET.....	34
2.7.4	Optimization Results – Metamodels.....	34
2.7.5	Optimization Results – Adjusted Constraints.....	36
2.7.6	Comparison of Computational Time.....	38
2.7.7	Discussion	39
2.8	Other Applications for WDS	40
2.9	Conclusions	41
2.10	Acknowledgments	42
Chapter 3	<i>Publication 2: Local Search</i>	43
3.1	Introduction	46
3.2	Metamodelling Procedure	50
3.3	Local Searches	52
3.3.1	Sequential Downward Mutation	53
3.3.2	Random Downward Mutation.....	54
3.3.3	Maximum Savings Downward Mutation.....	54
3.3.4	Triangular Mutation.....	55
3.3.5	Probabilistic Allele Swapping.....	57
3.3.6	Simulated Annealing.....	59

3.4	Case Study	61
3.4.1	Analysis Conducted	62
3.4.2	Results	65
3.5	Conclusion.....	68
Chapter 4	<i>Publication 3: Complex Hydraulic Systems</i>	70
4.1	Introduction	74
4.2	Proposed Methodology.....	76
4.2.1	Introduction	76
4.2.2	Complexity of Hydraulic Simulation Model.....	77
4.2.3	Complexity of Decision Space.....	77
4.2.4	Locations at which Simulation Model Outputs are Required.....	78
4.2.5	Summary of Proposed Methodology	81
4.3	Case Study: Wallan	83
4.3.1	Introduction	83
4.3.2	Problem Formulation	85
4.3.3	Development of ANN Metamodels	89
4.3.4	Results and Discussion.....	98
4.4	Conclusions.....	106
4.5	Acknowledgments.....	107
Chapter 5	<i>Publication 4: Data Uncertainty</i>	108
5.1	Background.....	111
5.1.1	Uncertainty	111
5.1.2	Metamodelling.....	112
5.2	Proposed Metamodelling Approach	116
5.2.1	Metamodel Scope Definition (Step 2)	118
5.2.2	Post-Optimization Solution Checking (Step 8).....	125
5.3	Application to Risk-Based Optimal Design of WDSs	127
5.3.1	Background.....	127
5.3.2	Metamodel Scope Definition.....	133
5.3.3	Case Studies.....	139

5.4	Summary and Conclusions	154
5.5	Acknowledgments.....	156
<i>Chapter 6</i>	<i>Conclusions and Recommendations</i>	<i>157</i>
6.1	Research Contributions.....	158
6.2	Recommendations for Future Work.....	161
	<i>Bibliography.....</i>	<i>165</i>
	<i>Appendix A: Systematic Approach applied to Simple Mathematical Functions ..</i>	<i>181</i>
	<i>Appendix B: Hammersley Sampling for Stochastic Variables.....</i>	<i>186</i>
	<i>Appendix C: Case Study Details</i>	<i>190</i>

Abstract

Evolutionary Algorithms (EAs) have been shown to apply well to optimizing the design and operations of water distribution systems (WDS). Recent research in the field has focussed on improving existing EAs and developing new ones so as to obtain better solutions (closer to the global optimum) and/or find solutions more efficiently.

The primary aim of this research, however, has been to broaden the scope of optimization to include a number of the many factors that planning engineers need to consider when designing or planning the operations of WDS. Those factors considered here are (1) water quality criteria, (2) real-world, complex systems, and (3) the incorporation of data uncertainty.

Incorporating each of these factors independently increases computational run-time of EA-based optimization of an algorithm that is already computationally intensive compared to other (inferior) algorithms that have been used in WDS optimization. Water quality models tend to run slower than hydraulic models due to the shorter timestep that is required to ensure sufficient accuracy, and the need for extended period simulations thereby increasing the simulation duration. Real-world models run slower due to their size. Data uncertainty is typically accounted for through the use of Monte Carlo simulations, that add several orders of magnitude to the computational requirements of optimization.

Considering each of these factors together compounds the computational requirements to a point where it is impossible to optimize WDS using EAs in a reasonable amount of time. In this research metamodels have been used in place of simulation models within an EA to reduce this computational burden. A metamodel is a model of a model that runs much faster than the said model, but is still a high-fidelity approximation of it. The particular type of metamodel used in this research is an Artificial Neural Network (ANN) due to its theoretical capabilities and demonstrated effectiveness in water resources applications.

The use of metamodels to act as surrogates for complex simulation models is not a trivial task. Therefore, guidelines have been developed on how best to incorporate them into the WDS optimization process.

The overall metamodel-empowered, EA-based optimization algorithm developed in this research was applied to several case studies. Two small case studies, both variations of the New York Tunnels problem were studied for proof-of-concept purposes. They demonstrated that near globally-optimal solutions could still

be found using the metamodel-based approach, i.e. there was minimal compromise in the effectiveness of the EA-based approach. Two larger, real-world problems were also studied: Wallan (operations planning) and Pacific City (system augmentation). These last two case studies were key to demonstrating the power of using metamodels in that they enabled a computational speed-up of up to 1375 times (137,500%) compared to a non-metamodel approach. This speed-up includes factoring in the computational overheads of using metamodels, i.e. time to generate calibration data and calibrate the metamodels.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

SIGNED:.....DATE:.....

List of Figures

Figure 1-1 Traditional, or non-metamodel-based optimization procedure	6
Figure 1-2 Metamodel-based optimization procedure with references to relevant publications for further detail.....	7
Figure 2-1. Steps in optimizing a water distribution system, (a) ANN metamodel approach, (b) simulation model approach.....	19
Figure 2-2. A simple WDS and its ANN metamodel.....	20
Figure 2-3. Strategies of obtaining the optimal solution through the evaluation of select solutions with EPANET, (a) top few solutions, (b) local search, (c) new optimal solutions.	25
Figure 2-4. ANN approximations to an EPANET model.....	27
Figure 2-5. RMSE using two different training sets for the critical nodes of the NYT-WQ problem.	33

Figure 2-6. Constraint adjustment for the NYT-WQ problem: ANN #3 of scenario A.....	37
Figure 3-1. Procedure for developing and using an ANN metamodel for the purpose of optimization.....	51
Figure 3-2. Pseudo-code for sequential downward mutation.....	53
Figure 3-3. Pseudo-code for random downward mutation.....	54
Figure 3-4. Pseudo-code for maximum-savings downward mutation	55
Figure 3-5. Comparison of different types of mutation.....	56
Figure 3-6. Pseudo-code for triangular mutation.....	57
Figure 3-7. Pseudo-code for probabilistic allele swapping	58
Figure 3-8. Pseudo-code for simulated annealing.....	60
Figure 3-9. The New York Tunnels Water Distribution System, with critical nodes.....	62
Figure 3-10. Run-times for each local search, with comparison to EPANET-GA and ANN-GA.....	68
Figure 4-1. Layout of Wallan case study area.	84
Figure 4-2. Metamodel structure for Stage 1 of HGC Case Study.....	90
Figure 5-1. Basic Sequential Framework steps to develop and use a metamodel for EA-based optimization, with new/modified steps from this paper in bold.....	117
Figure 5-2. Generic fitness evaluation and two possible metamodel scopes.....	119
Figure 5-3. Proposed process for determining best metamodel scope.....	122
Figure 5-4. Fitness evaluation for risk-based optimization of WDS with metamodel scope options.	134
Figure 5-5. Comparison between simulation model and metamodel for critical hydraulic node, 16, for the New York Tunnels case study.	145

Figure 5-6. Comparison between simulation model and metamodel for critical chlorine node, C4, for the Pacific City case study. 151

Figure 6-1 Methodology for the optimization of water distribution systems using metamodels..... 159

Figure B.1. Example of data sampling methods in 2 dimensions (96 points in $U[0,1]$). (a) RS, (b) LHS with 4 stratifications per dimension, (c) HS (dimension 1 and 2), (d) HS (dimension 5 and 10)..... 188

Figure B.2. Example of data sampling methods in 2 dimensions (96 points in $N[0,1]$). (a) RS, (b) LHS with 4 stratifications per dimension, (c) HS (dimension 1 and 2), (d) HS (dimension 5 and 10)..... 189

Figure C.1. Schematic of the Pacific City case study..... 194

List of Tables

Table 1.1 Contributions to WDS optimization by publication.....	4
Table 1.2. Publications, chapters and their main focus.....	10
Table 2.1. ANN Parameters	31
Table 2.2. GA Parameters	32
Table 2.3. Solutions found [\$million] for the NYT-WQ problem with different methods of checking solutions with EPANET, with the current best-known solution in italics and the number of times it was found out of 10 runs in brackets.....	35
Table 2.4. Optimal solutions found [\$million] by constraint adjustment for the NYT-WQ problem.	38
Table 2.5. Computational times for metamodeling and optimization of the NYT-WQ problem.	39

Table 3.1. ANN Parameters	63
Table 3.2. GA Parameters	64
Table 3.3. Calibrated local search parameters.....	65
Table 3.4. RMS error in the validation set and training data range from 30 ANN Metamodels	65
Table 3.5. Local Search results from 100 random initialisations in the GA population.....	67
Table 3.6. P values from a t-test, illustrating the significance of the local search results.....	67
Table 4.1. Different ANN metamodel scenarios considered.	85
Table 4.2. Allowable range for tank trigger levels.	86
Table 4.3. Objective function parameter values.....	89
Table 4.4. Comparative statistics between the original model and the skeletonized model.....	92
Table 4.5. Critical nodes for different ANN metamodeling scenarios.	95
Table 4.6. Impact of different stages of proposed critical node determination procedure for scenario 4.....	97
Table 4.7. GA Parameters used for the Wallan Case Study.....	99
Table 4.8. Average RMS errors of the validation set for the various ANN model development scenarios considered.....	100
Table 4.9. Average R2 values of the validation set for the various ANN model development scenarios considered.....	100
Table 4.10. Summary of optimal solution obtained.	102
Table 4.11. Single day energy costs (peak and off-peak tariffs) for the optimal solution with a comparison to current operations.	103
Table 4.12. Chlorine dose rates [mg/L] for optimal solution with a comparison to current operations.....	104
Table 4.13. Computational requirements optimization (hours).....	105

Table 5.1. Assessment of fitness calculation steps for single-objective risk-based optimal design of water distribution systems.....	135
Table 5.2. Scenarios used in the two case studies.....	142
Table 5.3. Metamodelling parameters used for NYT case study.	143
Table 5.4. Input and output variables for the various MLPs within the metamodel for New York Tunnels.....	144
Table 5.5. Metamodel development results for NYT, showing the RMSE and R ² for the validation set.....	144
Table 5.6. Genetic Algorithm parameters used for NYT.	146
Table 5.7. Number of fitness evaluations of each post-EA solution checking type for different scenarios for Pacific City.	146
Table 5.8. Optimization results for New York Tunnels. Statistics of NPV shown, as well as frequency that the best solution was found for 30 runs per scenario.	147
Table 5.9. CPU times (hours) of each metamodelling step and comparison to non-metamodelling approach for NYT.	148
Table 5.10. Input and output variables for the various MLPs within the metamodel for New York Tunnels.....	150
Table 5.11. Metamodel calibration results for Pacific City.	150
Table 5.12. Number of fitness evaluations of each post-EA solution checking type for different scenarios for Pacific City.	152
Table 5.13. Optimization results for Pacific City. Statistics of NPV shown, as well as frequency with which the best solution was found for 30 runs per scenario.	152
Table 5.14. CPU times (hours) of each metamodelling step and comparison to non-metamodelling approach (estimated) for Pacific City.....	154

Table A.1. Assessment of calculation steps of Bukin’s function N6.....	182
Table A.2. Assessment of calculation steps of Rastrigin’s function.....	184
Table C.1. Assumed data for calculating disinfection costs for NYT.....	191
Table C.2. Simulation model summary for the New York Tunnels case study.....	192
Table C.3. Summary of optimization decisions and search space for the New York Tunnels case study.....	192
Table C.4. Risk metric constraints used for NYT case study.....	193
Table C.5. Simulation model summary of the Pacific City case study.....	195
Table C.6. Pipe decision options for the Pacific City case study.....	196
Table C.7. Summary of optimization decisions and search space for the Pacific City case study.....	196

List of Publications

The following is a list of the publications related to the research presented in this thesis:

Journal Papers:

Broad, D. R., Dandy, G. C., and Maier, H. R. (2005). "Water Distribution System Optimization Using Metamodels." *Journal of Water Resources Planning and Management - ASCE*, 131(3), 172-180.

Broad, D. R., Maier, H. R., and Dandy, G. C. (2010). "Optimal Operation of Complex Water Distribution Systems Using Metamodels." *Journal of Water Resources Planning and Management - ASCE*, 136(4), 433-443.

Broad, D. R., Dandy, G. C., and Maier, H. R., (2014). "Systematic approach to determining metamodel scope for risk-based optimization and its application to water distribution system design", *Environmental Modelling & Software*, Submitted.

Refereed Conference Paper:

Broad, D. R., Dandy, G. C., Maier, H. R., and Nixon, J. B. (2006). "Improving Metamodel-based Optimization of Water Distribution Systems with Local Search." *IEEE World Congress on Computational Intelligence, 16-21 July 2006*, Vancouver, BC, Canada, on CD-ROM.

Unrefereed Conference Papers:

Broad, D. R. (2004). "Incorporating Water Quality and Reliability into the Optimisation of Water Distribution Systems." *Fourth Postgraduate Student Conference of the CRC for Water Quality and Treatment, 14-16 April 2004*, Noosa, Queensland, 181-187.

Broad, D. R., Dandy, G. C., and Maier, H. R. (2004). "A Metamodeling Approach to Water Distribution System Optimization." *EWRI World Water and Environmental Resources Congress, 27 June - 1 July 2004*, Salt Lake City, Utah, USA, on CD-ROM.

Broad, D. R., Maier, H. R., Dandy, G. C., and Nixon, J. B. (2005). "Estimating Risk Measures for Water Distribution Systems using Metamodels." *EWRI World Water and Environmental Resources Congress, 15 - 19 May 2005*, Anchorage, Alaska, USA.

Broad, D. R. (2006). "Optimising Water Distribution Systems using Metamodels." *Fifth Postgraduate Student Conference of the CRC for Water Quality and Treatment, 10-13 July 2006*, Melbourne, Victoria.

Broad, D. R., Maier, H. R., Dandy, G. C., and Nixon, J. B. (2006). "Optimal Design of Water Distribution Systems including Water Quality and System Uncertainty." *8th Annual International Symposium on Water Distribution Systems Analysis, 27-30 August 2006*, Cincinnati, Ohio, USA, on CD-ROM.

Report:

Gibbs, M., Broad, D. R., Dandy, G. C., and Maier, H. R. (2010). "Decision Support Systems for Water Quality Optimisation." Water Quality Research Australia.

Presentation:

Broad, D. R. (2005). "Improving Water Distribution System Optimisation through the use of Metamodels", AWA Computer Modelling Special Interest Group.

Acknowledgments

Thanks firstly to my excellent, very patient, supervisors Prof. Graeme Dandy and Prof. Holger Maier. You have given me fantastic guidance, wisdom and encouragement over the years that will serve me well for the rest of my career. Thanks also to my industry supervisor, Dr. John Nixon for his guidance.

Thanks to the former Co-operative Research Centre for Water Quality and Treatment for providing financial support for my PhD. Thanks to Greg Ryan (formerly with South East Water) and Asoka Jayaratne (Yarra Valley Water) for their assistance in developing a real-world case study. Thanks to Chris Saliba for collating the necessary data for the Wallan case study and the guided tour of the area.

Thanks also to the federal government for providing an Australian Postgraduate Award.

Thanks to fellow postgrads Matt Gibbs and Rob May for all their technical help along the way. Thanks to my fellow postgraduate students (especially Nicole Arbon, Joe Davis, Matt Haskett, Greer Humphrey, Kylie Hyde, Pedro Lee, Michael Leonard, Dalius Misiunas, Steve Need, Jakin Ravalico, Mark Rebentrost, Tim Rowan, Mark Stephens, Jerry Vaculik, John Vitkovsky, Julian Whiting, Craig Willis, Aaron Zecchin, Matt and Rob) with whom I shared many interesting discussions, Friday night drinks, whinge sessions, and if memory serves correctly 10 ACPGNAPCs.

Thanks to my colleagues at Optimatics for their encouragement and helping me significantly in my coding skills.

Thanks to my friends who frequently asked the dreaded question for any decade-long student: “How’s the PhD going?” Thanks especially to David McIver and Gerhard Bartodziej; you made me think about this PhD even when I didn’t want to, and that probably helped me get over the line.

Thanks to my family for their love and support, especially my Mum for financial support and letting me move back home... twice.

Thanks to my good mate Ryan Ogilvy who was a great supporter and listener.

Finally, thanks to God who made me, forgave me, sustains me and gives me purpose.

Soli Deo gloria