Real-Coded Genetic Algorithm Parameter Setting for Water Distribution System Optimisation

by Matthew S. Gibbs

Submitted in fulfilment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

MARCH 2008

FACULTY OF ENGINEERING, COMPUTER AND MATHEMATICAL SCIENCES
School of Civil, Environmental and Mining Engineering

THE UNIVERSITY OF ADELAIDE
AUSTRALIA

Table of Contents

Table of Contents **iii**

Abstract **ix**

Statement of Originality **xi**

Acknowledgements **xiii**

List of Figures **xv**

List of Tables **xix**

List of Symbols **xxv**

List of Acronyms **xxvii**

1 Introduction **1**

1.1 Goals of this research 3

1.2 Proposed Methodology 4

1.3 Layout and Contents of Thesis 7

2 Background **11**

2.1 WDS Optimisation 11

2.1.1 Constraint Handling 12

2.1.2 Previous WDS Optimisation Methods 14

2.1.3 Evolutionary Optimisation of WDS 15

2.1.4 Applications of EAs to WDS Optimisation 16

2.2 Genetic Algorithm Overview 17

2.2.1 Encoding Scheme 18

2.2.1.1 Binary Representation 18

2.2.1.2 Integer Representation 19

2.2.1.3 Real Value Representation 20
2.2.2 Selection 21
2.2.3 Crossover 22
2.2.4 Mutation 23
2.2.5 Advanced Operators 24
2.2.5.1 Elitism 24
2.2.5.2 Local Search Operators 24
2.2.5.3 Niching Operators 25
2.2.5.4 Linkage Learning 25
2.3 GA Theory 26
2.3.1 Schema Theory 26
2.3.2 Markov Chain Theory 27
2.3.3 Quantitative Genetics 28
2.3.4 Dimensional Analysis 29
2.4 GA Calibration Methods 30
2.4.1 Measuring Optimisation Problem Difficulty 31
2.4.1.1 Fitness Function Structure 33
2.4.1.2 Epistatic Interactions 35
2.4.2 Empirical Studies of GA Parameters 37
2.4.3 Dimensional Analysis 38
2.4.4 Parameter Control 40
2.4.4.1 Deterministic Parameter Control 41
2.4.4.2 Adaptive Parameter Control 42
2.4.4.3 Self-adaptive Parameter Control 42
2.4.5 Supervisory Algorithms 43
2.4.6 GA Calibration Methodologies 44
2.5 Summary and Proposed Methodology 45
2.5.1 GA Adopted for This Research 46
2.5.2 Relevance of the Literature 48

3 The Number of Generations Until Convergence 51
3.1 Observing an Optimal Number of Generations 52
3.1.1 Methodology 53
3.1.1.1 Test Functions 53
3.1.1.2 Parametric Study 55
3.1.2 Parametric Study Results 56

3.1.2.1 Maximal Generation Functions 59
3.1.2.2 Optimal Generation Functions 64
3.1.2.3 Problem Size Effects 69
3.1.2.4 Epistasis Effects 70
3.1.3 Discussion of Parametric Study Results 71
3.2 The Effect of Function Characteristics on the Number of Generations . . 73
3.2.1 Methodology 73
3.2.1.1 Test Functions 74
3.2.2 Function Characteristics Results 80
3.2.2.1 F3 80
3.2.2.2 F4 82
3.2.2.3 F6 83
3.2.3 Observed Effect of Characteristics on the Number of Generations 84
3.2.3.1 Roughness and Multimodality 84
3.2.3.2 Salience of Variables 85
3.2.3.3 Epistatic Interactions 85
3.2.3.4 Problem Size 87
3.2.4 Discussion of Characteristics Results 87
3.3 Summary 88

4 Development of Fitness Function Statistics 91
4.1 Spatial Correlation 92
4.1.1 Methodology 94
4.1.2 Results 98
4.1.3 Discussion 99
4.2 Epistatic Interactions 100
4.2.1 Testing the Gene Epistasis Measure 101
4.2.1.1 Method 1: Gene Epistasis by Joint Mutual Information 103
4.2.1.2 Method 2: Gene Epistasis by Fitness Function Residuals 103
4.2.2 Separability Measure 105
4.3 Decision Variable Salience 109
4.4 Summary 111

5 Predicting the Number of Generations Before Convergence 113
5.1 Background 113

5.2 The Relationship Between Fitness Function Characteristics and Population Variance . . . 115
5.2.1 Methodology . . . 116
5.2.1.1 Test Functions . . . 116
5.2.1.2 Determining the Decay in Population Variance . . . 118
5.2.1.3 Statistics Based on the Fitness Function Measures . . . 119
5.2.2 Relationship Results . . . 121
5.2.2.1 Input Determination . . . 121
5.2.2.2 Functional Form of the Relationship . . . 122
5.3 Validation of the Relationship . . . 125
5.3.1 Predicting the Decay in Population Variance . . . 125
5.3.1.1 F3 . . . 125
5.3.1.2 F4 . . . 126
5.3.1.3 F6 . . . 126
5.3.2 Predicting the Number of Generations Before Convergence . . . 128
5.3.2.1 Determining the Initial and Final Population Variance . 128
5.3.2.2 F3 . . . 129
5.3.2.3 F4 . . . 129
5.3.2.4 F6 . . . 130
5.4 Discussion . . . 130
5.5 Summary . . . 133

6 GA Calibration Methodology **135**
6.1 The Relationship Between GA Parameters . . . 136
6.1.1 Population Size . . . 138
6.1.2 Probability of Mutation . . . 138
6.1.3 Elitism . . . 141
6.1.4 Standard Deviation of Crossover . . . 142
6.1.5 Probability of Crossover . . . 144
6.2 GA Calibration Methodology . . . 147
6.3 Convergence Due to Genetic Drift . . . 150
6.4 Comparison of GA Calibration Methods . . . 151
6.4.1 Outline of the Methodologies . . . 151
6.4.2 Test Functions . . . 153
6.4.3 Function Characterisation . . . 156
6.4.4 Overall Solution Quality Comparison . . . 162

6.4.5 Function by Function Performance Comparison 167
6.5 Discussion 171
6.6 Summary 174

7 Application to WDS Optimisation 177
7.1 Cherry Hill-Brushy Plains Network 178
7.1.1 System Description 178
7.1.2 Cherry Hill-Brushy Plains Fitness Function 180
7.1.3 Cherry Hill-Brushy Plains Results 180
7.1.3.1 Fitness Function Characterisation 181
7.1.3.2 Comparison of GA Calibration Methods 184
7.1.3.3 Comparison of Best Solution Found 187
7.2 Woronora WDS 189
7.2.1 System Description 189
7.2.2 Model Calibration 191
7.2.3 Model Validation 200
7.2.4 Woronora WDS Fitness Function 206
7.2.5 Woronora Results 210
7.2.5.1 Fitness Function Characterisation 211
7.2.5.2 Comparison of GA Calibration Methods 214
7.2.5.3 Comparison of Best Solutions 216
7.3 Discussion 226
7.4 Summary 231

8 Conclusions and Further Work 233
8.1 Contributions of this Work 233
8.2 Conclusions 238
8.3 Recommended Future Work 241

References 245

Appendices 259

Appendix A Published Works 261

Appendix B Test Function Results 263

Appendix C Ranking of Median Solutions 359

Appendix D Woronora Results 367

Abstract

The management of Water Distribution Systems (WDSs) involves making decisions about various operations in the network, including the scheduling of pump operations and setting of disinfectant dosing rates. There are often conflicting objectives in making these operational decisions, such as minimising costs while maximising the quality of the water supplied. Hence, the operation of WDSs can be very difficult, and there is generally considerable scope to improve the operational efficiency of these systems by improving the associated decision making process. In order to achieve this goal, optimisation methods known as Genetic Algorithms (GAs) have been successfully adopted to assist in determining the best possible solutions to WDS optimisation problems for a number of years.

Even though there has been extensive research demonstrating the potential of GAs for improving the design and operation of WDSs, the method has not been widely adopted in practice. There are a number of reasons that may contribute to this lack of uptake, including the following difficulties: (a) developing an appropriate fitness function that is a suitable description of the objective of the optimisation including all constraints, (b) making decisions that are required to select the most appropriate variant of the algorithm, (c) determining the most appropriate parameter settings for the algorithm, and (d) a reluctance of WDS operators to accept new methods and approaches.

While these are all important considerations, the correct selection of GA parameter values is addressed in this thesis. Common parameters include population size, probability of crossover, and probability of mutation. Generally, the most suitable GA parameters must be found for each individual optimisation problem, and therefore it might be expected that the best parameter values would be related to the characteristics of the associated fitness function.

The result from the work undertaken in this thesis is a complete GA calibration methodology, based on the characteristics of the optimisation problem. The only input required by the user is the time available before a solution is required, which is beneficial in the

WDS operation optimisation application considered, as well as many others where computationally demanding model simulations are required. Two methodologies are proposed and evaluated in this thesis, one that considers the selection pressure based on the characteristics of the fitness function, and another that is derived from the time to convergence based on genetic drift, and therefore does not require any information about the fitness function characteristics.

The proposed methodologies have been compared against other GA calibration methodologies that have been proposed, as well as typical parameter values to determine the most suitable method to determine the GA parameter values. A suite of test functions has been used for the comparison, including 20 complex mathematical optimisation problems with different characteristics, as well as realistic WDS applications.

Two WDS applications have been considered: one that has previously been optimised in the literature, the Cherry Hills-Brushy Plains network; and a real case study located in Sydney, Australia. The optimisation problem for the latter case study is to minimise the pumping costs involved in operating the WDS, subject to constraints on the system, including minimum disinfectant concentrations. Of the GA calibration methods compared, the proposed calibration methodology that considered selection pressure determined the best solution to the problem, producing a 30% reduction in the electricity costs for the water utility operating the WDS.

The comparison of the different calibration approaches demonstrates three main results:

1. *that the proposed methodology produced the best results out of the different GA calibration methods compared;*
2. *that the proposed methodology can be applied in practice; and*
3. *that a correctly calibrated GA is very beneficial when solutions are required in a limited timeframe.*

Statement of Originality

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.

I give consent to this copy of my thesis being made available in the University Library.

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

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

Acknowledgements

I would like to acknowledge the support and guidance of my academic supervisors Prof. Holger R. Maier and Prof. Graeme C. Dandy. I appreciate the research freedom I have enjoyed under their supervision, I doubt any of us knew what what we were getting into when I first started this work. I would also like to acknowledge the support of my industry supervisor, Dr. John B. Nixon, especially for the typesetting suggestions that he provided.

The author was financially supported by the Cooperative Research Centre for Water Quality and Treatment (CRCWQT) and an Australian Postgraduate Award. As part of the CRCWQT, the author would like to thank Corinna Doolan, Robert Green and Dusko Mirkovic from the Sydney Water Corporation for their assistance in answering persistent questions regarding the Woronora Water Distribution System, and for supplying the data required to make the case study a success.

The author is indebted to the South Australian Partnership for Advanced Computing for the use of their facilities to generate the results presented in this thesis. I would like to thank the staff that maintain and support the nodes for making this work possible.

The author would like to thank his fellow postgrads for their support and company over the past four years. I would especially like to thank Robert May for the numerous conversations regarding the Mutual Information components of this work, and Darren Broad for organising Friday Drinks for the majority of my time as a postgraduate student at the University of Adelaide.

I would like to thank my parents, brothers and sister for their support. I must especially thank Mum for everything she has done for me over the past two years, it has made undertaking this work much easier.

Last, but definitely not least, I would like to thank Verity. The support, consideration, patience, and encouragement she has given me has been invaluable.

List of Figures

1.1 Flowchart depicting the relationship between the work presented in this thesis. 8

2.1 Representation the interaction between decision variables in terms of Mutual Information. 36
2.2 Probability distribution used for crossover 48

3.1 The different results obtained for the Optimal Generation Functions and the Maximal Generation Functions. 57
3.2 Functional form of F1 and F2. 59
3.3 Optimal populations sizes for F1 for different function evaluations. 60
3.4 Optimal populations sizes for F2 for different function evaluations. 61
3.5 Optimal populations sizes for F5 for different function evaluations. 62
3.6 Optimal populations sizes for F7 for different function evaluations. 63
3.7 Functional form of F5 and F7. 63
3.8 Functional form of F6 and F3. 64
3.9 Optimal populations sizes for F6 for different function evaluations. 65
3.10 Optimal populations sizes for F3 for different function evaluations. 66
3.11 Optimal populations sizes for F4 for different function evaluations. 67
3.12 Plot of F4 in $l = 2, 5, 10$, and 20 dimensions. The higher dimensional plots are 1–dimensional cross-sections of the function taken along the diagonal of the hypercube. 68
3.13 Flow chart depicting the fitness function calibration classes. 72
3.14 Plots of the effect of parameters A and f on the characteristics of F3B . . 76
3.15 The effect of increasing the salience of x_2 relative to x_1 for F3C 77
3.16 Plots of the effect of parameters A and p on the characteristics of F4A . . 79
3.17 Plots of the effect of parameter p on the characteristics of F6. 80
3.18 The effect of interactions on the solution found for F3A with $l = 30$. . . . 86

4.1 Extension of the Wiener–Khinchin Theorem into two dimensions. 95
4.2 The Rastrigin function and Fourier Series approximation to the Rastrigin Function used to compute the correlation statistics. 97
4.3 Temporal (R_t) and spatial (R_s) correlations compared to the true correlation (R). 98
4.4 Comparison of the spatial correlation function computed for variations of the approximated Rastrigin Function 100
4.5 Representation the interaction between decision variables in terms of Mutual Information. 101
4.6 Sampling method used for the separability measure. 106

5.1 The effect of applying the rotation matrix. 117
5.2 Example of the Fourier Series test functions used. 118
5.3 The computation of the correlation statistics from the spatial correlation measure. 120
5.4 Predicted and experimental values of k for the Fourier Series test functions. 124
5.5 Predicted and experimental values of k for F3. 126
5.6 Predicted and experimental values of k for F4. 127
5.7 Predicted and experimental values of k for F6. 127
5.8 Predicted and experimental values of g_{conv} for F3. 129
5.9 Predicted and experimental values of g_{conv} for F4. 130
5.10 Predicted and experimental values of g_{conv} for F6. 131

6.1 Best performing population sizes for Maximal Generation Functions . . . 139
6.2 Population size against probability of mutation for Optimal Generation Functions 140
6.3 Best performing (a) mutation rates and (b) number of elite solutions for the Maximal Generation Functions. 140
6.4 Number of elite solutions for each population size for Optimal Generation Functions 141
6.5 Number of elite solutions for each probability of mutation for Optimal Generation Functions 142
6.6 Standard deviation of crossover for each probability of mutation for Optimal Generation Functions 143
6.7 Best performing (a) fraction for the standard deviation of crossover and (b) probabilities of crossover for the Maximal Generation Functions. . . . 144

6.8 Probability of crossover for each probability of mutation for Optimal Generation Functions 145
6.9 Probability of crossover for each population size for Optimal Generation Functions 146
6.10 The proposed GA calibration methodology 148
6.11 Probability distribution used for crossover with $\sigma = (p_1 - p_2)/6$ 149
6.12 Functional form of f1 and f2. 155
6.13 Functional form of f3 and f5. 155
6.14 Functional form of f6 and f8. 156
6.15 Functional form of f9 and f10. 156
6.16 Functional form of f11 and f12. 157
6.17 Functional form of f13 and f14. 157
6.18 Functional form of f15 and f16. 158
6.19 Functional form of f18 and f19. 158
6.20 Functional form of f20 and f21. 159
6.21 Functional form of f22 and f23. 159
6.22 GA convergence for the different calibration methods, for f12 with $l = 10$ 166

7.1 Schematic of the Cherry Hill-Brushy Plains Network 179
7.2 Schematic of the Woronora WDS EPANET model 190
7.3 Monthly demand patterns for the demand node at Maianbar. 193
7.4 Spatial distribution of the daily demand around the Woronora Water Distribution System (WDS). 194
7.5 Electricity costs over a 24 hour period. 195
7.6 Total chlorine trend at the outlet of Menai reservoir for December 2005. . 197
7.7 Influence of tablet dosing for December 2005. 198
7.8 Influence of tablet dosing for July 2006. 199
7.9 Calcium hypochlorite tablet dosing model. 199
7.10 Simulated and observed Helensburgh reservoir profile. 201
7.11 Simulated and observed Engadine reservoir profile. 202
7.12 Simulated and observed Heathcote reservoir profile. 203
7.13 Simulated and observed Lucas Heights reservoir profile. 203
7.14 Simulated and observed Menai reservoir profile. 204
7.15 Simulated and observed total chlorine concentrations at the Engadine pumping station. 205

7.16 Simulated and observed total chlorine concentrations at the Menai reservoir inlet. . . . 206
7.17 Helensburgh reservoir profile with original initial levels . . . 218
7.18 Engadine reservoir profile with original initial levels . . . 219
7.19 Heathcote reservoir profile with original initial levels . . . 220
7.20 Lucas Heights reservoir profile with original initial levels . . . 220
7.21 Menai reservoir profile with original initial levels . . . 221
7.22 Total chlorine concentration at demand nodes with original initial reservoir levels . . . 222
7.23 Helensburgh reservoir profile with low initial levels . . . 223
7.24 Engadine reservoir profile with low initial levels . . . 224
7.25 Heathcote reservoir profile with low initial levels . . . 224
7.26 Lucas Heights reservoir profile with low initial levels . . . 225
7.27 Menai reservoir profile with low initial levels . . . 226
7.28 Total chlorine concentration at demand nodes with low initial reservoir levels . . . 227

List of Tables

3.1 The test functions used in the parametric study, and the interval used for the decision variables. . . . 54
3.2 The GA parameter values used for the parametric study. . . . 55
3.3 Optimal Generation Function results. . . . 57
3.4 Maximal Generation Function results. . . . 58
3.5 The test functions with controllable characteristics used in the parametric study. . . . 74
3.6 The function values used for F3A. . . . 75
3.7 The function values used for F3B. . . . 75
3.8 The function values used for F3C. . . . 77
3.9 The function values used for F4A and F4B. . . . 78
3.10 The function values used for F6A and F6B. . . . 78
3.11 g_{conv} values observed for F3A. . . . 81
3.12 g_{conv} values observed for F3B. . . . 81
3.13 g_{conv} values observed for F3C. . . . 82
3.14 g_{conv} values observed for F4A and F4B. . . . 83
3.15 g_{conv} values observed for F6A and F6B. . . . 84

4.1 Gene epistasis test functions. . . . 102
4.2 Interactions by mutual information and joint probability. . . . 103
4.3 Interactions by mutual information and residuals. . . . 104
4.4 Separability measure results. . . . 107
4.5 Separability measure for F3A with $m_{\text{BB}} = 1$, $\delta_{\text{BB}} = 1$. . . . 108
4.6 Separability measure for F3A with $m_{\text{BB}} = 1$, $\delta_{\text{BB}} = 2$. . . . 108
4.7 Separability measure for F3A with $m_{\text{BB}} = 1$, $\delta_{\text{BB}} = 4$. . . . 108
4.8 Separability measure for F3A with $m_{\text{BB}} = 2$, $\delta_{\text{BB}} = 1$. . . . 108
4.9 Separability measure for F3A with $m_{\text{BB}} = 2$, $\delta_{\text{BB}} = 2$. . . . 108
4.10 Separability measure for F3A with $m_{\text{BB}} = 4$, $\delta_{\text{BB}} = 1$. . . . 108

4.11 Salience results for F3C. 110
4.12 Salience results for F3C without epistatic interactions. 111

5.1 Parameters values used for the test functions. 116
5.2 Input selection results. 123

6.1 GA parameter interactions for Optimal Generation Functions. 137
6.2 GA parameter interactions for Maximal Generation Functions. 137
6.3 Test functions for the GA calibration methods. 154
6.4 Fitness function statistic values, and corresponding N for fitness functions with $l = 10$. 160
6.5 Fitness function statistic values for different sized search spaces for f6. . . 161
6.6 Predicted N for fitness functions with $l = 30$. 162
6.7 Predicted N for fitness functions with $l = 50$. 163
6.8 Population Sizes Predicted due to Genetic Drift 163
6.9 Overall rankings of the GA calibration methods 165
6.10 Average ranking of each GA calibration method for each fitness function. 168
6.11 Variance of best solutions found for $FE = 5 \times 10^5$ and $l = 50$. 170

7.1 Interactions 183
7.2 Results for different GA calibration methods with set parameters 185
7.3 Results for different GA calibration methods with self adaptive parameters 186
7.4 Average Rankings 186
7.5 Comparison of the solutions found by different optimisation methods . . . 187
7.6 Total mass of chlorine for different methods 188
7.7 Summary of reservoirs in Woronora EPANET model. 191
7.8 Woronora Decision Variables 208
7.9 Separability measure results for scenario 2 of the Woronora fitness function 212
7.10 Statistic values for the Woronora fitness function 213
7.11 Ranking of GA calibration methods for Woronora fitness function for $FE = 5 \times 10^3$ 214
7.12 Ranking of GA calibration methods for Woronora fitness function for $FE = 10^4$ 215
7.13 Ranking of GA calibration methods for Woronora fitness function for $FE = 10^5$ 215
7.14 Overall ranking of GA calibration methods for Woronora fitness function 216
7.15 Pumping cost of best solution found for different scenarios. 217

B.1 Predicted - Set Values. $l = 10$. f1-6 263
B.2 Predicted - Set Values. $l = 10$. f8-12 264
B.3 Predicted - Set Values. $l = 10$. f13-18 265
B.4 Predicted - Set Values. $l = 10$. f19-23 266
B.5 Predicted - Self Adaptive. $l = 10$. f1-6 267
B.6 Predicted - Self Adaptive. $l = 10$. f8-12 268
B.7 Predicted - Self Adaptive. $l = 10$. f13-18 269
B.8 Predicted - Self Adaptive. $l = 10$. f19-23 270
B.9 Drift - Set Values. $l = 10$. f1-6 271
B.10 Drift - Set Values. $l = 10$. f8-12 272
B.11 Drift - Set Values. $l = 10$. f13-18 273
B.12 Drift - Set Values. $l = 10$. f19-23 274
B.13 Drift - Self Adaptive. $l = 10$. f1-6 275
B.14 Drift - Self Adaptive. $l = 10$. f8-12 276
B.15 Drift - Self Adaptive. $l = 10$. f13-18 277
B.16 Drift - Self Adaptive. $l = 10$. f19-23 278
B.17 Parameterless - Set Values. $l = 10$. f1-6 279
B.18 Parameterless - Set Values. $l = 10$. f8-12 280
B.19 Parameterless - Set Values. $l = 10$. f13-18 281
B.20 Parameterless - Set Values. $l = 10$. f19-23 282
B.21 Parameterless - Self Adaptive. $l = 10$. f1-6 283
B.22 Parameterless - Self Adaptive. $l = 10$. f8-12 284
B.23 Parameterless - Self Adaptive. $l = 10$. f13-18 285
B.24 Parameterless - Self Adaptive. $l = 10$. f19-23 286
B.25 Typical - Set Values. $l = 10$. f1-6 287
B.26 Typical - Set Values. $l = 10$. f8-12 288
B.27 Typical - Set Values. $l = 10$. f13-18 289
B.28 Typical - Set Values. $l = 10$. f19-23 290
B.29 Typical - Self Adaptive. $l = 10$. f1-6 291
B.30 Typical - Self Adaptive. $l = 10$. f8-12 292
B.31 Typical - Self Adaptive. $l = 10$. f13-18 293
B.32 Typical - Self Adaptive. $l = 10$. f19-23 294
B.33 Predicted - Set Values. $l = 30$. f1-6 295
B.34 Predicted - Set Values. $l = 30$. f8-12 296
B.35 Predicted - Set Values. $l = 30$. f13-18 297

B.36 Predicted - Set Values. $l = 30$. f19-23 298
B.37 Predicted - Self Adaptive. $l = 30$. f1-6 299
B.38 Predicted - Self Adaptive. $l = 30$. f8-12 300
B.39 Predicted - Self Adaptive. $l = 30$. f13-18 301
B.40 Predicted - Self Adaptive. $l = 30$. f19-23 302
B.41 Drift - Set Values. $l = 30$. f1-6 303
B.42 Drift - Set Values. $l = 30$. f8-12 304
B.43 Drift - Set Values. $l = 30$. f13-18 305
B.44 Drift - Set Values. $l = 30$. f19-23 306
B.45 Drift - Self Adaptive. $l = 30$. f1-6 307
B.46 Drift - Self Adaptive. $l = 30$. f8-12 308
B.47 Drift - Self Adaptive. $l = 30$. f13-18 309
B.48 Drift - Self Adaptive. $l = 30$. f19-23 310
B.49 Parameterless - Set Values. $l = 30$. f1-6 311
B.50 Parameterless - Set Values. $l = 30$. f8-12 312
B.51 Parameterless - Set Values. $l = 30$. f13-18 313
B.52 Parameterless - Set Values. $l = 30$. f19-23 314
B.53 Parameterless - Self Adaptive. $l = 30$. f1-6 315
B.54 Parameterless - Self Adaptive. $l = 30$. f8-12 316
B.55 Parameterless - Self Adaptive. $l = 30$. f13-18 317
B.56 Parameterless - Self Adaptive. $l = 30$. f19-23 318
B.57 Typical - Set Values. $l = 30$. f1-6 319
B.58 Typical - Set Values. $l = 30$. f8-12 320
B.59 Typical - Set Values. $l = 30$. f13-18 321
B.60 Typical - Set Values. $l = 30$. f19-23 322
B.61 Typical - Self Adaptive. $l = 30$. f1-6 323
B.62 Typical - Self Adaptive. $l = 30$. f8-12 324
B.63 Typical - Self Adaptive. $l = 30$. f13-18 325
B.64 Typical - Self Adaptive. $l = 30$. f19-23 326
B.65 Predicted - Set Values. $l = 50$. f1-6 327
B.66 Predicted - Set Values. $l = 50$. f8-12 328
B.67 Predicted - Set Values. $l = 50$. f13-18 329
B.68 Predicted - Set Values. $l = 50$. f19-23 330
B.69 Predicted - Self Adaptive. $l = 50$. f1-6 331
B.70 Predicted - Self Adaptive. $l = 50$. f8-12 332

B.71 Predicted - Self Adaptive. $l = 50$. f13-18 333
B.72 Predicted - Self Adaptive. $l = 50$. f19-23 334
B.73 Drift - Set Values. $l = 50$. f1-6 335
B.74 Drift - Set Values. $l = 50$. f8-12 336
B.75 Drift - Set Values. $l = 50$. f13-18 337
B.76 Drift - Set Values. $l = 50$. f19-23 338
B.77 Drift - Self Adaptive. $l = 50$. f1-6 339
B.78 Drift - Self Adaptive. $l = 50$. f8-12 340
B.79 Drift - Self Adaptive. $l = 50$. f13-18 341
B.80 Drift - Self Adaptive. $l = 50$. f19-23 342
B.81 Parameterless - Set Values. $l = 50$. f1-6 343
B.82 Parameterless - Set Values. $l = 50$. f8-12 344
B.83 Parameterless - Set Values. $l = 50$. f13-18 345
B.84 Parameterless - Set Values. $l = 50$. f19-23 346
B.85 Parameterless - Self Adaptive. $l = 50$. f1-6 347
B.86 Parameterless - Self Adaptive. $l = 50$. f8-12 348
B.87 Parameterless - Self Adaptive. $l = 50$. f13-18 349
B.88 Parameterless - Self Adaptive. $l = 50$. f19-23 350
B.89 Typical - Set Values. $l = 50$. f1-6 351
B.90 Typical - Set Values. $l = 50$. f8-12 352
B.91 Typical - Set Values. $l = 50$. f13-18 353
B.92 Typical - Set Values. $l = 50$. f19-23 354
B.93 Typical - Self Adaptive. $l = 50$. f1-6 355
B.94 Typical - Self Adaptive. $l = 50$. f8-12 356
B.95 Typical - Self Adaptive. $l = 50$. f13-18 357
B.96 Typical - Self Adaptive. $l = 50$. f19-23 358

C.1 FE=10^3 - $l = 10$, f1-12 359
C.2 FE=10^3 - $l = 10$, f13-23 359
C.3 FE=10^3 - $l = 30$, f1-12 360
C.4 FE=10^3 - $l = 30$, f13-23 360
C.5 FE=10^3 - $l = 50$, f1-12 360
C.6 FE=10^3 - $l = 50$, f13-23 361
C.7 FE=10^4 - $l = 10$, f1-12 361
C.8 FE=10^4 - $l = 10$, f13-23 361
C.9 FE=10^4 - $l = 30$, f1-12 362

C.10 FE=10^4 - $l = 30$, f13-23 362
C.11 FE=10^4 - $l = 50$, f1-12 362
C.12 FE=10^4 - $l = 50$, f13-23 363
C.13 FE=10^5 - $l = 10$, f1-12 363
C.14 FE=10^5 - $l = 10$, f13-23 363
C.15 FE=10^5 - $l = 30$, f1-12 364
C.16 FE=10^5 - $l = 30$, f13-23 364
C.17 FE=10^5 - $l = 50$, f1-12 364
C.18 FE=10^5 - $l = 50$, f13-23 365
C.19 FE=3×10^5 - $l = 30$, f1-12 365
C.20 FE=3×10^5 - $l = 30$, f13-23 365
C.21 FE=5×10^5 - $l = 50$, f1-12 366
C.22 FE=5×10^5 - $l = 50$, f13-23 366

D.1 Predicted - Set Values 367
D.2 Rogers - Set Values 368
D.3 Parameterless - Set Values 369
D.4 Typical - Set Values 370
D.5 Predicted - Self Adaptive 371
D.6 Rogers - Self Adaptive 372
D.7 Parameterless - Self Adaptive 373
D.8 Typical - Self Adaptive 374

List of Symbols

Fitness Function Symbols

A	Amplitude
k_{BB}	Building Block Size
δ_{BB}	Defining Length of a Building Block
f	Frequency
l	Problem Size
M	Transformation Matrix
m_{BB}	Building Block Number
ϕ	Phase Shift
π	Mathematical Constant, Pi

Fitness Function Statistic Symbols

D	Dominance Statistic
λ	Separability Measure
R_{av}	Average Correlation
R_{l}	Correlation Length
R_{T}	Total Correlation

General Sampling Symbols

β	Fourier Series Rotation
d	Distance
h	Smoothing Parameter (Bin Width or Bandwidth)
I	Mutual Information
I_c	Moran's I
n	Number of Samples
NI	Normalised Mutual Information
PI	Partial Mutual Information
ρ	Standard Autocorrelation
R	Analytic Autocorrelation
R_{s}	Spatial Autocorrelation

R_{t}	Temporal Autocorrelation
w	Weighting Function
Genetic Algorithm Symbols	
c	Fraction Used for the Standard Deviation of Crossover Distribution
e	Number of Elite Solutions
FE	Function Evaluations
g	Number of Generations
g_{conv}	Number of Generations Before Convergence
N	Population Size
p_m	Probability of Mutation
p_c	Probability of Crossover
σ	Standard Deviation of Crossover Distribution
s	Selection Pressure
Quantitative Genetics Symbols	
i	Selection Intensity
k	Decay of Population Variance
σ_{pop}	Standard Deviation of the Population
σ_{P}	Standard Deviation of the Fitness Function
S	Selection Differential
Water Distribution System Symbols	
k_{cl}	Chlorine Decay Rate

List of Acronyms

ABCO	Aggregation-Based Crossover Operator
ACOA	Ant Colony Optimisation Algorithm
AICV	Automatic Inlet Control Valve
CWS	Clear Water Storage
DCO	Discrete Crossover Operator
EA	Evolutionary Algorithm
FDC	Fitness Distance Correlation
GA	Genetic Algorithm
HCO	Hybrid Crossover Operator
KDE	Kernel Density Estimation
MI	Mutual Information
MSE	Mean Squared Error
NBCO	Neighbourhood-Based Crossover Operator
pdf	probability density function
PMI	Partial Mutual Information
PRV	Pressure Reducing Valve
RCGA	Real Coded Genetic Algorithm
SCADA	Supervisory Control And Data Acquisition
SFLA	Shuffled Frog Leaping Algorithm
TCV	Throttle Control Valve
WDS	Water Distribution System
WFP	Water Filtration Plant
WTP	Water Treatment Plant
WQ	Water Quality