Developing Parsimonious and Efficient Algorithms for Water Resources Optimization Problems

by

Masoud Asadzadeh Esfahani

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Civil Engineering

Waterloo, Ontario, Canada, 2012

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In the current water resources scientific literature, a wide variety of engineering design problems are solved in a simulation-optimization framework. These problems can have single or multiple objective functions and their decision variables can have discrete or continuous values. The majority of current literature in the field of water resources systems optimization report using heuristic global optimization algorithms, including evolutionary algorithms, with great success. These algorithms have multiple parameters that control their behavior both in terms of computational efficiency and the ability to find near globally optimal solutions. Values of these parameters are generally obtained by trial and error and are case study dependent. On the other hand, water resources simulation-optimization problems often have computationally intensive simulation models that can require seconds to hours for a single simulation. Furthermore, analysts may have limited computational budget to solve these problems, as such, the analyst may not be able to spend some of the computational budget to fine-tune the algorithm settings and parameter values. So, in general, algorithm parsimony in the number of parameters is an important factor in the applicability and performance of optimization algorithms for solving computationally intensive problems.

A major contribution of this thesis is the development of a highly efficient, single objective, parsimonious optimization algorithm for solving problems with discrete decision variables. The algorithm is called Hybrid Discrete Dynamically Dimensioned Search, HD-DDS, and is designed based on Dynamically Dimensioned Search (DDS) that was developed by Tolson and Shoemaker (2007) for solving single objective hydrologic model calibration problems with continuous decision variables. The motivation for developing HD-DDS comes from the parsimony and high performance of original version of DDS. Similar to DDS, HD-DDS has a single parameter with a robust default value. HD-DDS is successfully applied to several benchmark water distribution system design problems where decision variables are pipe sizes among the available pipe size options. Results show that HD-DDS exhibits superior performance in specific comparisons to state-of-the-art optimization algorithms.

The parsimony and efficiency of the original and discrete versions of DDS and their successful application to single objective water resources optimization problems with discrete and continuous decision variables motivated the development of a multi-objective optimization algorithm based on DDS. This algorithm is called Pareto Archived Dynamically Dimensioned Search (PA-DDS). The algorithm parsimony is a major factor in the design of PA-DDS. PA-DDS has a single parameter from

its search engine DDS. In each iteration, PA-DDS selects one archived non-dominated solution and perturbs it to search for new solutions. The solution perturbation scheme of PA-DDS is similar to the original and discrete versions of DDS depending on whether the decision variable is discrete or continuous. So, PA-DDS can handle both types of decision variables. PA-DDS is applied to several benchmark mathematical problems, water distribution system design problems, and water resources model calibration problems with great success.

It is shown that hypervolume contribution, HVC1, as defined in Knowles et al. (2003) is the superior selection metric for PA-DDS when solving multi-objective optimization problems with Pareto fronts that have a general (unknown) shape. However, one of the main contributions of this thesis is the development of a selection metric specifically designed for solving multi-objective optimization problems with a known or expected convex Pareto front such as water resources model calibration problems. The selection metric is called convex hull contribution (CHC) and makes the optimization algorithm sample solely from a subset of archived solutions that form the convex approximation of the Pareto front. Although CHC is generally applicable to any stochastic search optimization algorithm, it is applied to PA-DDS for solving six water resources calibration case studies with two or three objective functions. These case studies are solved by PA-DDS with CHC and HVC1 selections using 1,000 solution evaluations and by PA-DDS with CHC selection and two popular multi-objective optimization algorithms, AMALGAM and ε-NSGAII, using 10,000 solution evaluations. Results are compared based on the best case and worst case performances (out of multiple optimization trials) from each algorithm to measure the expected performance range for each algorithm. Comparing the best case performance of these algorithms shows that, PA-DDS with CHC selection using 1,000 solution evaluations perform very well in five out of six case studies. Comparing the worst case performance of the algorithms shows that with 1,000 solution evaluations, PA-DDS with CHC selection perform well in four out of six case studies. Furthermore, PA-DDS with CHC selection using 10,000 solution evaluations perform comparable to AMALGAM and ε-NSGAII. Therefore, it is concluded that PA-DDS with CHC selection is a powerful optimization algorithm for finding high quality solutions of multi-objective water resources model calibration problems with convex Pareto front especially when the computational budget is limited.

Acknowledgements

I would like to take this opportunity to express my special gratitude to Dr. Bryan Tolson for his continuous guidance and support throughout my PhD studies. Bryan is an intelligent, enthusiastic, responsible and patient advisor. His encouraging and elegant comments and thoughts have always helped me step forward in my research. I would also like to thank my co-advisor, Dr. Donald Burn for his constructive comments on my research.

It is my pleasure to thank my defense committee members Dr. Patrick Reed from the Pennsylvania State University, Dr. Shawn Matott from the State University of New York at Buffalo, Dr. Kumaraswamy Ponnambalam and Dr. Liping Fu from the University of Waterloo.

I would like to thank my dear wife, Mahnaz, for being patient, helpful and encouraging. I would also like to express my sincere gratitude to my Parents Zohreh and Mehdi and my siblings Laleh and Mahmood. Without their encouragement, I would not have a chance to be at the University of Waterloo.

Table of Contents

AUTHOR'S DECLARATION ii
Abstract iii
Acknowledgements v
Table of Contents vi
List of Figures ix
List of Tables xiii
Chapter 1 Introduction 1
1.1 Problem Statement and Research Objectives 1
1.2 Thesis Structure and Research Contributions 4
Chapter 2 Hybrid Discrete Dynamically Dimensioned Search (HD-DDS) Algorithm for Water Distribution System Design Optimization 6
Summary 6
2.1 Introduction 7
2.2 Methodology 9
2.2.1 Components of the Hybrid Discrete Dynamically Dimensioned Search Algorithm 9
2.2.2 Benchmark Optimization Algorithms 19
2.2.3 Benchmark WDS Design Studies 19
2.2.4 Optimization Model Formulation for HD-DDS 20
2.2.5 Outline of Algorithm Comparisons 21
2.3 Results 22
2.3.1 HD-DDS Component Performance Assessment 22
2.3.2 HD-DDS Performance Relative to Benchmark Algorithms 24
2.4 Discussion and Conclusions 28
Chapter 3 Hybrid Pareto Archived Dynamically Dimensioned Search for Multi-Objective Combinatorial Optimization: Application to Water Distribution Network Design 31
Summary 31
3.1 Introduction 32
3.1.1 Hybridizing MO algorithms 33
3.1.2 Comparing MO algorithm performance 34
3.2 Methodology 34
3.2.1 Pareto Archived Dynamically Dimensioned Search (PA-DDS) 35

3.2.2 Hybrid PA-DDS for Combinatorial Multi-objective Optimization Problems 36
3.2.3 Optimization Problem Formulation 36
3.2.4 Benchmark Water Distribution System Design Problems 37
3.2.5 Selected performance metrics 39
3.2.6 Benchmark optimization algorithms 41
3.3 Results 42
3.3.1 Hybrid PA-DDS versus NSGAII and SPEA2 42
3.3.2 Local search performance assessment 46
3.4 Discussion and Conclusions 46
Chapter 4 Pareto Archived Dynamically Dimensioned Search with Hypervolume Based Selection for Multi-objective Optimization 50
Summary 50
4.1 Introduction 51
4.2 Methodology 52
4.2.1 Selection Metric 52
4.3 Numerical Experiments 55
4.3.1 Experiment 1: Choosing the Selection Metric for PA-DDS 55
4.3.2 Experiment 2: Solving a Water Resources Multi-objective Optimization Problem 56
4.3.3 Experiment 3: Solving CEC09 Problems with PA-DDS 59
4.3.4 Results Comparison Approach 59
4.4 Results 62
4.4.1 Results of Experiment 1 63
4.4.2 Results of Experiment 2 65
4.4.3 Results of Experiment 3 67
4.4.4 PA-DDS Algorithm Runtime Analysis and Limitations 69
4.5 Discussion and Conclusions 72
Chapter 5 Convex Hull Contribution, a Novel Selection Metric for Convex Multi-objective Optimization Problems: Application to Water Resources Calibration Problems 74
Summary 74
5.1 Introduction 75
5.2 Methodology 78
5.2.1 Convex Hull Background 78

5.2.2 Convex Hull Contribution (CHC) Selection Metric 82
5.3 Numerical Experiments 84
5.3.1 Mathematical Multi-objective Optimization Problems 86
5.3.2 MO Water Resources Model Calibration 87
5.3.3 Results Comparison Approach 92
5.4 Results 93
5.4.1 CHC versus HVC for Mathematical Problems with Convex Pareto Front 93
5.4.2 PA-DDS Performance Assessment in Water Resources Calibration Problems 95
5.4.3 Calibrated Model Simulation Results Comparison 103
5.5 Conclusions 107
Chapter 6 Summary, Conclusions and Recommendations for Future Work 108
6.1 Summary 108
6.2 Conclusions and Guidelines for Users of the Developed Algorithms 109
6.3 Recommendations for Future Work 113
References 116
Appendix A Pseudo Codes 125

List of Figures

Figure 2-1 Example discrete DDS probability mass functions for candidate option numbers $xinew$ for a single decision variable with 16 possible options (A and B) and 6 possible options (C and D) under various values for x_i^{best}. Default Discrete DDS parameter $r = 0.2$. 12
Figure 2-2 Progress as of the end of each step of the HD-DDS algorithm (see appendix A-Figure A-5) versus average number of solution evaluations for various WDS case studies and corresponding total computational budget input to HD-DDS. 50 optimization trials are shown for each case study except for NYTP2 with 20 optimization trials. The numbers in brackets count the number of trials where the best solution so far is equal to the best known solution. 23
Figure 2-3 Empirical CDF of best solutions from HD-DDS and MMAS ACO algorithm for the A) NYTP, B) HP and C) NYTP2 for approximately the same number of solution evaluations (see the figure for NFEtotal in brackets). MMAS results are for 20 optimization trials. HD-DDS results for NYTP and HP are for 50 optimization trials while NYTP2 are for 20 optimization trials. 25
Figure 2-4 HD-DDS performance with different computational budgets (NFEtotal) compared to other algorithm performance on Balerma network. HD-DDS results show all 10 trials. Other algorithm results are for a single trial or show the range of results from multiple trials. 26
Figure 3-1 Normalized Hypervolume Performance Metric for Minimization of Two Objectives 39
Figure 3-2 CNHV performance metric for minimization of two objectives 41
Figure 3-3 Comparison between the best attained front and the worst CNHV trials of all three algorithms solving each case study 44
Figure 3-4 Comparison between the best CNHV trial of Hybrid PA-DDS and NSGAII for solving (A) HP, Extreme solutions (17678.5 m, $1.803M) and (0 m, $7.470M) are used to normalize the fronts, and (B) BP, Extreme solutions (5213.7.5 m, €0.724M) and (0 m, €2.3620M) are used to normalize the fronts. 45
Figure 4-1 Comparison between CD (line segment lengths), HVC1 and HVC2 (both are areas) for solution $\boldsymbol{x^p}$ in an example bi-objective optimization problem. 54
Figure 4-2 ZDT4, empirical CDF plots for Additive epsilon Indicator based on final results of 50 independent trials of the PA-DDS with RND, CD, HVC1 or HVC2 selection. (A) computational budget = 2,500, (B) computational budget = 25,000. Vertical line at 0 represents perfect result. 63
Figure 4-3 Empirical CDFs for 1-NHV, IGD, GD and ε+ Indicator performance metrics based for Experiment 2 (3-objective hydrologic model calibration). Vertical lines at 0 represent ideal results. 66

Figure 4-4 Portion of total PA-DDS runtime solely spent on evaluation of objective functions based on solving 10 different MO test problems with 2, 3 and 5 objectives and with four computational budgets of 10^3, 10^4, 10^5, 10^6. White dots represent timings based on actual optimization trials with observed average solution evaluation times while the solid black dots are based on the same optimization trials using artificially increased average solution evaluation times. The solid and dashed lines represent the most and the least algorithmically complex trials: a 3-objective MO problem solved with the budget of 10^6 and a BOP solved with the budget of 10^3, respectively..... 72
Figure 5-1 a set of non-dominated points in the objective space of three example bi-objective minimization problems and the Convex approximation of Pareto front based on these points. A) the convex curve accurately approximates the Pareto front, B) the convex curve does not some represent a portion of the Pareto front, C) the convex curve does not represent the whole Pareto front..... 76
Figure 5-2 Convex Hull of a set of points in a two dimensional space 79
Figure 5-3 Convex Hull of a set of non-dominated points in a normalized bi-objective minimization problem..... 80
Figure 5-4 Convex Hull Contribution (CHC) calculations for an active non-dominated solution in the normalized objective space of an example bi-objective optimization problem 83
Figure 5-5 The true Pareto front of DTLZ2 with two objective functions..... 87
Figure 5-6 C-Town water distribution system layout..... 88
Figure 5-7 UF3, empirical CDF plot based on ε+ Indicator and final results (Pareto approximate fronts) of 50 independent trials of PA-DDS_CHC and PA-DDS_HVC and with computational budget of 2,500 and 25,000 solution evaluations. Vertical line at 0 represents perfect result. 94
Figure 5-8 EPANET2 calibration problem, Pareto approximate fronts of A) selected best trial and B) selected worst trial of PA-DDS with CHC and HVC selections with 1,000 solution evaluations and PA-DDS with CHC selection, ε-NSGAII and AMALGAM with 10,000 solution evaluations. The different scale of axes is due to the significant difference in the quality of results in A and B 97
Figure 5-9 SWAT 2000 calibration problem, Pareto approximate fronts of A) selected best trial and B) selected worst trial of PA-DDS with CHC and HVC selections with 1,000 solution evaluations and PA-DDS with CHC selection, ε-NSGAII and AMALGAM with 10,000 solution evaluations..... 98
Figure 5-10 SAC-SMA calibration problem, Pareto approximate fronts of A) selected best trial and B) selected worst trial of PA-DDS with CHC and HVC selections with 1,000 solution evaluations and PA-DDS with CHC selection, ε-NSGAII and AMALGAM with 10,000 solution evaluations..... 99

Figure 5-11 HYMOD calibration problem, Pareto approximate fronts of A) selected best trial and B) selected worst trial of PA-DDS with CHC and HVC selections with 1,000 solution evaluations and PA-DDS with CHC selection, ε-NSGAII and AMALGAM with 10,000 solution evaluations. 101
Figure 5-12 SWAT 2003 calibration problem, Pareto approximate fronts of A) selected best trial and B) selected worst trial of PA-DDS with CHC and HVC selections with 1,000 solution evaluations and PA-DDS with CHC selection, ε-NSGAII and AMALGAM with 10,000 solution evaluations. 102
Figure 5-13 SAC-SMA, time series of observed versus simulated flow in the Leaf River watershed for selected solutions from the worst trials of PA-DDS_CHC_1000 and alternative solution from the worst MO trials with 10,000 solution evaluations. 105
Figure 5-14 EPANET2, time series of observed versus simulated data in C-Town for selected solutions from the worst trials of PA-DDS_CHC_1000 and alternative solution from the worst MO trials with 10,000 solution evaluations. 105
Figure 5-15 SWAT 2000, time series of observed versus simulated Flow (A), Sediment (B) and Phosphorous (C) in the Town Brook watershed for selected solutions from the worst trials of PA-DDS_CHC_1000 and alternative solution from the worst MO trials with 10,000 solution evaluations. 106
Figure A- 1 Discrete Dynamically Dimensioned Search (Discrete DDS) algorithm. $\mathbf{N}(0,1)$ generates random numbers from the standard Normal Distribution. $\mathbf{U}lb, ub$ generates random numbers from the Uniform distribution in the lb, ub interval. 125
Figure A- 2 Evaluating the objective function in HD-DDS. Note that $\mathrm{cost}(\boldsymbol{x})$ calculates the cost of the network based on the diameter and length of pipes, $f(\boldsymbol{x})$ is the objective function, $g\boldsymbol{x}$ is the summation of pressure violations at all nodes in the network. 126
Figure A- 3 Outline of local search L1 for constrained WDS optimization problem that can identify a local minimum such that no further improvement is possible by changing one decision variable (pipe) at a time 126
Figure A- 4 Outline of local search L2 for constrained WDS optimization problem that can identify a local minimum such that no further improvement is possible by changing two decision variables (pipe) at a time 127
Figure A- 5 The Hybrid Discrete Dynamically Dimensioned Search (HD-DDS) algorithm. 128
Figure A- 6 Pareto Archived Dynamically Dimensioned Search (PA-DDS) Algorithm. $\mathbf{U}lb, ub$ generates random numbers from the Uniform distribution in the lb, ub interval. 129
Figure A- 7 Archive strategy of PA-DDS 129

Figure A- 8 Select function based on roulette wheel selection. $\mathbf{U}(0,1)$ generates random numbers from Uniform distribution in the $(0,1)$ interval. 130
Figure A- 9 Perturb function (search strategy) of PA-DDS. $\mathbf{U}(0,1)$ Uniform distribution in the $(0,1)$ interval. $\mathbf{N}(0,1)$ generates random numbers from Standard Normal distribution. 131
Figure A- 11 Local Search L to polish one solution on the approximate front 133

List of Tables

Table 2-1 Characteristics of various optimization algorithms recently applied to water distribution system optimization. 8
Table 2-2 Summary of HD-DDS and other algorithm performance for the five WDS case studies investigated in this study. 27
Table 3-1 Algorithm comparison based on NHV and CNHV. **BOLD** numbers represent the best result for each case study, and *ITALIC* numbers corresponds to trials whose Pareto approximate front is plotted in Figure 3-3 43
Table 3-2 Evaluating the effectiveness of the local search for improving the extreme points of the approximate fronts based on the average percent improvement in CNHV. 46
Table 4-1 Algorithm parameter values for AMALGAM and ε-NSGAII for solving the MO hydrologic model calibration problem 58
Table 4-2 Statistical comparison of selection metric performance from Experiment 1 based on the Wilcoxon rank-sum test. P-values are based on the sample size of 50 and compare the results of PA-DDS with two most effective selection metrics (referred to as Preferred and Alternative) designated based on stochastic dominance analysis (visual comparison) of CDF plots for all four performance metrics. **BOLD** names highlight selection metrics that are unambiguously preferred because the preferred metric yields a significantly different empirical CDF (P-value < 0.05) and stochastically dominates the alternative selection metric. 64
Table 4-3 Statistical comparison of PA-DDS, ε-NSGAII and AMALGAM in Experiment 2 based on the two-sided Wilcoxon rank-sum test. P-values are based on a sample size of 10 and compare of PA-DDS with HVC1 selection to each of the ε-NSGAII and AMALGAM algorithms. **BOLD** P-values highlight when PA-DDS is unambiguously preferred because it yields a significantly different empirical CDF (P-value < 0.05) and stochastically dominates the alternative selection metric. 66
Table 4-4 PA-DDS performance in comparison with 13 MO algorithms based on IGD for solving problems UF1 to UF10 from CEC09 competition. Algorithms sorted by average rank. IGD value is the average of 30 optimization trials each using a computational budget of 300,000. **BOLD** numbers highlight the best IGD value in each test problem. 67
Table 4-5 PA-DDS performance in comparison with 13 MO algorithms based on IGD for solving problems UF11 to UF13 from CEC09 competition. Algorithms sorted by average rank. IGD value is

the average of 30 optimization trials each using a computational budget of 300,000. **BOLD** numbers highlight the best IGD value in each test problem. 68

Table 4-6 Statistics of IGD, GD, ε+ Indicator, and $1 - {NHV_{PA-DDS}}/{NHV_{ref-set}}$ performance metrics for solving problems UF1 to UF13 from CEC09 competition using PA-DDS. Minimum, average, maximum, and standard deviation based on 30 optimization trials each using 300,000 solution evaluations. 69

Table 5-1 Parameter name and range, lower bound (lb) to upper bound (ub) used for the hydrologic model calibration problems SAC-SMA, HYMOD, and SWAT2003 case studies. 91

Table 5-2 Statistical comparison of selection metric performance for solving seven mathematical problems based on the Wilcoxon rank-sum test. Numbers are P-values based on the sample size of 50 and compare the results of PA-DDS_CHC to PA-DDS_HVC. P-value<0.05 with an asterisk shows clear preference. 94

Table 5-3 Statistical analysis of results comparing PA-DDS with CHC selection to ε-NSGAII and AMALGAM at the budget of 10,000. P-values from the two-sided Wilcoxon rank-sum test are based on a sample size of 10. P-value<0.05 with an asterisk shows clear preference. 95

Table 5-4 Statistical analysis of results comparing CHC and HVC selections for PA-DDS at the budget of 1,000. P-values from the two-sided Wilcoxon rank-sum test are based on a sample size of 50 for C-Town and Town Brook case studies and 10 for other problems. P-value<0.05 with an asterisk shows clear preference. 96

Table 5-5 Objective function (simulation error metric) values of selected solutions (calibrated models) from PA-DDS_CHC_1000 and the alternative solution by optimization runs with 10,000 solution evaluations. Solutions from the best and worst performed optimization are separated..... 103

Table 5-6 Some statistics of simulation error in calibrated model by PA-DDS_CHC_1000 versus the alternative calibrated model with 10,000 solution evaluations. The selected solutions from the best and worst performed optimization trials are separately studied. 104

Chapter 1
Introduction

1.1 Problem Statement and Research Objectives

Water resources optimization problems can often be cast as simulation-optimization problems. This thesis represents the development of two heuristic optimization algorithms tailored to work effectively for solving such problems in the form of single- and multi-objective optimization problems. The optimization problems considered in this thesis are hydrologic model calibration and water distribution network design and calibration problems in the form of single- or multi-objective optimization problems with continuous and/or discrete valued decision variables.

Calibration is the process of tuning parameters of a mathematical simulation model that simulates the behaviour of an environmental or water resources system. The parameters can vary within a given plausible range and impact the performance of the model, e.g. goodness of fit between simulations and observations of system response. The calibration task can be done manually via a trial-and-error effort by the analyst, or automatically, via an optimization algorithm. The calibration objective is to optimize a single or multiple error functions that measure the goodness of fit between the observed and simulated system response values. Therefore, the automatic calibration of a model can be set up as a single- or multi-objective optimization problem with continuous decision variables to optimize the error metrics.

Water distribution network design problems can also be set up and solved in a simulation-optimization framework. In such problems, the analyst is designing a new network or rehabilitating an existing network by identifying the optimal pipe diameters which are chosen from a finite set of available pipe size options in the market. The optimal design can be the least-cost design which maintains some design constraints such as minimum allowable nodal pressure throughout the network. So, the problem can be solved as a single objective constrained optimization problem with discrete-valued decision variables or a bi-objective unconstrained optimization problem where one of the objectives is to minimize an aggregation of all constraint violations.

Water resources optimization problems are well-known to be highly complex and non-linear problems (Nicklow et al. 2010). So, the direct application of traditional linear and non-linear optimization techniques to these problems is limited, and the majority of current literature in the field of water resources systems optimization report using heuristic global optimization algorithms,

including evolutionary algorithms, with great success. These algorithms have multiple control parameters whose value can significantly impact the algorithm performance both in terms of computational efficiency and the ability to find near globally optimal solutions. For example, Genetic Algorithm-based optimization tools often require the user set the value of at least three algorithm parameters namely population size, probability of mutation, and probability of crossover. The value of these parameters can be fine-tuned to maximize the algorithm performance. However, in practice most studies employ an ad-hoc approach for conducting these parameter setting experiments and such experiments can be extremely computationally expensive, as many of the parameters are dependent. This issue is exacerbated for problems with long objective function simulation times. In addition, there are many situations where the objective function runtime and/or the available computational budget (time) to solve the problem dictate that algorithm parameters should be assigned knowing that only a limited number of solutions can be evaluated during optimization. In other words, optimal or even reasonable algorithm parameter values typically change (especially for evolutionary algorithms) as the total number of solutions evaluated during optimization changes, clearly demonstrated in Reed et al. (in press).

During my PhD, I was involved in the optimization of several water resources problems relying on numerical simulation models to evaluate objectives that took seconds to minutes to evaluate for a single solution with Intel ® Core ™2 Quad CPU Q 6600 @ 2.4 GHz with 3.25 GB of RAM. These problems include hydrologic model calibration problems with a single simulation runtime from a fraction of a second to more than 15 minutes, reservoir operation problems with a simulation runtime of 40 minutes, water distribution system design problems with a simulation runtimes ranging from a fraction of a second to more than 5 seconds and sorptive barrier design problems with a simulation runtime of 5 to 32 seconds on the abovementioned computer. Considering the above runtimes and a computational budget of only one day, an analyst would be able to conduct only 36 simulation model runs for the abovementioned reservoir operation problem. This number would be 96, 2700, and 17280 for the longest simulation runtime of the abovementioned hydrologic model calibration, sorptive barrier design, and water distribution system design problems, respectively. The range of simulation runs available if the computational budget was increased to seven days becomes roughly 250 to 121000. Clearly, the above examples show that analysts will be challenged to spend some of their computational budget to fine-tune the algorithm settings and parameter values.

In actuality, the abovementioned optimization problems were solved using computational budgets resulting in between 500 and 10,000,000 solutions being evaluated and due to computational burden,

many were solved by evaluating fewer than 10,000 solutions. With computational budgets measured in units of total solutions evaluated, the current heuristic optimization literature clearly indicates at least two very different types of optimization algorithms to solve problems on the range of computational budgets. For example, in multi-objective optimization studies on a limited budget (i.e., 500 or fewer solution evaluations), meta-modelling approaches such as ParEGO (Knowles 2006) and SMS-EGO (Ponweiser et al. 2008) are examples of state-of-the-art. For higher budgets (10,000 and higher), evolutionary algorithms are common and in water resources optimization studies, ε-NSGAII (Kollat and Reed 2005) and AMALGAM (Vrugt and Robinson 2007) appear to be the most popular. Therefore, for many practical multi-objective optimization studies limited to a budget of under 10,000 and above 500 solution evaluations, it is unclear what optimization algorithms should be utilized. In general though, the aforementioned optimization algorithms, as well as many others, would require algorithm parameter adjustment to work reasonably well on computational budgets ranging from a few hundred to 10,000 or more solutions.

The above discussion indicates that, in general, algorithm parsimony in the number of parameters is an important factor for the applicability and performance of optimization algorithms for solving problems with long simulation times. The most obvious evidence in the literature for this claim is the research attempts to make algorithm parameters self-adaptive. In this regard, evolution strategy based algorithms are unique because they embed algorithm parameters as decision variables (Beyer 1995). Schaffer and Morishima (1987) appear to be the first study that proposed adaptive parameter setting in genetic algorithms. Parameter control is the terminology that is used for heuristic parameter adjustment during the search of evolutionary algorithms in contrast to the fixed parameter setting which requires parameter fine tuning before the search which is usually very time consuming (Eiben et al. 1999). Moreover, "Parameter setting in evolutionary algorithms" (Lobo et al. 2007) is an entire book that aims to address challenges in parameter tuning of optimization algorithms. Beyond self-adaptive parameters, Reed et al. (in press) conclude the future of MOEAs is self-adaptive search algorithms that adapt operator selection as well as parameters during the search.

The main goal of this thesis was to develop new parsimonious algorithms that are robust under default configurations for solving water resources optimization problems across a wide range of computational budgets. Here, robust means an algorithm is at least competitive, perhaps better, relative to the performance of selected state-of-the-art optimization algorithms. Robustness at low computational budgets implicitly makes the algorithms appear computationally efficient. Although parsimonious algorithms eliminate or reduce the demands of parameter-tuning experiments,

parsimony does not guarantee algorithm insensitivity to the parameter values. Importantly, this thesis did not investigate how to change algorithm parameter values of existing algorithms like ε-NSGA and AMALGAM, for problems to be solved with limited computational budgets (i.e., < 10,000 evaluations). The algorithms developed in this thesis are novel derivatives of DDS (Tolson and Shoemaker 2007). Each algorithm has a single parameter that has a robust default value intended to eliminate the need for analysts to fine tune the algorithm setting for a new problem. The developed algorithms are designed to handle both types of continuous and discrete decision variables without any changes to the algorithm settings. Also, one of the two developed algorithms is designed to handle multiple objective functions. The developed algorithms are designed to be robust when the computational budget changes by self-adapting algorithm behaviour to the pre-specified computational budget.

1.2 Thesis Structure and Research Contributions

This thesis is structured around published or accepted manuscripts. Chapter 2 and Chapter 3 are based on and thus very closely mirror two published manuscripts, Tolson et al. (2009) and Asadzadeh and Tolson (2012) respectively. Chapter 4 is based on an accepted manuscript in the Engineering Optimization journal by Asadzadeh and Tolson but the Chapter will be much more detailed (more results) than the eventual published manuscript. Chapter 5 is written in the format of a journal manuscript to be submitted to a water resources related journal. Although each of the manuscripts has one or more co-authors, my contribution to the design of the algorithms and writing of the manuscripts equalled or exceeded 50% in each case. Except for very brief concluding paragraphs in each chapter, extensive conclusions and recommendations for future work are detailed in Chapter 6.

This research has made the following main contributions to the water resources optimization research area:

- The development of Hybrid Discrete Dynamically Dimensioned Search or the HD-DDS algorithm that uses two straightforward local search techniques. HD-DDS is an efficient and parsimonious optimization algorithm for solving single objective optimization problems with discrete decision variables. The two local search techniques in this work are specifically developed for discrete optimization problems where the objective function and constraints behave monotonically by decreasing the value of each decision variable. HD-DDS is applied to water distribution system design problems where the decision maker desires a set of available pipe sizes that costs as low as possible but keeps the hydraulic pressure in the

system above the minimum requirements. Details about the HD-DDS algorithm and results are reported in Chapter 2.

- The development of Pareto Archived DDS or the PA-DDS parsimonious multi-objective optimization algorithm that can handle both discrete and continuous decision variables. Its development is motivated by the efficiency and parsimony of HD-DDS and the original version of DDS (Tolson and Shoemaker 2007). In Chapter 3, the PA-DDS algorithm and its successful application to water distribution system design problems are reported. The algorithm is hybridized by a straightforward local search technique specifically designed for multi-objective optimization problems with discrete decision variables. The local search technique is shown to be applicable to other algorithms with great success, especially when the computational budget for the global search by the main algorithm is limited. Furthermore, in Chapter 4, it is shown that hypervolume contribution as introduced in Knowles et al. (2003) is the most effective selection metric for solving general multi-objective optimization problems based on the results of solving several benchmark test problems with up to five objective functions and a hydrologic model calibration problem with three objective functions.

- The development of Convex Hull Contribution or CHC selection metric for solving multi-objective optimization problems with convex Pareto fronts. CHC makes the multi-objective heuristic optimization algorithms focus on solutions that form the convex shape of the Pareto front and therefore approximate them accurately. The scientific literature of multi-objective water resources calibration problems shows that these problems are expected to have convex Pareto fronts. So, CHC is utilized as the selection metric in PA-DDS, and this version of PA-DDS is successfully applied to several multi-objective water resources calibration problems. Details about this contribution are reported in Chapter 5.

Consistently in this thesis, all equations use the same notation. A lower case italic single letter represents a single value variable (e.g., a decision variable value, x_i), a lower case italic bold single letter represents a set of single value variables such as a solution (e.g., $\boldsymbol{x} = \{x_1, \cdots, x_n\}$), a capital italic bold single letter represents a subset of space (e.g., $\boldsymbol{X} \subseteq R^n$), and a capital multi-letter is an acronym such as PS for Pareto optimal set of solutions and PF for Pareto optimal Front (objective function values).

Chapter 2

Hybrid Discrete Dynamically Dimensioned Search (HD-DDS) Algorithm for Water Distribution System Design Optimization

This chapter is based on the published article with the same title in the Water Resources Research journal, by Tolson B. A., Asadzadeh M., Maier H. R., and Zecchin A., December 2009, volume 45, W12416 in 15 pages. References are unified at the end of thesis.

Summary

The Dynamically Dimensioned Search (DDS) continuous global optimization algorithm by Tolson and Shoemaker (2007) is modified to solve discrete, single-objective, constrained Water Distribution System design problems. The new global optimization algorithm is called Hybrid Discrete Dynamically Dimensioned Search (HD-DDS) and combines two local search heuristics with a discrete DDS search strategy adapted from the continuous DDS algorithm. The main advantage of the HD-DDS algorithm compared with other heuristic global optimization algorithms, such as genetic and ant colony algorithms, is that its searching capability (i.e., the ability to find near globally optimal solutions) is as good, if not better, while being significantly more computationally efficient. The algorithm's computational efficiency is due to a number of factors, including the fact that it is not a population-based algorithm and only requires computationally expensive hydraulic simulations to be conducted for a fraction of the solutions evaluated. This chapter introduces and evaluates HD-DDS by comparing its performance with that of three other algorithms (specific versions of the Genetic Algorithm, Ant Colony Optimization, and Particle Swarm Optimization) on four water distribution system design case studies (21- to 454-dimensional optimization problems) on which these algorithms have been found to perform well.

This chapter is organized as follows. Section 2.1 is devoted to the review of related studies. The detailed information about the HD-DDS algorithm and the methodology to assess its performance are explained in section 2.2. Results and findings are reported in Section 2.3 followed by discussions in Section 2.4.

2.1 Introduction

The optimal design and rehabilitation of Water Distribution Systems (WDS) is an important research area as improved optimization methods save substantial infrastructure capital and operational costs. Historically, traditional optimization methods such as linear programming (Schaake and Lai 1969; Alperovits and Shamir 1977; Bhave and Sonak 1992), nonlinear two phase decomposition methods (Fujiwara and Khang 1990; Eiger et al. 1994) and nonlinear programming (Varma et al. 1997) have been applied to a continuous version of the WDS optimization problem. These methods are sophisticated in terms of their use of the fundamental hydraulic equations to recast the form of the optimization problem and to yield gradient and hessian Hessian information; however, their inability to restrict the search space to discrete pipe sizes is a significant practical limitation (Cunha and Sousa 1999). Reca and Martinez (2006) provide a more detailed review of classical optimization techniques as applied to WDS optimization problems.

The majority of current single-objective WDS optimization literature report using heuristic global optimization algorithms, including evolutionary algorithms, with great success. Genetic Algorithms (GAs) are probably the most well-known combinatorial evolutionary algorithm. Example applications of GAs to WDS optimization include Simpson et al. (1994), Savic and Walters (1997), Wu et al. (2001) and Tolson et al. (2004). Ant Colony Optimization (ACO) algorithms have also received attention in the recent WDS literature (Maier et al. 2003; Zecchin et al. 2006; Zecchin et al. 2007). Other new and promising approaches applied to WDS optimization include the Shuffled Frog Leaping Algorithm (SFLA) in Eusuff and Lansey (2003), the Harmony Search (HS) algorithm in Geem (2006), the Cross Entropy (CE) method in Perelman and Ostfeld (2007), Particle Swarm Optimization (PSO) in Suribabu and Neelakantan (2006) and another PSO variant in Montalvo et al. (2008). Additionally, more traditional heuristic search strategies that are not population-based, such as Simulated Annealing (Cunha and Sousa 1999) and Tabu Search (Lippai et al. 1999; Cunha and Ribeiro 2004) continue to be applied to some WDS optimization problems.

One of the major problems associated with the use of heuristic global optimization algorithms is that their performance, both in terms of computational efficiency and their ability to find near globally optimal solutions, can be affected significantly by the settings of a number of parameters that control their searching behavior (e.g., population size, probability of mutation, probability of crossover in the case of GAs), as well as penalty functions that are commonly used to account for system constraints.

The algorithm parameters for eight WDS optimization studies are summarized in Table 2-1. As can be seen, the reported number of total parameters (algorithm+penalty) in these eight algorithms ranges from three to eight, and seven of these eight studies report that a subset of algorithm parameters were either experimented with or modified for each of the case studies they were applied to. However, such algorithm parameter setting experiments can be extremely computationally expensive, as many of the parameters are dependent. In addition, this problem is exacerbated for problems with long WDS simulation times, particularly those requiring extended period simulations (e.g., when water quality considerations are important). In addition to requiring a large amount of computational time, there is no well accepted methodology for conducting these parameter setting experiments, which are therefore typically implemented on an ad-hoc basis. For example, the methodology used to determine the optimal parameter settings for many of the studies in Table 2-1 is not described.

Table 2-1 Characteristics of various optimization algorithms recently applied to water distribution system optimization.

Optimization Algorithm (Reference)	Number of Reported Algorithm + penalty function parameters[a]	Do algorithm parameter sets vary by case study?[b]	Are parameter values determined by case study specific experimentation or optimization?[b]
GA, GENOME (Reca and Martinez 2006)	8	NO	YES (for 1 of 3 cases)
PSO variant, (Montalvo et al. 2008)	8	YES	YES
SFLANET, (Eusuff and Lansey 2003)	6	YES	YES
HS, (Geem 2006)	5	YES	YES
MSATS, (Reca et al. 2008)	5	NO[c]	YES
PSO, (Suribabu and Neelakantan 2006)	5	YES	YES
MMAS ACO, (Zecchin et al. 2007)	4	YES	YES
CE, (Perelman and Ostfeld 2007)	3	YES	NO

a) Algorithm parameter counts do not include stopping criteria for algorithms (such as maximum number of generations or maximum number of solution evaluations) since these can be specified based on project specific timelines. All studies (excluding this chapter) report using a standard penalty function. All studies with a penalty function report either 1 or 2 penalty parameters except Perelman and Ostfeld (2007) who do not report on the form of the penalty function used.
b) Usually only a subset of reported algorithm parameters is varied or optimized by case study.
c) The authors report that parameters were only transferred from the first to the second case study because the extreme computational burden of the second case study precluded parameter experiments.

The issue of tuning parameters of heuristic optimization algorithms has been investigated in many studies. In relation to GAs, one approach is to self-adapt the parameters as part of the optimization procedure itself, e.g. Srinivasa et al. (2007). Alternative approaches are based on parameterless GA calibration methodologies, e.g. Lobo and Goldberg (2004) and Minsker (2005), and GA convergence behavior associated with genetic drift, e.g. Rogers and Prugel-Bennett (1999) and Gibbs et al. (2008). (Gibbs et al. 2010) compared the performance of the above approaches on the Cherry Hills-Brushy

Plains WDS optimization problem (Boccelli et al. 1998) and found that the approach based on genetic drift performed best overall. In relation to ACO, Zecchin et al. (2005) used a mixture of theoretical and sensitivity analyses to derive expressions for seven ACO parameters, which have been shown to perform well for a number of benchmark WDS optimization problems (Zecchin et al. 2007). One approach for eliminating penalty parameters or multipliers in single objective WDS optimization problems is to convert hydraulic constraints into objectives and solve a multi-objective optimization problem without hydraulic constraints e.g. Wu and Simpson (2002) and Farmani et al. (2005).

Despite these efforts, common practice in many current heuristic optimization methods for WDS design (see Table 2-1) still involves case study specific experimentation for tuning algorithm parameters. As discussed previously, such experimentation is undesirable, as it has the potential to increase computational effort significantly. In addition, it is unclear how much experimentation is sufficient and what impact limited experimentation has on algorithm performance for a particular problem. In order to address these issues, the Hybrid Discrete Dynamically Dimensioned Search (HD-DDS) algorithm is introduced in this chapter. Not only is the performance of HD-DDS reliant on only one parameter, but the value of this parameter does not have to be adjusted for different case studies. This is in contrast to existing heuristic optimization methods (Table 2-1). The HD-DDS algorithm builds on the Dynamically Dimensioned Search (DDS) algorithm introduced by Tolson and Shoemaker (2007) for continuous optimization problems and can be used for solving constrained, single-objective combinatorial WDS optimization.

2.2 Methodology

The specific WDS optimization problem addressed in this chapter is to determine pipe diameters from a discrete set of available options such that the total pipe material cost is minimized and pressure constraints are met across the network. All other network characteristics are known. This is a classical WDS design problem that the majority of the above WDS optimization references also solve.

2.2.1 Components of the Hybrid Discrete Dynamically Dimensioned Search Algorithm

The Hybrid Discrete Dynamically Dimensioned Search (HD-DDS) algorithm for WDS design optimization is described in the following sections. The HD-DDS algorithm utilizes global and local search strategies, as such a hybrid approach has been shown to be successful previously, e.g. Broad et al. (2006). Sections 2.2.1.1 through 2.2.1.4 describe how each strategy functions independently and then section 2.2.1.5 describes how they are combined to form the overall HD-DDS algorithm.

2.2.1.1 Discrete Dynamically Dimensioned Search Algorithm

The discrete Dynamically Dimensioned Search (discrete DDS) algorithm is the most important component of the HD-DDS algorithm and is a discrete adaptation of the DDS algorithm introduced by Tolson and Shoemaker (2007) for continuous optimization problems. DDS was designed as a simple and parsimonious algorithm with only one algorithm parameter to solve computationally intensive environmental simulation model automatic calibration problems. One DDS design goal was to have the algorithm automatically adjust and exhibit good performance within the user's timeframe for optimization (maximum number of solution evaluations) rather than require the user to modify and/or experiment with algorithm parameters to match their timeframe. A related DDS design goal was to eliminate the need for algorithm parameter adjustment when the case study or number of decision variables changes. While it is acknowledged that this is unlikely to result in the identification of globally optimal solutions, the DDS algorithm is simple to use in practice, while being able to consistently find near globally optimal solutions. In fact, Tolson and Shoemaker (2007) demonstrate better overall performance of DDS relative to other benchmark automatic calibration algorithms on optimization problems ranging from 6- to 30-dimensions with 1,000 to 10,000 solution evaluations per optimization trial while using the same DDS algorithm parameter value.

The parsimonious nature of DDS provides an attractive alternative for discrete WDS optimization problems given the very recent set of single objective WDS optimization algorithms reviewed in section 1.1 (see Table 2-1), all of which have from three to eight algorithm parameters of which a subset is usually modified and even optimized for different case studies. The discrete DDS algorithm is identical to the original DDS algorithm except for two modifications. The first of these enables the proposed algorithm to sample discrete valued candidate solutions, whereas the second is the addition of a new algorithm stopping criterion. The paragraphs below describe the DDS algorithm largely from Tolson and Shoemaker (2007) and are followed by a description of the two modifications that distinguish discrete DDS from DDS.

The DDS algorithm is a simple, stochastic, single-solution based, heuristic, global search algorithm that was developed for the purpose of finding a good approximation of the globally optimal solution within a specified maximum number of solution evaluations. The algorithm is designed to scale the search to a user-specified number of maximum solution evaluations. In short, DDS searches globally at the start of the search, transitioning to a more local search as the number of solution evaluations approaches the maximum allowable limit. The adjustment from global to local search is achieved by dynamically and probabilistically reducing the number of dimensions in the neighborhood (i.e., the

set of decision variables modified from their best value). Candidate solutions are sampled from the neighborhood by perturbing only the randomly selected decision variables from the current solution. These perturbation magnitudes are randomly sampled from a normal distribution with a mean of zero for the continuous version of DDS. These features of the DDS algorithm ensure that it is as simple and parsimonious as possible. DDS is a greedy type of algorithm since the current solution, also the best solution identified so far, is never updated with a solution that has an inferior value of the objective function. The algorithm is unique compared with current optimization algorithms because of the way the neighborhood is dynamically adjusted by changing the dimension of the search. The DDS perturbation variances remain constant and the number of decision variables perturbed from their current best value decreases as the number of solution evaluations approaches the maximum function evaluation limit. This key feature of DDS was motivated by experience with manual calibration of watershed models where early in the calibration exercise relatively poor solutions suggested the simultaneous modification of a number of decision variables but as the calibration results improved, it became necessary to only modify one or perhaps a few decision variables simultaneously so that the current gain in calibration results was not lost.

The discrete DDS algorithm pseudo code is given in appendix A-Figure A- 1, and the changes relative to the original DDS algorithm are in Steps 3 and 5. The only user-defined algorithm parameter is the scalar neighborhood size perturbation parameter (r), which defines the standard deviation of the random perturbation size as a fraction of the decision variable range. In the discrete DDS algorithm, the decision variables are integers from 1 to the number of discrete options (x_i^{max}) for each decision variable i. The objective function must translate or map these option numbers to discrete pipe diameters in the WDS optimization case. Consistent with the continuous version of DDS, a default value of the r parameter for discrete DDS is recommended as 0.2 (and used in this study). In the original DDS algorithm, the perturbation magnitude for each decision variable is sampled from a normal probability distribution. In Step 3 of discrete DDS (appendix A-Figure A- 1), the perturbation magnitude is randomly sampled from a discrete probability distribution that approximates a normal distribution (see example probability mass functions in Figure 2-1). In continuous and discrete DDS algorithms, $r = 0.2$ yields a sampling range that practically spans the normalized decision variable range for a current decision variable value halfway between the decision variable bounds. This sampling region size is designed to allow the algorithm to escape regions around poor local minima. The r parameter is a scaling factor assigning the same relative variation to each decision variable (relative to decision variable range).

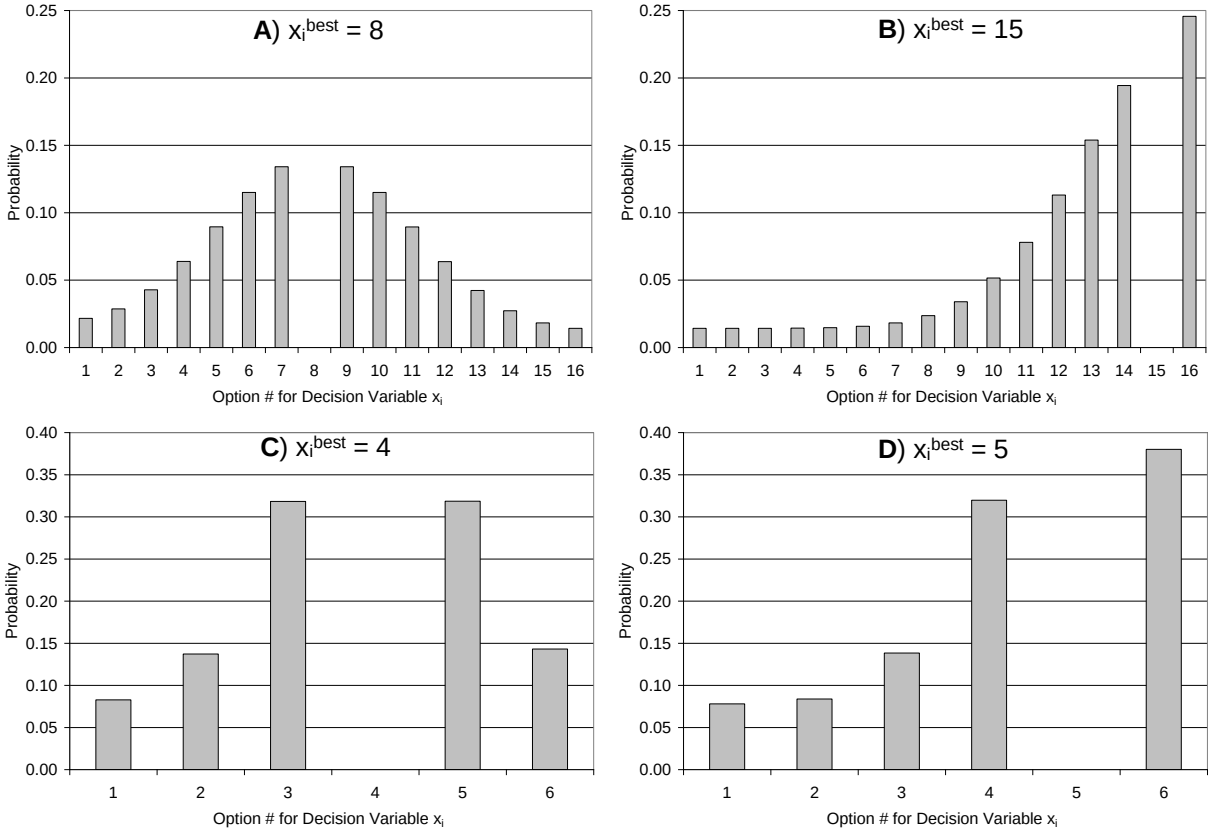

Figure 2-1 Example discrete DDS probability mass functions for candidate option numbers (x_i^{new}) for a single decision variable with 16 possible options (A and B) and 6 possible options (C and D) under various values for x_i^{best}. Default Discrete DDS parameter $r = 0.2$.

The one-dimensional decision variable perturbations in Step 3 of the DDS algorithm can generate new decision variable values outside of the decision variable bounds (e.g., 1 and x_i^{max}). In order to ensure that each one-dimensional perturbation results in a new decision variable that respects the bounds, the DDS and discrete DDS algorithms define reflecting boundaries (See Step 3 of appendix A-Figure A- 1). In the discrete DDS algorithm, the candidate decision variable values are initially treated as continuous random variables and the reflecting boundaries for the decision variables within the algorithm are defined to be 0.5 and $0.5 + x_i^{max}$. Once a candidate decision variable value within these reflecting boundaries is sampled, it is rounded to the nearest integer value to represent the discrete option number (e.g., $1, 2, 3, \ldots, x_i^{max}$). This reflecting boundary approach allows decision variable values to more easily approach their minimum or maximum values in comparison with a simple perturbation resampling approach (truncated normal distribution) for ensuring decision variable boundaries are respected. In the event that a perturbed decision variable has a candidate discrete option number that is the same as the option number of the current best solution (i.e., the

decision variable is not perturbed), a new discrete option number is sampled from a discrete uniform probability distribution. The resultant probability mass function for the decision variable option number is given for four examples in Figure 2-1 based on Step 3 of appendix A-Figure A- 1.

The maximum number of solution evaluations (NFE) is an algorithm input (like the initial solution) rather than an algorithm parameter because it should be set according to the problem specific available (or desired) computational time, see Gibbs et al. (2008). The value of NFE therefore depends on the time to compute the objective function and the available computational resources or time. Except for the most trivial objective functions, essentially 100% of discrete DDS execution time is associated with the solution evaluation. Recall that discrete DDS scales the search strategy from global in the initial stages of the search to more local in the final stages of the search regardless of whether NFE is 1,000 or 1,000,000 solution evaluations. In the absence of a specific initial solution, discrete DDS is initialized to the best of a small number (maximum of $0.005 \times \text{NFE}$ and 5) of uniformly sampled random solutions.

In the continuous DDS algorithm, depending on the problem dimension and NFE, a significant proportion of NFE in the latter half of the search would evaluate a change in only one decision variable relative to the current best solution. Given the general constrained WDS problem, it is known a priori that the feasible high quality solutions discrete DDS identifies by the later stages of the search will be very close to infeasibility. It is also known that the only way to minimize the objective function defined for discrete DDS is to select a smaller pipe diameter and thus reduce nodal pressures somewhere in the network. Being close to infeasibility means that the original DDS perturbation strategy that happens late in the search (changing only one decision variable) will quickly become futile and all possible single pipe diameter reductions will result in infeasible solutions. Therefore, a new primary stopping condition was added to discrete DDS in Step 5 to stop the search before all NFE is utilized (and thus save the remaining computation time for more productive search strategies) based on the probability of decision variable perturbation, p.

The number of dimensions selected for perturbation at each discrete DDS iteration, as determined by the first two bullets of Step 2 in appendix A-Figure A- 1, follows a binomial probability distribution parameterized by p. It is desired that more than one dimension is selected for perturbation at each iteration, which implies that an appropriate termination point is when the expected value of the binomial random variable is one. The corresponding probability value at this expected value is

$1/n$ (n is the number of decision variables) leading to a termination criterion of $p < 1/n$, or simply when the iteration count becomes $NFE^{(n-1)/n}$.

For the moderately sized WDS case studies solved (21-42 dimensions) with computational budgets ranging from 10,000 to 1,000,000 solution evaluations, this new stopping condition terminates discrete DDS after 52% to 80% of NFE are conducted. In networks with an order of magnitude of more decision variables (more than 200), for any computational budget fewer than 10,000,000 solution evaluations, discrete DDS will only stop after 90% or more of NFE. The complete HD-DDS algorithm and its other component search strategies (sections 2.2.1.3, 2.2.1.4, and 2.2.1.5) define how the remaining NFE is utilized after an initial optimization with discrete DDS terminates.

A preliminary version of discrete DDS was first defined in Tolson et al. (2008). However, discrete DDS in appendix A-Figure A- 1 of this study differs from Tolson et al. (2008) in two ways. First, the neighborhood definition is slightly modified to be more consistent with the original DDS algorithm in (Tolson and Shoemaker 2007). Another distinction is that Tolson et al. (2008) allowed the preliminary discrete DDS to utilize all NFE rather than stopping the algorithm early.

2.2.1.2 Constraint-Handling with Discrete DDS

Discrete DDS is not a population based algorithm and only one solution (the best found so far) influences the region of decision variable space that is sampled. As shown in Step 5 of appendix A-Figure A- 1, discrete DDS (like DDS) is not impacted by objective function scaling. Only the relative ranks of the best solution found so far and the candidate solution determine whether the best solution should be updated. Any update to the best solution moves it to a different point in decision space, around which future candidate solutions are cantered. This aspect of DDS and discrete DDS makes handling constraints very straightforward without the need for penalty function parameters. The constraint handling approach described below is equivalent to the method described in Deb (2000) for constraint handling in GAs using tournament selection (where objective function scaling also has no impact on the optimization). The key to the approach in Deb (2000) is that the objective function is defined such that any infeasible solution always has an objective function value that is worse than the objective function value of any feasible solution. In a WDS design problem where costs are to be minimized subject to minimum pressure requirements across the network, the worst cost feasible solution is known before any optimization as the cost of the solution specifying all pipes at their maximum diameter. The only other requirement for the constraint handling technique in Deb (2000) is to quantify the relative magnitude of constraint violations for infeasible solutions so that the

relative quality of two infeasible solutions can be compared. The steps to evaluate the objective function in discrete DDS are outlined in appendix A-Figure A- 2.

The evaluation of the objective function in appendix A- Figure A- 2 has a penalty term, but there is no need to scale it to a different magnitude with a case study specific penalty multiplier and/or exponent. Instead, the objective function definition implements the following logic:

- Between two infeasible solutions, the one with the least total pressure violation is always assigned a better objective function value.
- Between an infeasible and a feasible solution, the feasible one is always assigned a better objective function value.
- Between two feasible solutions, the one which is less costly is always assigned a better objective function value.

In addition to the benefit of requiring no case study specific penalty parameter values (which generally require experimentation to determine), the objective function definition in this study yields the significant benefit of eliminating the need to evaluate the hydraulics of a large number of solutions generated by discrete DDS.

2.2.1.3 One-pipe Change Local Search (L_1)

This local search technique starts at a feasible solution and cycles through all the possible ways to perturb the current best solution by reducing the diameter of one pipe at a time (thus saving costs). Each perturbed solution is evaluated for cost and feasibility. The L_1 search continues until it has evaluated a maximum number of solution evaluations or until it has enumerated all possible one-pipe changes without finding any improvement to the current best solution and therefore confirms that a local minimum has been located. The solution is a local minimum in that no better solution exists that differs from the current best solution in only one decision variable (pipe diameter). Pseudo code for implementing L_1 is shown in appendix A-Figure A- 3. Note that L_1 is implemented such that it is always evaluating one-pipe changes relative to the current best known solution. In other words, whenever L_1 finds a new best solution, the current best solution is updated and L_1 searches in the vicinity of the updated current best solution. Therefore, the order of enumeration (starting at pipe D instead of pipe 1) can change the final best solution returned by L_1.

2.2.1.4 Two-pipe Change Local Search (L_2)

L_2 starts at a feasible solution and cycles through all possible ways to perturb the current best solution by only two pipes where one pipe has the diameter increased and the other pipe has the diameter decreased. L_2 is only initiated at feasible solutions that are already confirmed to be a local minimum with respect to L_1. As Gessler (1985) suggested in his pipe design enumeration strategy, solutions that have a higher cost than the current best solution are not evaluated for their hydraulics in L_2. L_2 continues until it has evaluated a maximum number of solution evaluations, or it has enumerated all possible two-pipe changes without finding any improvement to the current best solution and confirms that a local minimum has been located. The solution is a local minimum in that no better solution exists that differs from the current best solution in only two decision variables (pipe diameters). Pseudo code of L_2 is shown in appendix A-Figure A- 4. L_2 enumerates all possible two-pipe changes from the current best solution to identify the best two-pipe change, so it is a different enumeration approach compared to L_1. Only after all two-pipe change possibilities are enumerated in L_2 is the current best solution updated and the enumeration process repeated. Hence, L_2 is defined so that different orders of enumeration do not affect the results if L_2 converges to a local minimum. In L_2, EPANET2 is run only if the enumerated solution is less expensive than the current best L_2 solution. This is quickly determined only by comparing the costs associated with the two decision variables (pipe diameters) being changed relative to the corresponding costs in the current best L_2 solution.

2.2.1.5 HD-DDS Algorithm Definition

The explicit design goals that motivated the development of the Hybrid Discrete Dynamically Dimensioned Search (HD-DDS) Algorithm were, in highest to lowest priority, to develop an algorithm that 1) reliably returned good quality approximations of the global optimum 2) was capable of solving WDN design problems to optimality and 3) could achieve these goals in a way that was computationally efficient. In addition, the algorithm was designed to be parsimonious in the number of algorithm parameters such that achieving the above goals did not necessarily require parameter tuning experiments for each case study. The definition of a good quality approximation is clearly specific to each case study and even the computational budget used to solve the problem, but in general, is taken to be a solution whose cost is small enough to satisfy the designer such that further optimization would be deemed unnecessary. A design characteristic that the algorithm inherits from the continuous DDS algorithm is the ability to effectively scale to various computational budgets such that it is capable of generating good results without algorithm parameter tuning.

The HD-DDS algorithm component is designed to overcome the shortcomings of other individual component search strategies. First and foremost, as described in section 2.2.1.1, the discrete DDS search is predictably unproductive with the original DDS algorithm stopping criterion, and thus HD-DDS invokes alternative search strategies after discrete DDS reaches this unproductive stage (when $p < 1/n$). Unlike other hybridized algorithms (such as a ACO + a local search), the point at which to terminate a global search in favour of a new search strategy in HD-DDS is known a priori and thus does not require the definition of various algorithm specific convergence measures. The component search combination strategy defining HD-DDS is given in appendix A-Figure A- 5. Initially, HD-DDS requires all problem inputs to be specified. This includes the user-defined computational budget, $\mathrm{NFE}_{\mathrm{total}}$, which can be expended in solving the problem. $\mathrm{NFE}_{\mathrm{total}}$ is a quantity that users should determine based on their problem specific available computation time and the average computation time required for a solution evaluation that evaluates both cost and hydraulics. As discussed later in section 2.4, because HD-DDS can skip the hydraulics evaluation for a large number of candidate solutions, HD-DDS will terminate much quicker than what would be estimated in this way. The only algorithm parameter in HD-DDS is the discrete DDS parameter r, which has a default of 0.2 that is utilized for all HD-DDS analyses presented in this chapter.

The first search strategy implemented in Step 1 of appendix A-Figure A- 5 is a global search with discrete DDS for a maximum of NFE solution evaluations, where NFE in HD-DDS is a variable tracking the remaining available number of solution evaluations. The second search strategy in Step 2 of HD-DDS is defined to be the L_1 local search (one-pipe change), which polishes the discrete DDS solution to a local minimum with respect to L_1. Results show this enumerative search is typically very quick in polishing a discrete DDS solution. L_1 often requires less than n (decision variable count) evaluations of the simulation model EPANET2 (USEPA, 2002, version 2.00.12), simulation model for hydraulics (pipe flows and nodal pressures) and/or water quality in pressurized pipe networks.

Step 3 in HD-DDS (appendix A-Figure A- 5) is defined as a second independent discrete DDS search from a new initial solution followed by another L_1 search to polish the second discrete DDS solution. This second discrete DDS search has a smaller NFE than the first and therefore will utilize fewer solution evaluations. At the end of Step 3, a second potentially new solution that is a local minimum with respect to L_1 has been located, which is referred to as x_B^{best}. Step 4 in HD-DDS (appendix A-Figure A- 5) implements the first L_2 local search to polish the best of the two HD-DDS solutions found so far. With very large networks, L_2 can be extremely slow to converge and confirm

that a local minimum has been reached. However, as results will show, with a reasonably good quality solution, L_2 typically converges in a few thousand solution evaluations in HD-DDS for case studies with up to 42 decision variables. Step 5 in HD-DDS implements the second L_2 local search to polish the other HD-DDS solution that is confirmed to be a local minimum with respect to L_1.

HD-DDS can terminate when either the total computational budget (NFE_{total}) is exceeded or when all five steps are completed in fewer than NFE_{total} solution evaluations. For a fixed and reasonably large budget (e.g., 100,000), as the network size decreases (number of decision variables decrease), the likelihood that HD-DDS will be able to complete all five steps increases. However, as network size increases, it becomes more and more likely that HD-DDS will terminate without a second discrete DDS search. In other words, HD-DDS reduces to only Steps 1 and 2 when the number of decision variables becomes extremely large because the user defined computation limit will be exceeded before Step 2 finishes. This behavior is based on the assumption that for a fixed and practical budget, the globally optimal solution to an extremely large problem is virtually impossible to find, and the best approach would be to conduct the longest possible global search.

HD-DDS was designed to always perform a second global search (Step 3) before the first L_2 local search polishing step. The main purpose of a second global search is to act as a safety mechanism to guard against a very poor solution from the first discrete DDS trial. To maximize the chances the safety mechanism works, this second global search requires the largest possible computation budget and therefore must be executed prior to L_2 (which may take an indeterminate amount of time). Furthermore, there is a general belief that the best chance to significantly improve upon a particularly poor solution is to conduct a new global search rather than polish the particularly poor solution. This design choice has no impact on final solution quality if the computation budget is large enough and/or the problem dimension is small enough to allow all five steps in HD-DDS terminate completely in fewer than NFE_{total} solution evaluations. There are some situations (particularly with problems having hundreds of decision variables) where the choice to conduct a second global search (followed by L_1) uses all or most of the remaining computational budget and thus precludes or reduces computation time available for the L_2 local search. In these situations with such large networks, the L_2 local search can require an incredible number of solution evaluations to enumerate all possible two-pipe changes such that L_2 terminates before it confirms the solution is a local minimum. Despite this, the optimal design choice between L_2 or second global search is certainly case study dependent. Results will show that this design choice has very little impact on HD-DDS results.

2.2.2 Benchmark Optimization Algorithms

The main focus of this chapter is to introduce the HD-DDS algorithm and its performance comparison with alternative algorithms. The design case studies and benchmark algorithms were selected from the literature based on whether the WDS case study could be replicated exactly, and whether the algorithm results were presented for multiple optimization trials. Exact replication of a WDS case study meant that comparative algorithm results must have been generated with EPANET2 using metric units. Most WDS optimization algorithms are stochastic in that they can and often do generate a different final solution for a different random seed and/or different initial solution. So, objectively assessing relative algorithm performance requires multiple independent optimization trials.

Three studies that meet the above criteria and thus provide the majority of the comparative algorithm results used in this study are Reca and Martinez (2006), Zecchin et al. (2007), and Montalvo et al. (2008). Zecchin et al. (2007) test the performance of five different Ant Colony Optimization (ACO) algorithms on some standard WDS case studies. They report that the MMAS ACO algorithm outperforms all other algorithms applied to their case studies in the literature. Reca and Martinez (2006) apply a GA called GENOME to a much larger scale WDS optimization problem. Montalvo et al. (2008) introduce a Particle Swarm Optimization (PSO) algorithm variant to some standard WDS case studies and show that their PSO variant outperforms a standard discrete PSO algorithm. Details of the WDS case studies selected for HD-DDS testing from the above studies are given in the following section.

2.2.3 Benchmark WDS Design Studies

The first three example WDS design problems in this study, the New York Tunnels, Doubled New York Tunnels, and Hanoi problems, are equivalent to those used in Zecchin et al. (2007). For complete details of these three case studies, readers are referred to Zecchin et al. (2005). The GoYang WDS problem was introduced by Kim et al. (1994) and is also utilized as the fourth case study. The fifth and final case study is the Balerma WDS problem, which was first proposed by Reca and Martinez (2006). The next few paragraphs provide an overview of each WDS case study, and readers should consult the references given for case study details (e.g., discrete pipe diameter options, pipe length, nodal head requirements etc.). Note that all EPANET2 units are metric in all the case studies.

The New York Tunnels Problem (NYTP), originally considered in Schaake and Lai (1969), involves the rehabilitation of an existing WDS by specifying design options for the 21 tunnels in the system. There are 16 design options for each tunnel (parallelization with one of 15 tunnel sizes or a

do-nothing option), thus defining a search space size of 16^{21} (approximately 1.93×10^{25}). The Doubled New York Tunnels problem (NYTP2) was originally proposed by Zecchin et al. (2005). This network has 42 pipes to be sized, and each has 16 options. This defines a search space of 16^{42} (approximately 3.74×10^{50}) for NYTP2.

The Hanoi Problem (HP) was first reported on in Fujiwara and Khang (1990) and has a larger search space size than NYTP, 2.87×10^{26} solutions based on 32 pipes to be sized from 6 options. HP problem details are available in Wu et al. (2001). The GoYang problem (GYP) represents a South Korean WDS and was first proposed by Kim et al. (1994) and has 30 pipes to be sized with 8 diameter options per pipe to define a search space size of 8^{30} (approximately 1.24×10^{27}).

The Balerma irrigation network problem (BP) from Reca and Martinez (2006) is a large and complex network that has 443 demand nodes, 454 pipes, 8 loops, and 4 reservoirs. Each of the 454 pipes to be sized has ten possible diameters and thus defines a search space size of 10^{454}.

Based on a review of the current literature, the best current known feasible solutions (using EPANET2) are \$38.638 million for the NYTP (Maier et al. 2003), \$77.276 million for the NYTP2 (Zecchin et al. 2005), \$6.081 million for the HP (Perelman and Ostfeld 2007), and €2.302 million for the BP (Reca and Martinez 2006). Currently, published optimization algorithm results for GYP with EPANET2 as the hydraulic solver are unavailable and thus HD-DDS is not compared to any other algorithm for the GYP.

2.2.4 Optimization Model Formulation for HD-DDS

The HD-DDS algorithm solves the optimization problem in equation (2-1) for each WDS case study. In this problem, $f(\boldsymbol{x})$ is the objective function for decision vector $\boldsymbol{x} = (x_1, \ldots, x_n)$ as determined by the procedure in appendix A-Figure A- 2, and x_i is the integer-valued pipe diameter option number for pipe i and is between option 1 (the smallest diameter) and the maximum diameter option, x_i^{max}, for all n pipes in the network to be sized. The minimum required nodal pressure heads (e.g., h_j^{min} for node j in appendix A-Figure A- 2) and nonlinear cost equations vary with each case study and must be specified (readers can find these case study specific values in the case study references noted in section 2.2.3).

$$\begin{aligned} &\min_{\boldsymbol{x}} f(\boldsymbol{x}) \\ &st.\ x_i \in \{1,2,\ldots,x_i^{max}\}, \qquad \forall i = 1,\ldots,n \end{aligned} \tag{2-1}$$

2.2.5 Outline of Algorithm Comparisons

The performance of different algorithms is generally compared for a similar computational budget as measured by the total number of solution evaluations utilized in the optimization run. The computational budget reported for all algorithms is measured as the budget required for the algorithms to terminate. Although this is not the measure that is reported in Zecchin et al. (2007) for the MMAS ACO algorithm, the measure used in this study reflects the fact that in any new problem, with an unknown best solution, the user will not generally have any knowledge or reason to stop the algorithm before it terminates normally.

Due to the stochastic nature of most heuristic global optimization algorithms, their relative performance must be assessed over multiple independent optimization trials, each initialized to independent populations or solutions. Previously published algorithm results are compared to HD-DDS results using 10 to 50 optimization trials. Optimization algorithm comparisons cover 21-, 30-, 32-, 42- and 454-dimensional (i.e., the number of decision variables or pipes to be sized) problems, and the size of the search spaces ranges from 1.93×10^{25} to 1.0×10^{454} possible solutions. The maximum number of solution evaluations, $\mathrm{NFE}_{\mathrm{total}}$, per HD-DDS optimization trial varies from 10,000 to 10,000,000. Given that average algorithm performance across multiple independent optimization trials does not provide a complete picture of results, the distribution or range of the HD-DDS and benchmark algorithm solutions are also assessed.

The initial solution for HD-DDS is generated as described in section 2.2.1.1, and the neighborhood size parameter, r, is set to the default value of 0.2 for all HD-DDS trials (no parameter setting experimentation was performed on r). In some of the case studies, L_1 or L_2 local searches initiated in HD-DDS before the remaining available computational budget is exceeded were typically allowed to terminate at a local minimum even if that meant exceeding $\mathrm{NFE}_{\mathrm{total}}$ by a few thousand solution evaluations on average. In contrast, comparative MMAS ACO algorithm results from Zecchin et al. (2007) are based on MMAS ACO algorithm parameters that were optimized independently for each case study using millions of EPANET2 simulations. Similarly, the comparative results for the GENOME GA and PSO variant algorithms were also derived from some experimental optimization runs to identify good algorithm parameters.

2.3 Results

Results are presented in two sub-sections. In Section 2.3.1, the importance and impact of each step of the HD-DDS algorithm are assessed in the results of four case studies. Comparative algorithm performance results are presented in section 2.3.2.

2.3.1 HD-DDS Component Performance Assessment

For all HD-DDS optimization trials, at the end of each HD-DDS step in appendix A-Figure A- 5, the best HD-DDS solution found so far and the corresponding number of solution evaluations are recorded. This information is shown in Figure 2-2 for the NYTP, NYTP2, HP, and GYP benchmarks. In addition to presenting results from all individual optimization trials, the average cost of the best solution found so far is shown in Figure 2-2 where objective function values are plotted against the average number of solution evaluations for clarity because for Steps 2, 3, 4 and 5, the number of utilized solution evaluations varies between optimization trials. The difference between each step on the x-axis in Figure 2-2 shows the average computational requirements of each Step in HD-DDS.

For each case study in Figure 2-2, Step 2 in HD-DDS (L_1 one-pipe change local search) requires a negligible number (~10) of solution evaluations to determine that discrete DDS terminated at a local minimum with respect to L_1 In fact, in all HD-DDS optimization trials, L_1 never improved upon a discrete DDS solution, indicating that discrete DDS always terminated at a local minimum with respect to L_1 With a much smaller computational budget and/or much larger network (e.g., see BP results for $\mathrm{NFE}_{\mathrm{total}} \leq 10{,}000$ in section 2.3.2.1), discrete DDS will not always terminate at a local minimum with respect to L_1, and thus L_1 can improve upon discrete DDS solutions.

The computational budget for the second discrete DDS search (Step 3) is significantly smaller than that for the first discrete DDS search for all four case study results in Figure 2-2 (21% - 41% of the budget of the first discrete DDS search). Despite the substantial decrease in computational budget, results of Steps 2 and 3 in Figure 2-2 show that this second, quicker discrete DDS search does sometimes improve upon the first DDS search result (for all four case studies, Step 3 improves best solution in 20% to 32% of optimization trials). Most notable in Figure 2-2 is that the second DDS search operates to eliminate the worst cost solutions from Step 2 in all four case studies.

Results in Figure 2-2, at each Step of HD-DDS, also present a count of the number of optimization trials that have located the best known solution for each case study. Discrete DDS (Steps 1 and 3 of HD-DDS) generally does not find the best known solution (except for a small number of trials in

NYTP and GYP). Results for Steps 4 and 5 clearly show that the L_2 (two-pipe change) is capable of polishing discrete DDS (Steps 1 and 2) solutions and returning the best known solution with 85% frequency for NYTP and NYTP2. The average computational burden and percentage of optimization trials with improved solutions due to L_2 in Step 4 (which terminates with a local minimum with respect to L_2) is 1421, 9031, 3250, and 586 solution evaluations and 90%, 56%, 100%, and 66% for NYTP, NYTP2, HP, and GYP, respectively. The average computational burden for Step 5 (second L_2 which terminates with a local minimum with respect to L_2) increases to 2302, 15403, 4487, and 649 solution evaluations on average for NYTP, NYTP2, HP, and GYP, respectively, because the second L_2 is typically starting from a lower quality solution. Results for NYTP and NYTP2 in Figure 2-2 demonstrate that, Step 5 can improve results indicating that a second L_2 starting from a lower quality initial solution can be fruitful. The percentage of optimization trials with improved solutions due to L_2 in Step 5 is 32%, 25%, 14%, and 6% for NYTP, NYTP2, HP, and GYP, respectively.

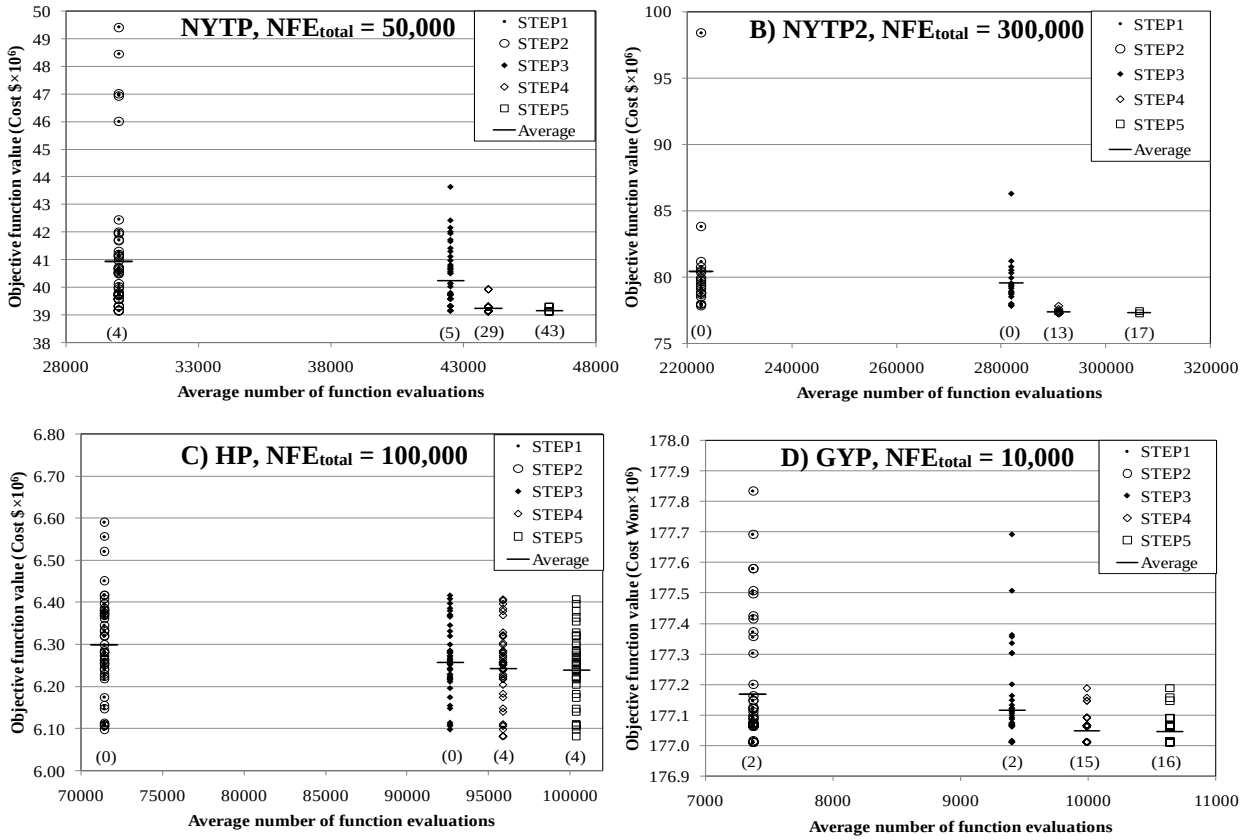

Figure 2-2 Progress as of the end of each step of the HD-DDS algorithm (see appendix A-Figure A-5) versus average number of solution evaluations for various WDS case studies and corresponding total computational budget input to HD-DDS. 50 optimization trials are shown for each case study except for NYTP2 with 20 optimization trials. The numbers in brackets count the number of trials where the best solution so far is equal to the best known solution.

2.3.2 HD-DDS Performance Relative to Benchmark Algorithms

The first comparative results are for HD-DDS and the MMAS ACO algorithm results in Zecchin et al. (2007) for the NYTP, HP and NYTP2. Figure 2-3 shows the empirical cumulative distribution function (CDF) for HD-DDS and MMAS ACO of the final best objective function values (costs) from all optimization trials. HD-DDS results stochastically dominate MMAS ACO results in all three case studies because for any desirable cost target identified by a decision-maker, HD-DDS always has an equal or higher probability of satisfying the cost target than MMAS ACO. The near vertical lines for NYTP and NYTP2 indicate that HD-DDS yields the best known solution with a high reliability (more than 80%). For HP, eight HD-DDS optimization trials were better than the best MMAS ACO solution. HD-DDS avoided the worst solutions identified by MMAS ACO. The superior performance of HD-DDS over MMAS ACO is even more noteworthy considering there was no algorithm parameter experimentation with HD-DDS, and there were extensive case study specific parameter setting experiments conducted for MMAS ACO, as discussed previously.

2.3.2.1 Large Scale WDS: Balerma Case Study

The Balerma network has 454 pipes to be sized ($n = 454$), and as a result, Step 1 in HD-DDS (see appendix A-Figure A- 5) uses more than 96% of any of the computational budgets specified in this study ($1{,}000 \leq \mathrm{NFE}_{\mathrm{total}} \leq 10{,}000{,}000$). Note that HD-DDS Step 1 (first discrete DDS search) terminates only when p, as calculated in Step 2 of discrete DDS (see appendix A-Figure A- 1), is less than $1/n$. The remaining 4% or less of the computational budget is utilized by Step 2 (L_1 search) and then any remaining budget is dedicated to Step 3 of HD-DDS (second global search followed by another typically incomplete L_1 search).

Figure 2-4 shows all of the HD-DDS results generated for Balerma as the number of solutions are increased from 1000 to 10,000,000. All HD-DDS results are based on 10 optimization trials. As expected, Figure 2-4 shows that HD-DDS performance improves with a larger computational budget. It is important to recall that HD-DDS scales the search to the user input $\mathrm{NFE}_{\mathrm{total}}$, and thus separate independent optimization trials are used to generate results for different computational budgets. For example, the best solution after 10,000 solution evaluations in HD-DDS with $\mathrm{NFE}_{\mathrm{total}} = 100{,}000$ will not be equal to the final best solution in HD-DDS with $\mathrm{NFE}_{\mathrm{total}} = 10{,}000$.

Comparative algorithm performance is also shown in Figure 2-4 using results for MSATS and GENOME GA reported in Reca et al. (2008) and GENOME GA results reported in Reca and Martinez (2006). For the same computational budget ($\mathrm{NFE}_{\mathrm{total}} = 10{,}000{,}000$), HD-DDS clearly

outperforms the GENOME GA, as the worst HD-DDS solution is better than the best GENOME GA solution by nearly €200,000. Even with 1 percent of this computational budget, HD-DDS still outperforms the GENOME GA as the worst HD-DDS result after 100,000 solution evaluations costs nearly €100,000 less than the best GENOME GA result obtained by 10,000,000 solution evaluations.

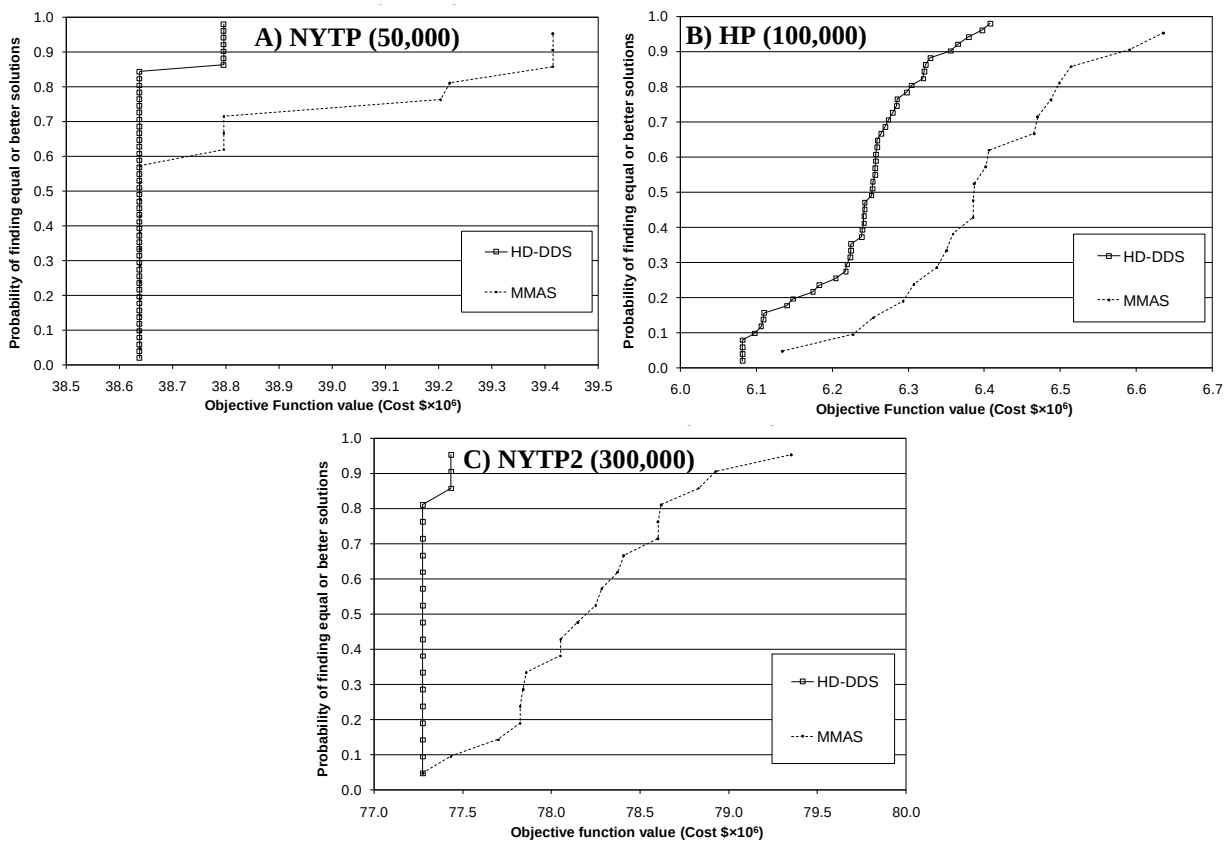

Figure 2-3 Empirical CDF of best solutions from HD-DDS and MMAS ACO algorithm for the A) NYTP, B) HP and C) NYTP2 for approximately the same number of solution evaluations (see the figure for NFE_{total} in brackets). MMAS results are for 20 optimization trials. HD-DDS results for NYTP and HP are for 50 optimization trials while NYTP2 are for 20 optimization trials.

In the interest of determining the best possible solution, the best solution of HD-DDS with $NFE_{total} = 10{,}000{,}000$ (cost of €1,956,226) was passed onto the L_2 local search to polish the solution until it was a confirmed local minimum with respect to L_2. This L_2 polishing was manually terminated after an additional 20 million EPANET2 evaluations with a local minimum having a cost of €1,940,923 (also shown in Figure 2-4).

The HD-DDS results for $NFE_{total} = 10{,}000$ and $NFE_{total} = 1{,}000$ in Figure 2-4 demonstrate that even with a severely restricted computational budget, HD-DDS can generate reasonable solutions, all of which are feasible. Even the median performed HD-DDS trial is much better than GENOME GA

and MSATS results from Reca et al. (2008) after 45,000 solution evaluations. Note that, HD-DDS results for $NFE_{total} = 1{,}000$ utilized an average of 2,900 solution evaluations rather than 1000 because the L_1 one-pipe change search was only terminated when it returned a local solution with respect to L_1 (which required approximately 1900 additional solution evaluations). In fact, the L_1 search drastically improved the average discrete DDS solution quality from 5.021 to 3.080 million Euros. Application of only L_1 without the preliminary discrete DDS search was ineffective in large part because without the discrete DDS solution, L_1 could not quickly identify a feasible solution.

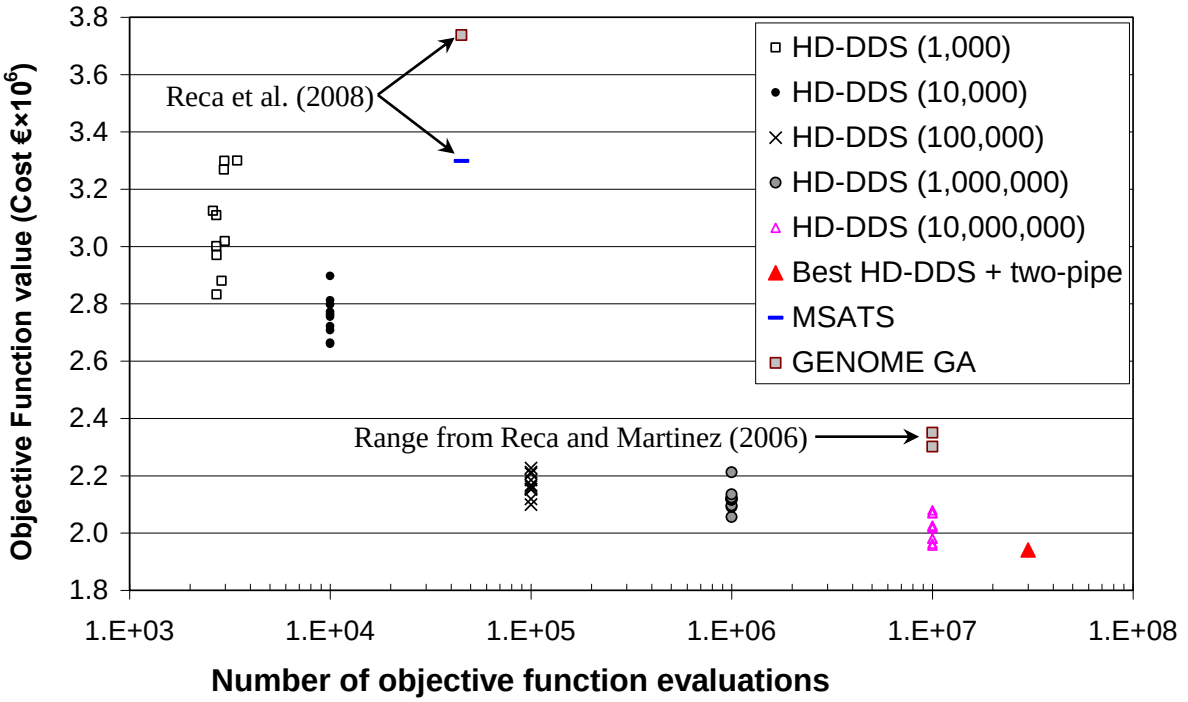

Figure 2-4 HD-DDS performance with different computational budgets (NFE_{total}) compared to other algorithm performance on Balerma network. HD-DDS results show all 10 trials. Other algorithm results are for a single trial or show the range of results from multiple trials.

Some final tests were performed to compare HD-DDS performance if L_2 was performed instead of the second global search. Performing L_2 yields reliable but very small improvements (0.1%-0.2% on average) in the best cost solution after Step 2 of HD-DDS in appendix A-Figure A- 5. However, performing the second global search is capable of yielding significantly larger best cost improvements (more than 2%) much less frequently. Therefore, as designed the algorithm forgoes very small improvements under L_2 for significantly higher but less frequent improvements capable with second global search.

2.3.2.2 HD-DDS Performance Comparison Summary

Table 2-2 summarizes and compares the results of HD-DDS and other algorithms previously noted in sections 2.3.2. The main difference in addition to compressing all previous graphical results into one table is that the solution quality is also measured with respect to the percent deviation from the best known solution. Table 2-2 also includes new results for a few other algorithms and HD-DDS computational budgets. Results for other algorithms in Table 2-2 are only included where it was possible to confirm that the algorithms were applied to the exact same optimization formulation (e.g., EPANET2 was used with metric units).

Table 2-2 Summary of HD-DDS and other algorithm performance for the five WDS case studies investigated in this study.

Algorithms [a]	WDS Case Study	Solution Evaluations ×Number of optimization trials [b]	% trials found best known solution	Best Cost [c]		
				Minimum	Median	Maximum
–MMAS ACO	NYTP	50,000 × 20	60	38.638 (0.0)	38.638 (0.0)	39.415 (2.0)
HD-DDS		50,000 × 50	86	38.638 (0.0)	38.638 (0.0)	38.769 (0.3)
PSO variant		80,000 × 2000	30	38.638 (0.0)	38.83 (0.5)	47.0[d] (21.6)
–MMAS ACO	HP	100,000 × 20	0	6.134 (0.9)	6.386 (5.0)	6.635 (9.1)
HD-DDS		100,000 × 50	8	6.081 (0.0)	6.252 (2.8)	6.408 (5.4)
PSO variant		80,000 × 2000	5	6.081 (0.0)	6.31 (3.8)	6.55[d] (7.7)
CE		97,000 × 1	-	6.081 (0.0)	-	-
GENOME GA		150,000 × 10	10	6.081 (0.0)	6.248 (2.7)	6.450 (6.1)
HD-DDS		150,000 × 50	2	6.081 (0.0)	6.260 (2.9)	6.393 (5.1)
–MMAS ACO	NYTP2	300,000 × 20	5	77.275 (0.0)	78.199 (1.2)	79.353 (2.7)
HD-DDS		300,000 × 20	85	77.275 (0.0)	77.275 (0.0)	77.434 (0.2)
GENOME GA	BP[e]	10,000,000 × 10	0	2.302 (18.7)	2.334 (20.3)	2.35 (21.1)
HD-DDS		100,000 × 10	0	2.099 (8.2)	2.165 (11.6)	2.212 (14.0)
MSATS[f]		45,000 × 1	-	3.298 (69.9)	-	-
HD-DDS		10,000 × 10	0	2.660 (37.0)	2.759 (42.2)	2.897 (49.3)

a) See Table 2-1 for references

b) Unlike all other algorithms, the majority of HD-DDS solution evaluations do not require evaluating the hydraulics with EPANET2. See discussion in section 2.4.

c) Monetary units times 10^6 and percent deviation from best known solution (in parentheses).

d) Based on 100 optimization trials, Figure 3 for HP and Figure 5 for NYTP in Montalvo et al. (2008).

e) The best known solution to Balerma based on 30 million solution evaluations is 1.9409 million Euro (see section 2.3.2.1).

f) MSATS (mixed simulated annealing tabu search) is the best of 4 metaheuristics on this problem from Reca et al. (2008).

Results in Table 2-2 pair each HD-DDS result with another comparative algorithm result, where HD-DDS is typically applied with a similar number of solution evaluations, and demonstrate the excellent overall performance of HD-DDS. In all algorithm comparisons, the median best costs found by HD-DDS are always equal to or lower than the costs obtained using the other algorithms. Importantly, the maximum cost solutions found by HD-DDS are always better than those found by

the comparative algorithms for all case studies (HD-DDS is 1.2% to 21.3% closer to the best known solution). For example, the worst HD-DDS solution for NYTP2 is nearly 2 million dollars less than the worst MMAS ACO solution and the worst HD-DDS solution for NYTP is just over 8 million dollars less than the worst PSO variant solution. The minimum best costs found by HD-DDS are equal to or lower than the costs obtained using all other algorithms in all comparisons. With the exception of the 150,000 × 10 results for HP, Table 2-2 shows HD-DDS always returns the best known solution with a higher frequency than other algorithms. For example, HD-DDS finds the best known NYTP solution in 86% of optimization trials compared to 30% for the PSO variant despite HD-DDS using 30,000 fewer solution evaluations.

The best HD-DDS solutions for NYTP and NYTP2 are the same as the ones reported in Zecchin et al. (2007). The best HD-DDS solution for HP is the same as one reported in Perelman and Ostfeld (2007). The best HD-DDS solution for BP is a new best known solution.

2.4 Discussion and Conclusions

The HD-DDS algorithm was developed and successfully applied to a range for WDS case studies in this chapter. The default value of the neighborhood perturbation size parameter, r, of 0.2 produced excellent results compared to all other algorithms for the WDS case studies reported on in section 2.3. These good results cover 21- to 454-dimensional problems and are based on computational budgets ranging from 1,000 to 10,000,000 solution evaluations. Therefore, the default value for r appears robust and is suggested for future HD-DDS applications. Results also showed that each search strategy (or step) in HD-DDS played an important role in at least one case study.

HD-DDS has a very large computational efficiency advantage over most other WDS optimization algorithms that is not obvious from results reported in section 2.3. For consistency with previous studies, the computational budget of HD-DDS was defined with respect to the total number of solution evaluations. However, the solution evaluation strategy with constraint handling (see appendix A-Figure A- 2) and the computationally efficient implementation of L_2 enable HD-DDS to evaluate the solution quality to a sufficient level without simulating the hydraulics of the solution (e.g., an EPANET2 simulation). Therefore, even though, for example, HD-DDS utilized approximately 46,000 solution evaluations to optimize the NYTP with $\mathrm{NFE}_{\mathrm{total}} = 50{,}000$, EPANET2 simulations were not required for approximately 33,000 or 72% of all HD-DDS solutions evaluated. Based on actual run times for HD-DDS, results show that the HD-DDS algorithm as implemented runs 50% faster than it would have if EPANET2 was used to evaluate all 46,000

solutions identified in HD-DDS. Balerma results for $NFE_{total} = 100{,}000$ were similar in that EPANET2 simulations were not required for approximately 71% of all HD-DDS solutions evaluated. Based on HD-DDS run times for Balerma, results show that the HD-DDS algorithm as implemented runs 67% faster than it would have if EPANET2 was used to evaluate all solutions identified in HD-DDS. The relative computational efficiency gain of HD-DDS over other optimization algorithms that must evaluate hydraulics of every candidate solution becomes larger as the computational demand of the hydraulics simulation increases. If the EPANET2 evaluations accounted for nearly 100% of HD-DDS computation time then HD-DDS could be more than 70% faster than other WDS optimization algorithms like MMAS ACO (Zecchin et al. 2007) and the roulette wheel selection based GENOME GA (Reca and Martinez 2006) in evaluating the same number of objective functions.

This chapter represents the paper in which the constraint handling strategy in Deb (2000) was applied for the first time to a WDS optimization application, and overall, the excellent results of HD-DDS suggest the strategy works very well. All performed discrete DDS optimization trials returned a final solution that was feasible, even in the HP for which multiple studies report difficulty in locating any feasible solution due to its small feasible search space (Eusuff and Lansey 2003; Zecchin et al. 2005; Zecchin et al. 2007). In order to further demonstrate the excellent performance of HD-DDS with this constraint handling approach for HP, 10 independent trials initialized at the most infeasible solution (all pipes at their minimum diameter) with $NFE_{total} = 10{,}000$ were conducted. Each of these runs were terminated at exactly 10,000 solution evaluations, and all returned feasible solutions with an average cost of \$6.299 million and a worst cost of \$6.375 million.

For a fixed computational budget and specific case study, it is possible that there are better ways to combine discrete DDS, one-pipe and two-pipe change algorithms than the HD-DDS algorithm presented. In other words, it is not claimed that HD-DDS is the optimal way to configure these three search strategies across all case studies and all computational budgets. The optimal configuration of these three search strategies is almost certainly case study and computational budget specific. Instead, a very robust and parsimonious way to combine these strategies is demonstrated. HD-DDS performs two independent global searches before spending an unknown amount of effort to polish the best solution with the two-pipe change local search. Results show that relative to available benchmark algorithm performance, for similar computational budgets, HD-DDS performs equivalent to or better than all algorithms in the comparison in the five case studies considered here. Therefore, experimenting with alternative ways to combine the search strategies in HD-DDS is unnecessary. Instead of experimenting with HD-DDS component configurations to improve upon HD-DDS

performance for a new WDS problem, it is recommended that HD-DDS users utilize the available time they have for performing multiple independent HD-DDS optimization trials or implement alternative local search strategies.

In practice, when users apply the HD-DDS algorithm as suggested in this chapter, results have shown they will have a solution much more quickly (by 50-70%) than the worst case they plan for (e.g., under the worst case assumption that all solutions will have their hydraulics evaluated). In deciding how to utilize any available remaining computational budget after their first HD-DDS trial terminates, it is suggested that if users are satisfied with the current HD-DDS solution quality (cost), and HD-DDS terminated before the L_2 local searches both converge, they give HD-DDS the extra time to polish the available solutions with L_2. In any other cases (users are dissatisfied with best cost returned or HD-DDS terminates because all five algorithm steps were completed), users can perform a second HD-DDS optimization trial with the remaining budget.

It is important to note that the benchmark WDS case studies solved here (which minimize cost subject to some design constraints) are gross simplifications of real-world WDS design problems. Walski (2001) discusses why real world problems need to be solved by minimizing net benefits and thus considering multiple different objectives. Future work is to extend the HD-DDS methodology to multi-objective optimization benchmarks that are better representations of real-world WDS design problems. Now that HD-DDS has been shown to be effective relative to benchmark single-objective optimization algorithms, future studies should focus on the application of HD-DDS to real-world WDS design problems as formulated by practicing engineers designing the system. In such a study, it is expected that the one-pipe and/or especially the two-pipe change local searches could be modified or replaced with alternate and perhaps case study specific local search strategies that replicate the logic practicing engineers employ when evaluating alternative system designs by trial and error. For example, the two-pipe change local search could be replaced with the grouping method search strategy that Gessler (1985) suggested for large real-world WDS design problems. A more promising and incredibly efficient deterministic local search strategy that could replace or even precede L_2 in HD-DDS for polishing discrete DDS solutions, especially for case studies with hundreds of decision variables, is the cellular automaton network design algorithm (CANDA) for WDS optimization introduced by Keedwell and Khu (2006). CANDA enables expert designers to encapsulate their case study knowledge in the optimization procedure.

Chapter 3

Hybrid Pareto Archived Dynamically Dimensioned Search for Multi-Objective Combinatorial Optimization: Application to Water Distribution Network Design

This chapter is based on the published article with the same title in the Journal of Hydroinformatics, by Asadzadeh M., and Tolson, B. A., October 2011, volume 14, issue 1, pp. 192-205. References are unified at the end of thesis.

Summary

Pareto archived dynamically dimensioned search, PA-DDS multi-objective optimization algorithm is introduced in this chapter. PA-DDS uses DDS as its search engine and inherits the simplicity and parsimonious characteristics of DDS, so it has only one algorithm parameter and adjusts the search strategy to the user-defined computational budget. It also can handle discrete and/or continuous decision variables by combining the decision variable perturbation strategies of discrete and original versions of DDS. PA-DDS is hybridized by a general discrete local search strategy to improve its performance near the end of the search. The algorithm is applied to five benchmark water distribution network design problems and its performance is assessed in comparison with NSGAII and SPEA2. This comparison is based on a revised hypervolume metric that is introduced in this chapter. The revised metric measures the algorithm performance relative to the observed performance variation across all algorithms in the comparison. The revised metric is improved in terms of detecting clear differences between approximations of the Pareto optimal front. Despite its simplicity, Hybrid PA-DDS shows high potential for approximating the Pareto optimal front, especially with limited computational budget. Independent of the PA-DDS results, the new local search strategy is also shown to substantially improve the final NSGAII and SPEA2 Pareto fronts with minimal additional computational expense.

This chapter is organized as follows. In Section 3.1 a review of related studies is presented. In Section 3.2, PA-DDS algorithm is explained in detail followed by the description of the methodology to assess the performance of PA-DDS. Results are reported in Section 3.3 followed by discussion in Section 3.4.

3.1 Introduction

Multi-objective optimization (MO) algorithms can be categorized as classical and ideal methods (Deb 2001). Each trial of a classical method transforms a MO problem into a single objective optimization problem (e.g., methods that aggregate all objectives into one or epsilon constraint method that optimizes one of the objectives while other objectives are constrained with a degree of satisfaction) and solves the single objective optimization problem to approximate a single solution of the original MO problem. Therefore, several trials of a classical method are required to approximate several solutions that represent the Pareto front (tradeoff) between objectives. In contrast, ideal methods aim to approximate the Pareto front in a single trial. Perhaps MO evolutionary algorithms (MOEAs) are the most commonly applied ideal MO algorithms. In general, ideal MO algorithms search throughout the decision space and archive high quality solutions during the search.

Developing MO algorithms has been an active area of research for many years. Schaffer (1984) introduced the vector evaluated genetic algorithm based on the genetic algorithm to solve MO problems. Since then, several optimization algorithm development studies introduced MO versions of single objective optimization algorithms motivated by special characteristics of single objective optimization algorithms. Example MO algorithms include Pareto Archived Evolution Strategy or PAES by Knowles and Corne (2000), Strength Pareto Evolutionary Algorithm or SPEA2 by Zitzler et al. (2001), Nondominated Sorted Genetic Algorithm or NSGAII by Deb et al. (2002), Multi-objective Particle Swarm Optimization or MOPSO by Parsopoulos and Vrahatis (2002), Multi-objective Shuffled Complex Evolution Metropolis or MO-SCEM by Vrugt et al. (2003) and Multi-objective Cross Entropy by Perelman et al. (2008).

This chapter introduces a new MO algorithm based on the single objective optimization algorithm, Dynamically Dimensioned Search (DDS). The simplicity, parsimony and efficiency of original and discrete versions of DDS motivated the development of this MO algorithm. DDS was first introduced by Tolson and Shoemaker (2007) for solving computationally intensive automatic model calibration problems. DDS was originally designed based on the experience in single objective manual calibration of hydrologic models by Tolson and Shoemaker (2007). In such problems, an initial poor solution can be improved by perturbing multiple decision variables simultaneously, but, a good solution near the end of the search needs to be fine-tuned by perturbing only one or perhaps a few decision variables. DDS mimics this experience by stochastically decreasing the number of decision variables perturbed per iteration. DDS aims to find high quality global solutions (as opposed to

globally optimal solutions). In Chapter 2, the discrete version of DDS was introduced and successfully applied to water distribution system design problems with discrete decision variables. Inspired by the Pareto Archived Evolution Strategy or (1+1)-PAES (Knowles and Corne 2000), Pareto Archived DDS or PA-DDS uses the original and/or discrete versions of DDS as its search engine and archives all non-dominated solutions during the search. More details about PA-DDS are reported in section 3.2.1.

3.1.1 Hybridizing MO algorithms

Local search techniques can be used to improve proximity (Ishibuchi and Murata 1996; Jaszkiewicz 2002) and/or diversity (Talbi et al. 2001; Bosman and de Jong 2006) of the results of MO algorithms. Heuristic neighborhood search strategies such as hill-climbing, simulated annealing, and Tabu search are probably the most common approaches to hybridize MO algorithms; example applications are Ishibuchi and Murata (1996), Knowles and Corne (2000), Deb and Goel (2001), Talbi et al. (2001), Jaszkiewicz (2002), Kleeman et al. (2007). Brown and Smith (2003) combined the steepest-descent MO theory and evolutionary computation to guide the search towards the dominating search direction in each generation. Also, Bosman and de Jong (2006) successfully combined three different gradient techniques (local search strategies) with typical genetic operators. Moreover, Jourdan et al. (2005) introduced LEMMO that uses the learnable evolution model (LEM) to characterize some rules during the search of a MO algorithm such as NSGAII to improve its convergence speed. However, modelling time is required to develop these problem specific rules which can highly affect the quality of the coupled NSGAII and LEM solutions.

Deb and Goel (2001) introduced some hybridization terminology to distinguish between applying local search at the end of the genetic search (posteriori) and during the genetic search (online). Example posterior hybridization include Talbi et al. (2001) and Deb and Goel (2001) and example online hybridization include Ishibuchi and Murata (1996), Jaszkiewicz (2002), Brown and Smith (2003) and Bosman and de Jong (2006). Goel and Deb (2001) compared posteriori and online approaches for hybridizing NSGAII and concluded that the posteriori approach is more efficient since the online approach places too much emphasis on the local search. Also, (Ishibuchi et al. 2003) pointed out the need for a balance between global and local searches and applied the local search to only a few offspring in each generation. Therefore, a new parameter is often required to divide the computational budget between local and global searches. Similar to the other algorithm parameters, finding a proper value for this parameter might be problem specific and time-consuming.

In this chapter, a simple posteriori neighborhood search strategy replaces the DDS component of PA-DDS at the point in the search when the expected value of number of perturbed decision variables per iteration becomes one. This approach is based on the hybridization approach in the previous chapter and removes the need for a new algorithm parameter to change the search strategy. More details about this approach are provided in section 3.2.2.

3.1.2 Comparing MO algorithm performance

Performance metrics that assess the quality of a MO solution by a single number are referred to as unary metrics while binary metrics assess the quality of any two solutions by assigning one number to compare each of the two solutions to the other one; a detailed list of MO performance metrics can be found in Coello et al. (2007). Zitzler et al. (2003) defined compatibility and completeness for performance metrics based on the dominance relation. Based on this definition, only if a performance metric is complete and compatible in terms of weak dominance relation, can its result show whether one solution outperforms the other one. Zitzler et al. (2003) studied various performance metrics and showed that neither a single performance metric nor a combination of finite number of performance metrics can represent a complete and compatible metric with respect to weak dominance relation. Therefore, they concluded that based on a single performance metric or a combination of them it would not be possible to indicate if a MO solution outperforms the other one. However, a performance metric can at best indicate if a MO solution is not worse than (weakly dominated by) another one. As shown in Zitzler et al. (2003), hypervolume or HV introduced by Zitzler and Thiele (1998) is a complete metric with respect to weak dominance relation, that is, preferred solution by HV is not weakly dominated by its opponent.

In this chapter, a slightly modified version of HV is used to assess the results. The revised HV evaluates the algorithm performance relative to the best and worst observed performance across all algorithms in the comparison. Similar to the original HV, the revised HV is complete. Moreover, it is more interpretable than HV when the difference between the worst and the best solutions are practically meaningful.

3.2 Methodology

In general, a multi-objective optimization (MO) problem formulation can be presented as in equation (3-1). Without loss of generality, minimization is considered. A set of m objectives f_i maps the feasible decision space ($X \subseteq \mathbb{R}^n$) into the feasible objective space ($Z \subseteq \mathbb{R}^m$).

$$\min \boldsymbol{f}(\boldsymbol{x}) = f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots x_n)$$
$$\boldsymbol{x} \in \boldsymbol{X} \subseteq \mathbb{R}^n, \boldsymbol{f}(\boldsymbol{x}) \in \boldsymbol{Z} \subseteq \mathbb{R}^m \quad \textbf{(3-1)}$$

As opposed to single objective optimization problems with a single optimal objective function value, MO problems often have conflicting objectives meaning that no single solution $\boldsymbol{x}$ (set of decision variable values) can minimize all objectives simultaneously. The solution of a MO problem consists of a set of solutions without priority to each other with respect to all objectives (i.e., they are not dominated by any other solution). The dominance concept is the core concept of MO research. Comparing two solutions $\boldsymbol{x}$ and $\boldsymbol{y}$ $(\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X} \subseteq \mathbb{R}^n)$ of a MO problem, $\boldsymbol{x}$ weakly dominates $\boldsymbol{y}$ $(\boldsymbol{x} \preccurlyeq \boldsymbol{y})$ as in equation (3-2), $\boldsymbol{y}$ weakly dominates $(\boldsymbol{y} \preccurlyeq \boldsymbol{x})$, or $\boldsymbol{x}$ and $\boldsymbol{y}$ might be non-dominated or incomparable $(\boldsymbol{x} \parallel \boldsymbol{y})$ as in equation (3-3). The set of optimal solutions is usually referred to as Pareto optimal set $(\text{PS} \subseteq \boldsymbol{X})$ and its map to the objective space is called Pareto optimal front $(\text{PF} \subseteq \boldsymbol{Z})$. In general, solving a MO problem means identifying a set of non-dominated solutions $(\text{PS}^a \subseteq \boldsymbol{X})$ that approximates PS and its map to the objective space $(\text{PF}^a \subseteq \boldsymbol{Z})$ that approximates the PF.

$$\boldsymbol{x} \preccurlyeq \boldsymbol{y} \mid \forall i \in \{1, \dots, m\}: f_i(\boldsymbol{x}) \leq f_i(\boldsymbol{y}) \quad \textbf{(3-2)}$$

$$\boldsymbol{x} \parallel \boldsymbol{y} \mid \exists i \in \{1, \dots, m\}: f_i(\boldsymbol{x}) < f_i(\boldsymbol{y}), \exists j \in \{1, \dots, m\}: f_j(\boldsymbol{y}) < f_j(\boldsymbol{x}) \quad \textbf{(3-3)}$$

3.2.1 Pareto Archived Dynamically Dimensioned Search (PA-DDS)

PA-DDS is a heuristic MO algorithm that uses the original and discrete versions of DDS as its search engine to sample from candidate solutions with a single parameter which is the perturbation size r with a default value of 0.2 (as in original and discrete versions of DDS). The two major added components to DDS for designing PA-DDS are archiving and selection strategies. PA-DDS (appendix A- Figure A- 6) uses an unbounded archive and archives (appendix A- Figure A- 7) all non-dominated solutions during the search. PA-DDS uses the parameter-less roulette wheel (De Jong 1975) stochastic selection scheme to select (appendix A- Figure A- 8) one archived solution per iteration and then perturbs (appendix A-Figure A- 9) the selected solution to generate a candidate solution. If the perturbed solution is dominating or non-dominated compared to archived solutions, this recently perturbed solution is archived and then itself perturbed in the next iteration. Otherwise, the selection strategy chooses another archived solution based on a selection metric. Similar to its early version by Asadzadeh and Tolson (2009), PA-DDS uses the crowding distance, CD as in (Deb et al. 2002) measure of all archived non-dominated solutions as the selection metric.

3.2.2 Hybrid PA-DDS for Combinatorial Multi-objective Optimization Problems

A general local search strategy is designed to improve the quality of the PA-DDS results for solving MO problems with discrete decision variables. The pseudo code in appendix A-Figure A- 10 represents the details of the hybrid PA-DDS used in this work. Based on the decision to hybridize the D-DDS in Tolson et al. (2009), PA-DDS is hybridized when the expected number of perturbed decision variables per iteration becomes 1 ($p \leq {}^{1}/_{n}$ in STEP1 of appendix A-Figure A- 10). In PA-DDS, this happens when the iteration count reaches $\mathrm{NFE}^{(n-1)/n}$. At this stage, discrete DDS is replaced by the local search referred to as L which is defined in appendix A-Figure A-11. L is designed to polish a current non-dominated solution by cycling through all possible ways for decreasing or increasing one decision variable at a time by one discrete option.

In hybrid PA-DDS, L is called iteratively in Steps 2 and 3 of appendix A-Figure A- 10. L is first invoked to polish solutions corresponding to the extreme points of the current Pareto approximate front. For polishing extreme points, L restarts the search at decision variable 1 to ensure convergence to a solution that can no longer be improved by L. In Step 3 of appendix A- Figure A- 10, L polishes a portion of the other non-dominated solutions depending on the remaining computational budget. Unlike for extreme solutions, L does not restart the search to converge for other non-dominated solutions. Instead, L spends at most $2n$ (twice the number of decision variables) iterations to polish each of these solutions and this limit helps to ensure that the local search occurs along the entire front. If the remaining computational budget in hybrid PA-DDS is not enough for polishing all archived non-dominated solutions, the range of the first objective function is divided into j equal intervals $\left(j = \mathrm{abs}(\text{remaining budget})/(2n)\right)$ and at least one randomly selected non-dominated solution from each nonempty interval is selected to be polished by L.

3.2.3 Optimization Problem Formulation

The single objective constrained water distribution system (WDS) design problem solved in Chapter 2 is solved as an unconstrained bi-objective optimization problem in this chapter. As proposed by Parmee and Purchase (1994), one way to handle the constraints in a single objective constrained optimization problem without adding any penalty parameter to the problem is to aggregate all constraint violations in an objective function which is being minimized simultaneously with the other objective functions in a multi-objective optimization approach. Fonseca and Fleming (1998) noted that if all constraints cannot be satisfied simultaneously, by using this approach analyst can choose

the final solution from a set of solutions that are not dominated by any other solution though they violate some of the constraints. However, if all constraints can be satisfied simultaneously, a feasible solution with the best possible objective function must be selected.

The problem defined in equation (3-4) represents an unconstrained bi-objective formulation of the constrained single objective WDS design problem solved in Chapter 2. Indices i and j represent pipes and demand nodes respectively, l is the pipe length, $c(\,)$ is the pipe cost per unit length as a function of the decision variable x_i that is a discrete pipe diameter option number and is between option 1 (the smallest diameter) and the maximum diameter option, x_i^{max} for all n pipes in the network to be sized, h_j^{req} is the minimum required pressure head for each demand node j in the network, and h is the pressure at each demand node as a function of $\boldsymbol{x} = [x_1, \dots, x_n]$ and is determined by the network hydraulic simulator which is EPANET2 in this chapter. The same problem formulation is utilized in Farmani et al. (2005), Atiquzzaman et al. (2006), Perelman et al. (2008), and di Pierro et al. (2009). These two objectives are conflicting since pipes of larger diameter cost more and usually reduce the pressure deficit. Decision variables of the problem are pipe diameters that can be selected from a finite set of available pipe sizes and all other network characteristics are known.

$$\begin{aligned} &\min_{\boldsymbol{x}} f_1 = \left(\sum_{i=1}^{n} l_i c(x_i)\right) \\ &\min_{\boldsymbol{x}} f_2 = \max\left(0, \max_j\left\{h_j^{req} - h_j(\boldsymbol{x})\right\}\right) \\ &st.\, x_i \in \{1,2,\dots,x_i^{max}\}, \qquad \forall i = 1,\dots,n \end{aligned} \tag{3-4}$$

3.2.4 Benchmark Water Distribution System Design Problems

The following five WDS design problems are selected from the literature, modelled in EPANET2, and solved in the bi-objective formulation (3-4).

3.2.4.1 New York tunnels problem (NYTP)

NYTP (Schaake and Lai 1969; Zecchin et al. 2005) involves rehabilitation of an existing WDS with 21 pipes and 16 design options per pipe (parallelization with one of 15 tunnel sizes or a do-nothing option). This defines a search space size of 16^{21} (approximately 1.93×10^{25}). The best known least-cost design of NYTP costs \$38.638 million (Maier et al. 2003). So, (0 m, \$38.638 million) is one of the extreme points of the Pareto optimal front in problem formulation (3-4) for NYTP. The other extreme point is (47.6 m, \$0) corresponding to no additional pipe to the current network and accepting the maximum of 47.6 m pressure deficit.

3.2.4.2 Doubled New York tunnels problem (NYTP2)

NYTP2 (Zecchin et al. 2005) is twice as big as NYTP with 42 pipes to be sized from 16 options. This defines a search space size of 16^{42} (approximately 3.74×10^{50}). The best known least-cost design of NYTP2 costs \$77.276 million (Zecchin et al. 2005). So, the best known Pareto front for NYTP2 in problem formulation (3-4) has the extreme points (0m, \$77.276 million) and (47.6 m, \$0).

3.2.4.3 Hanoi problem (HP)

HP (Fujiwara and Khang 1990) has 32 pipes with six options resulting in a search space of 6^{32} combinations (2.87×10^{26} solutions). The single objective version of this problem is reportedly difficult to simply find a feasible solution for (Eusuff and Lansey 2003; Zecchin et al. 2005; Zecchin et al. 2007). Moreover, Farmani et al. (2005) noted that in the bi-objective optimization problem formulation (3-4), finding a fully feasible solution remains difficult for HP. The best known least-cost design of HP costs \$6.081 million (Perelman and Ostfeld 2007) that suggests (0 m, \$6.081 million) as an extreme point of the best known Pareto front in problem formulation (3-4). The other true extreme point that corresponds to the smallest pipe size for all pipes is (17678.5 m, \$1.802 million). Readers are referred to Zecchin et al. (2005) for detailed information about these first three networks.

3.2.4.4 GoYang problem (GYP)

GYP (Kim et al. 1994) is a WDN in South Korea with 30 pipes that should be sized from eight diameter options. This defines a search space size of 8^{30} (approximately 1.24×10^{27}). The best known least-cost design of the network costs 177.01 million Won (Tolson et al. 2009). So, one extreme point of the best known Pareto front is (0 m, 177.01 million Won) and the other extreme point is (125.1 m, 174.673 million Won).

3.2.4.5 Balerma irrigation network problem (BP)

BP is a large and complex WDN with 443 demand nodes, 454 pipes, eight loops, and four reservoirs (Reca and Martinez 2006). Each of the 454 pipes must be sized from 10 possible diameters that define a search space size of 10^{454}. The best known least-cost-design of BP costs €1.9409 million (Tolson et al. 2009). So, the best known Pareto front of BP has extreme points (0m, €1.9409 million) and (5213.7m, €0.724 million); the latter corresponds to the smallest size of all pipes.

3.2.5 Selected performance metrics

3.2.5.1 Normalized hypervolume (NHV)

Hypervolume (HV) was first introduced by Zitzler and Thiele (1998) as a unary performance metric. In principle, HV is the volume of the feasible objective space dominated by a Pareto approximate front. More precisely, HV is the volume of the partition of the objective space that is bounded between the Pareto approximate front and a reference point $\boldsymbol{r}$. HV for an example Pareto approximate front of a bi-objective minimization problem is presented as the shaded area in Figure 3-1. Deb (2001) suggested the calculation of HV in the normalized objective space as in equation (3-5) where coordinates of points, f_i, are normalized, f_i^N, by the maximum and minimum of all m objectives f_i^{max} and f_i^{min} respectively. After this normalization, all results are inside the unit hypercube $[0,1]^m$. This metric is called normalized hypervolume (or simply NHV). Also Van Veldhuizen (1999) proposed the hypervolume ratio (HVR) that is the ratio of HV for the approximate front to the HV for the Pareto optimal front. Hence, the HVR shows the quality of an approximate front in comparison with the Pareto optimal front.

$$f_i^N = \frac{f_i - f_i^{min}}{f_i^{max} - f_i^{min}} \quad \forall i = 1, \dots, m \qquad \textbf{(3-5)}$$

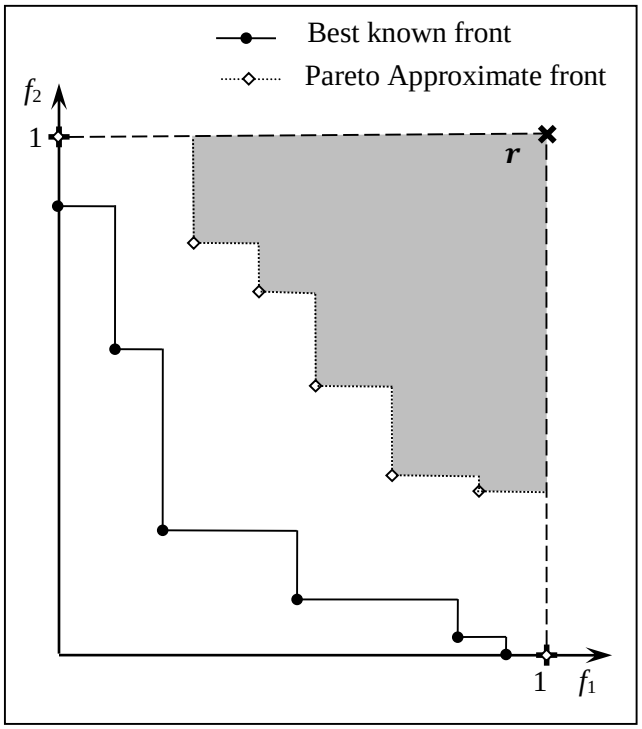

Figure 3-1 Normalized Hypervolume Performance Metric for Minimization of Two Objectives

Although theoretically NHV and HVR can take any value between 0 for the worst result or simply reference point and 1 for the best possible result, they are not necessarily close to 0 for all poor approximate fronts. In fact, very different approximate fronts can have NHV and RHV values with very small numerical differences. As a result, the interpretation of differences in NHV values between algorithms is not always straightforward. Therefore, a slightly modified version of the NHV is proposed to make the interpretation of NHV easier.

3.2.5.2 Comparative Normalized Hypervolume (CNHV)

The modified MO performance metric is referred to as the comparative NHV (CNHV) and is calculated in the normalized objective space for comparing multiple optimization trials of multiple algorithms. With reference to areas A and B in Figure 3-2, the CNHV for an approximate front is equal to $B/(A + B)$. The main difference between CNHV and its precedent performance metrics NHV and HV can be summarized as follows:

- The single reference point is replaced by a set of reference points corresponding to the worst attained front that can be constructed from all algorithm results in the comparison.
- Both the best and the worst attained fronts used in the CNHV are extracted from the results of all MO algorithms that are included in the comparison while in HV, NHV, and RHV, reference point and the best known front are fixed.

CNHV calculation requires the best and worst attained fronts (solid and dashed lines in Figure 3-2, respectively). To this end, final approximate fronts of all the trials of MO algorithms in the comparison are collected in a set. The best attained front is then identified as the subset of solutions from this set that are non-dominated. The worst attained front contains all solutions in this set that are weakly dominated by at least one solution from each optimization trial in the comparison.

Similar to the original HV, the CNHV is complete with respect to the weak dominance relation, that is, it always prefers a Pareto approximate front that is weakly non-dominated. In other words, comparing two Pareto approximate fronts, the better value of the CNHV indicates that the corresponding front is not weakly dominated by the other one.

The value of the CNHV is more directly interpretable than the value of HV, NHV or RHV since it determines how much of all attained results are dominated by each approximate front. As such, CNHV values close to 0 are relatively poor and values close to 1 are relatively good. However, CNHV is not recommended if the best and the worst results are not practically different. Because in

that situation, all results are practically the same quality and it might be misleading to assign a value close to 0 to one or more of the algorithms in the comparison.

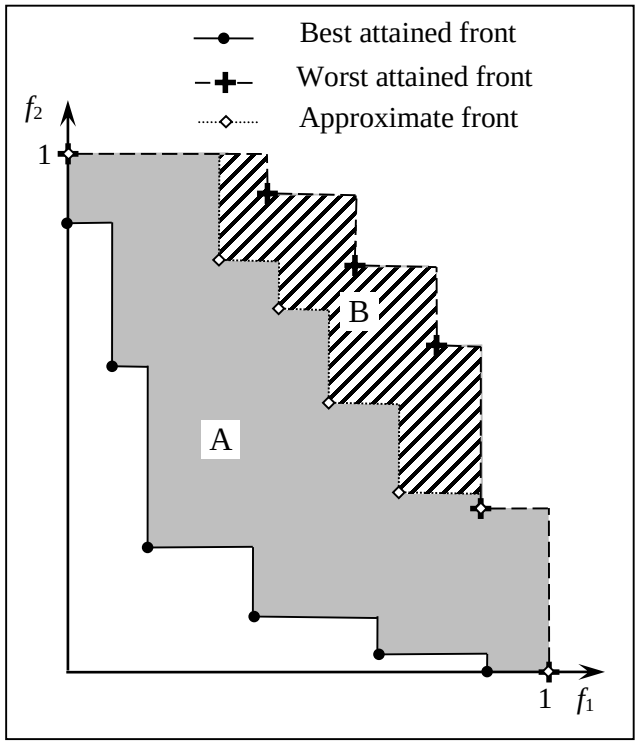

Figure 3-2 CNHV performance metric for minimization of two objectives

3.2.6 Benchmark optimization algorithms

In order to assess the performance of hybrid PA-DDS, NSGAII and SPEA2 are implemented and applied to the same bi-objective WDS design problems. The search engine of NSGAII and SPEA2 is an integer coded GA with a 90 percent chance of uniform crossover and average mutation rate of one decision variable. The other parts of NSGAII and SPEA2 are implemented as in Deb et al. (2002) and Zitzler et al. (2001), respectively. The population size in both NSGAII and SPEA2 is set to 100 except for solving GYP with the limited budget of 2,000 model evaluations where a smaller population size of 50 is used. The population size of 100 and the probability of crossover and mutation were selected based on the parameter values specified in Deb et al. (2002) and no further effort is spent on fine-tuning them.

3.3 Results

Results are presented in two subsections. In section 3.3.1, results of PA-DDS, NSGAII and SPEA2 are compared for solving problem formulation (3-4) for all five case studies. In section 3.3.2, the effectiveness of the local search is evaluated by applying it to both NSGAII and SPEA2.

3.3.1 Hybrid PA-DDS versus NSGAII and SPEA2

Each of the five case studies is solved by hybrid PA-DDS, NSGAII, and SPEA2, with a rather large computational budget (determined from typical budgets utilized for the case studies in previous publications) and with a more limited computational budget (one order of magnitude less). Hybrid PA-DDS has two termination criteria, computational budget and the local search convergence. For NYTP, NYTP2 and GYP solved with the higher computational budget, the local search usually converges before spending the whole computational budget. This is why the computational effort for these three cases is less than the total budget. Table 3-1 summarizes some statistics of the two performance metrics called NHV and CNHV proposed in this chapter. With the limited computational budget, Hybrid PA-DDS achieves the best NHV and CNHV values for all case studies. But, with the higher computational budget, NSGAII performs better than Hybrid PA-DDS in NYTP2 and BP.

Figure 3-3 compares the best attained front and the worst CNHV fronts corresponding to the italic numbers in Table 3-1 for all three algorithms solving each of the five case studies with the higher computational budget. The best attained front is obtained from results of all trials of all three algorithms and is a discrete front; however, its points are piecewise linearly connected for illustrative purposes. For NYTP, NYTP2, and GYP, the best known endpoints (see section 3.2.4) are captured in the best front. However, the best obtained endpoint corresponding to least-cost design of HP is (0 m, \$6.096M) instead of (0 m, \$6.081M) and for BP (0 m, €2.115M) instead of (0 m, €1.9409M). The other best known endpoints of these two cases are identified in the best attained front.

Table 3-1 Algorithm comparison based on NHV and CNHV. **BOLD** numbers represent the best result for each case study, and *ITALIC* numbers corresponds to trials whose Pareto approximate front is plotted in Figure 3-3

Case Study (# of trials)	Budget*	Statistic	NHV			CNHV		
			Hybrid PA-DDS	NSGAII	SPEA2	Hybrid PA-DDS	NSGAII	SPEA2
NYTP (50)	50,000	Avg.	**0.85**	0.85	0.85	**1.00**	0.96	0.95
		Best	**0.85**	0.85	0.85	**1.00**	0.98	0.99
		Worst	**0.85**	0.85	0.85	***0.99***	*0.92*	*0.80*
	5,000	Avg.	**0.82**	0.74	0.76	**0.90**	0.53	0.62
		Best	**0.84**	0.79	0.81	**0.96**	0.75	0.83
		Worst	**0.79**	0.67	0.70	**0.78**	0.27	0.38
NYTP2 (50)	300,000	Avg.	0.85	**0.85**	0.85	0.87	**0.91**	0.84
		Best	**0.85**	**0.85**	0.85	0.95	**0.95**	0.92
		Worst	0.84	**0.85**	0.84	0.77	**0.83**	0.69
	30,000	Avg.	**0.82**	0.79	0.79	**0.82**	0.60	0.62
		Best	**0.83**	0.82	0.83	**0.91**	0.81	0.84
		Worst	**0.80**	0.74	0.73	**0.69**	0.29	0.21
HP (50) trials	100,000	Avg.	**1.00**	0.98	0.80	**0.99**	0.96	0.65
		Best	**1.00**	0.98	0.93	**1.00**	0.97	0.88
		Worst	**0.99**	0.95	0.67	**0.99**	0.91	0.43
	10,000	Avg.	**0.92**	0.90	0.89	**0.96**	0.54	0.51
		Best	**0.96**	0.94	0.93	**0.98**	0.84	0.75
		Worst	**0.88**	0.86	0.82	**0.93**	0.30	0.04
GYP (50)	20,000	Avg.	**0.93**	**0.93**	0.93	**1.00**	1.00	0.99
		Best	**0.93**	**0.93**	**0.93**	**1.00**	1.00	1.00
		Worst	**0.93**	0.93	0.91	**1.00**	0.99	0.58
	2,000	Avg.	**0.89**	0.62	0.65	**0.94**	0.43	0.49
		Best	**0.92**	0.79	0.77	**0.99**	0.74	0.71
		Worst	**0.83**	0.39	0.42	**0.83**	0.02	0.07
BP (10)	1,000,000	Avg.	0.96	**0.97**	0.93	0.77	0.88	0.35
		Best	0.96	**0.97**	0.94	0.82	0.92	0.45
		Worst	0.96	**0.96**	0.93	0.72	0.83	0.30
	100,000	Avg.	**0.91**	0.83	0.81	**0.85**	0.35	0.28
		Best	**0.92**	0.84	0.83	**0.87**	0.42	0.38
		Worst	**0.90**	0.81	0.78	**0.78**	0.24	0.10

* Hybrid PA-DDS converged before spending the entire computational budget in NYTP, NYTP2 and GYP for the higher computational budget. Therefore, its computational effort for these cases is less than the computational budget. 35,000 for NYTP, 263,400 for NYTP2 and 17,330 for GYP

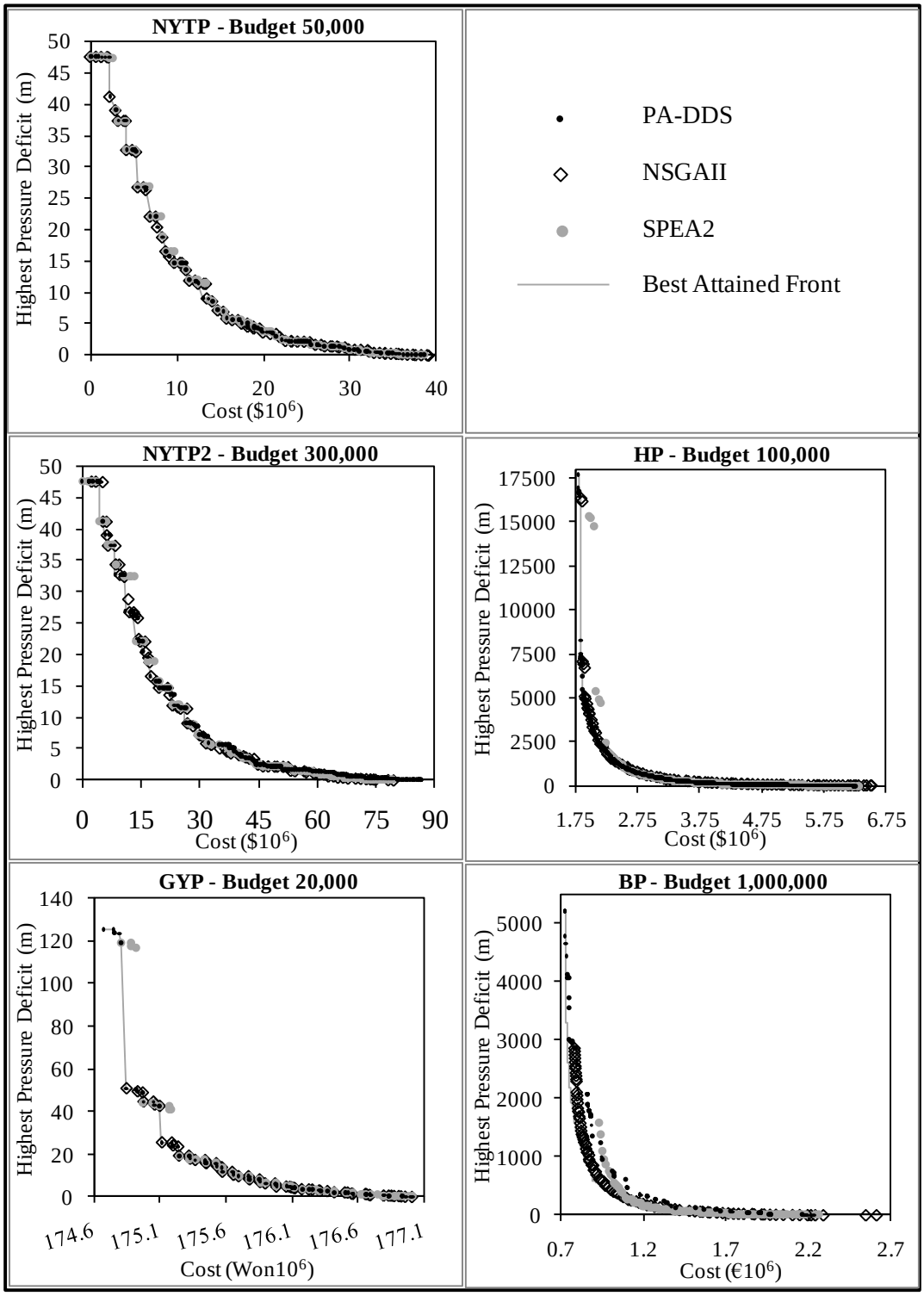

Figure 3-3 Comparison between the best attained front and the worst CNHV trials of all three algorithms solving each case study

Comparing NHV and CNHV values in each row of Table 3-1, CNHV more clearly detects the difference between the approximate fronts and therefore between algorithms. For example, in the HP case study with a computational budget equal to 10,000, the best trial of hybrid PA-DDS has a NHV equal to 0.96 while it is 0.94 for the best trial of NSGAII. This value means that the best trial of Hybrid PA-DDS covered (dominated) 96% of the area between the best attained front and the reference point while the best trial of NSGAII covered 94% of this area. Although this difference seems negligible, Figure 3-4-A shows that there is a considerable difference between the corresponding approximate fronts. For the same trials, the CNHV values are 0.98 and 0.84 denoting that the best trial of hybrid PA-DDS dominated 98% of the area between the best and worst attained fronts considering all 150 approximate fronts from 50 trials of each algorithm compared to only 84% for the best trial of NSGAII. Also comparing the best trial of these two algorithms for solving BP with a computational budget equal to 1,000,000, NHV is equal to 0.96 and 0.97 for Hybrid PA-DDS and NSGAII respectively, while CNHV magnifies the difference and results in 0.82 and 0.92, respectively. Figure 3-4-B demonstrates that the difference detected by CNHV is really considerable and NSGAII performed better than Hybrid PA-DDS.

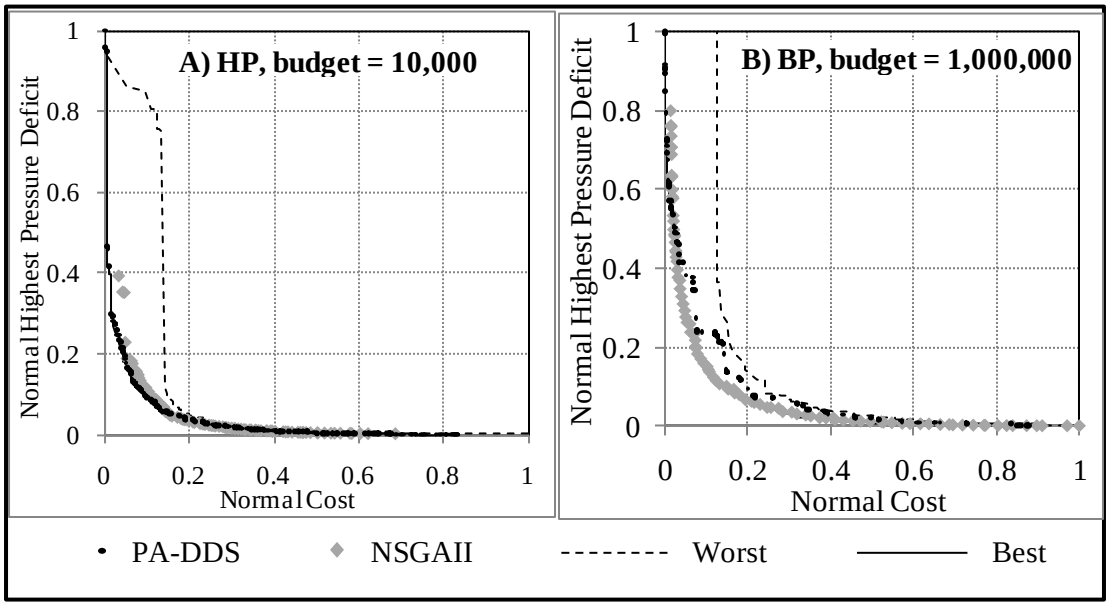

Figure 3-4 Comparison between the best CNHV trial of Hybrid PA-DDS and NSGAII for solving (A) HP, Extreme solutions (17678.5 m, $1.803M) and (0 m, $7.470M) are used to normalize the fronts, and (B) BP, Extreme solutions (5213.7.5 m, €0.724M) and (0 m, €2.3620M) are used to normalize the fronts.

3.3.2 Local search performance assessment

To evaluate the effectiveness of the proposed local search strategy, it was also applied to the results of NSGAII and SPEA2. However, as these two algorithms spent the whole computational budget for the global search, the local search is only applied to the extreme points of the resultant fronts with the computational budget equal to the average budget of local search in Hybrid PA-DDS. This computational budget and the average improvement in results based on the CNHV are summarized in Table 3-2. Although the local search is only applied to the extreme points of the front with very limited computational budget, it highly improved the results of NSGAII and SPEA2 especially when the total computational budget is limited.

Table 3-2 Evaluating the effectiveness of the local search for improving the extreme points of the approximate fronts based on the average percent improvement in CNHV.

		NSGAII		SPEA2	
Case study	Global search budget	Local search budget	Avg. imp. CNHV (%)	Local search budget	Avg. imp. CNHV (%)
NYTP	50,000	49	1.0	48	1.0
	5,000	183	34.0	160	17.7
NYTP2	300,000	111	0.0	127	0.0
	30,000	413	20.3	376	19.4
HP	100,000	166	1.0	251	36.9
	10,000	808	72.2	788	84.3
GYP	20,000	56	0.0	67	1.0
	2,000	171	107.0	191	83.7
BP	1,000,000	27,600	6.8	24,200	114.3
	100,000	2,500	193.8	2,500	256.0

3.4 Discussion and Conclusions

A simple and parsimonious optimization algorithm for solving multi-objective optimization problems was introduced in this chapter. The algorithm is called PA-DDS hybridized with a straightforward neighbourhood search. It was successfully applied to five benchmark bi-objective Water Distribution system design problems. The hypervolume performance metric was modified to define the new comparative normalized hypervolume metric which makes the hypervolume metric more interpretable for comparing multiple trials of multiple algorithms. Results show the comparable performance of PA-DDS with two well-known MO algorithms, NSGAII and SPEA2.

It should be noted that the local search L is implemented such that it always evaluates one option change at a time relative to the current solution. Therefore, the local search order (starting for example at decision variable n instead of decision variable 1 as implemented in this chapter) can

change the results. However, finding the best order is not the purpose of this chapter. Moreover, it is not claimed that the proposed neighborhood search strategy is the most efficient local search, but perhaps the simplest one that adequately improves the algorithm efficiency.

If the purpose of MO algorithm comparison is to assess the performance of various algorithms with various computational budgets, the best and worst attained fronts corresponding to each computational budget can be used. This recommendation makes the comparison hard to replicate. Nonetheless, calculating computational budget specific CNHV (and thus the best and the worst attained fronts) is more appropriate since the attainable objective space obviously varies as a function of computational budget.

The comparison of algorithms was based on default configurations/parameter settings that WDN modellers would most likely utilize when trying to solve their own MO design problems with these algorithms. Comparative results might change if each algorithm was fine tuned to optimally solve each problem but that would generally require substantial computational experiments. Instead, parameters of NSGAII and SPEA2 are set to the recommended values from literature while the design decisions and single parameter value of the Hybrid PA-DDS algorithm were based on previous decisions for DDS (Tolson and Shoemaker 2007) and HD-DDS (Tolson et al. 2009).

NSGAII and SPEA2 have a fixed-size archive and if it becomes full of non-dominated solutions, they ignore some current non-dominated solutions. Although this strategy controls and limits the complexity of the algorithm, it may have a disadvantage. Laumanns et al. (2002) showed that the standard archiving strategy of NSGAII or SPEA2 allows the algorithm eliminate some high quality solutions that in the long run might even dominate some of the non-dominated solutions in the final archive. To deal with this issue some new archiving strategies have been proposed; see for example Laumanns et al. (2002) and Beume et al. (2007). The current version of PA-DDS archives all non-dominated solutions during the search. Therefore, it does not lose any non-dominated solution; however, this leads to two related challenges for PA-DDS.

The first challenge corresponds to the algorithm efficiency (runtime) for solving large scale problems. In PA-DDS, any new non-dominated solution is checked against all current non-dominated solutions. Hence, the higher the number of archived solutions the more time required for the dominance check. This may affect the efficiency of PA-DDS for solving problems with many objective functions and a huge computational budget since either of these factors can generate excessive non-dominated solutions. However, in real-world engineering problems, where simulation

is very time consuming and hence the total computational budget is not so huge to allow numerous solutions in the archive, time of dominance check may not have a considerable impact on the algorithm runtime. For these five WDN bi-objective problems, PA-DDS always had a shorter serial runtime than SPEA2 while NSGA2 was always the quickest among all three algorithms. Over all of the five case studies with both computational budgets, PA-DDS runtimes were on average only 13 percent (with extremes of 7 and 36 percent) longer than NSGAII. Therefore, the first challenge was not an important issue for solving these five problems.

The second possible challenge is related to the selection process. The current selection scheme of PA-DDS is designed to sample more from less crowded parts of the approximate front. However, the most interesting parts of the tradeoff which may not coincide with the less crowded part of the front. So, the performance of PA-DDS can be improved by modifying the selection process to guide the search towards the most interesting parts of the tradeoff. This improvement can be significant especially for solving computationally intensive problems where a limited computational budget is available. The selection scheme of PA-DDS is investigated later in this thesis.

Although the relative PA-DDS performance is not investigated on large distribution networks (i.e., thousands of pipes/decision variables), PA-DDS with the same parameter value and configuration is applied to problems with 21 to 454 decision variables and a computational budget ranging from 2,000 to 1,000,000 hydraulic model evaluations. As such, PA-DDS could be applied to even larger distribution networks and in such a case it would be suggested to apply it without any algorithmic and/or parameter modifications. For larger distribution networks, the efficiency of local search L to refine the extreme solutions will degrade substantially relative to efficiencies reported in Table 3-2. Future algorithm comparison studies focused only on very large distribution networks are necessary to properly assess relative PA-DDS performance in this context.

It should be noted here that, problem formulation (3-4) is an artificial WDN bi-objective problem. Therefore, its result cannot be used for designing real WDN problems. For example, all Pareto optimal solutions returned in this chapter have a pressure deficit (some even have negative pressures). Even the extreme point corresponding to the least-cost design of the network is impractical since it tends to reduce pipe sizes or completely eliminate some pipes. This may lead to an insufficient capacity to handle system failures (Walski 2001). The problem formulation (3-4) is solved as the benchmark WDN problem type to assess the relative performance of the proposed MO algorithm just as many previous studies have done, e.g. Farmani et al. (2005), Atiquzzaman et al. (2006), Perelman

et al. (2008), di Pierro et al. (2009). Although not demonstrated here, Hybrid PA-DDS can be applied to more realistic WDN design studies that have more than two objectives. Further work should be conducted to compare Hybrid PA-DDS performance relative to other algorithms on problems with more than two objectives.

Chapter 4

Pareto Archived Dynamically Dimensioned Search with Hypervolume Based Selection for Multi-objective Optimization

This chapter is based on the accepted article with the same title in the Engineering Optimization journal, by Asadzadeh M., and Tolson B. A., 23 May 2012. References are unified at the end of thesis.

Summary

This chapter represents the continued development of PA-DDS by assessing the impact of selection metric on PA-DDS performance. It is shown that hypervolume contribution (HVC) is a very effective selection metric for PA-DDS when solving problems with two or three objective functions and Monte-Carlo-Sampling-based HVC is very effective for higher dimensional problems (5-objective in this chapter). Since PA-DDS is designed to solve computationally intensive problems (e.g., available computational budget limits the total number of solutions that can be evaluated), the performance of PA-DDS with HVC-based selection is empirically assessed for solving a hydrologic model calibration problem with a somewhat limited computational budget. Although solving problems with limited budget is the main concern of this chapter, the performance of PA-DDS with HVC-based selection is empirically evaluated for solving recently proposed mathematical multi-objective optimization problems with a sufficiently large computational budget to determine if PA-DDS can compete with recently developed MO algorithms.

This chapter is organized as follows. Section 4.1 is devoted to the review of related studies. In section 4.2, the alternative selection metrics are introduced. Section 4.3 provides detailed information about the numerical experiments and results and findings are reported in Section 4.4.

4.1 Introduction

The solution archiving process of multi-objective optimization (MO) algorithms is mainly based on the proximity of solutions and is most often measured by the dominance rank which prefers solutions dominated by fewer other solutions. Laumanns et al. (2002) listed three reasons for bounding the archive size of MO algorithms. First, the computation time grows as the number of archived solutions grows (e.g., due to the dominance test). Second, it is not necessary to have a large number of solutions in the process of decision making. Third, limiting the archive size allows the MO algorithm focus on regions of attraction rather than the whole set of non-dominated solutions. Therefore, when the number of first-rank solutions (i.e., non-dominated solutions) is more than the size of bounded archive, some extra solutions should be discarded from the archive. Environmental selection functions to discard extra solutions from highly populated parts of the archive and is often based on a measure of diversity of solutions such as niching in NSGA (Srinivas and Deb 1994), crowding distance (CD) in NSGAII (Deb et al. 2002), strength in SPEA2 (Zitzler et al. 2001) and hypervolume contribution (HVC1) in SMS-EMOA (Emmerich et al. 2005).

Hanne (1999) studied the convergence of MO Evolutionary Algorithms (MOEA), i.e. if all archived solutions converge to the Pareto optimal set, and concluded that convergence depends on the archiving strategy of the MO algorithm and can only be guaranteed if deterioration is avoided. Deterioration may occur in a bounded size archive if in a generation some non-dominated solutions are discarded from archive and later on in the search, some worse solutions are generated and archived. Hanne (1999) resolved this issue by archiving only dominating solutions (dominating a current archived solution). However, Laumanns et al. (2002) noted that this approach does not guarantee good diversity of solutions because when all archived solutions are Pareto optimal, no new solution can enter the archive to improve the diversity of solutions.

Laumanns et al. (2002) demonstrated the deterioration behavior of NSGAII and proposed an archiving strategy that avoids the deterioration. This strategy discretizes the objective space into grid cells and archives at most one solution in each cell. In this strategy, the dominance filter is applied to the grid cells to distinguish non-dominated grid cells. Inside each grid cell, only the non-dominated solution closest to the dominating corner of the cell is archived. This strategy archives a new solution only if it is in an empty non-dominated grid cell or dominates one archived solution. Convergence of this archiving strategy is guaranteed because deterioration is avoided. Also, grid cells guarantee the

good distribution of solutions in the objective space. This archiving strategy is used in some MO algorithms including ε-MOEA (Deb et al. 2003) and ε-NSGAII (Kollat and Reed 2005).

The most straightforward strategy to update the archive during the search is the linear dominance test that in the worst case compares each new solution with all archived solutions. However, dominance test based on data structures is more efficient when the archive size is large (Mostaghim et al. 2002). Most MOEAs have bounded archives, but some studies including Fieldsend et al. (2003), Chen and Lee (2007), Yang (2007), Smith et al. (2008), and Kaylani et al. (2010) proposed unbounded archives for all non-dominated solutions. Performance of MOEAs with bounded and unbounded archives and linear and data-structured dominance tests were compared in Fieldsend et al. (2003) and results show that the unbounded archive significantly improves the performance of MOEAs and using a data structure significantly reduces the computation time of unbounded archives.

Besides environmental selection, MO algorithms often have a mating selection process to select some archived solutions for generating new solutions. Mating selection in most MOEAs is based on the dominance rank. In mating selection, the lower the dominance rank the better and ties are often broken based on another metric such as CD in NSGAII, a secondary rank as in MO-CMA-ES (Igel et al. 2007) or random selection as in SMS-EMOA (Emmerich et al. 2005).

4.2 Methodology

The main goal of this chapter is to identify the most effective selection metric for PA-DDS applied to general MO problems without any prior knowledge about the shape of the Pareto front. To this end, the performance of PA-DDS for solving mathematical MO problems is evaluated under four alternative selection metrics namely crowding distance (CD) as in Deb et al. (2002), hypervolume contribution (HVC1) as in Knowles et al. (2003), hypervolume contribution (HVC2) as in Bader and Zitzler (2011), and uniform random (RND) selection as in Emmerich et al. (2005).

4.2.1 Selection Metric

PA-DDS uses an unbounded archive and its environmental selection is simply to archive all non-dominated solutions (see appendix A- Figure A- 7). So, PA-DDS does not suffer from deterioration as defined by Hanne (1999). The mating selection of PA-DDS is the roulette wheel as in appendix A- Figure A- 8 (De Jong 1975) based on a selection metric other than dominance rank because all archived solutions in PA-DDS are non-dominated or first-rank. The original version of PA-DDS used CD for mating selection. In this chapter, the effectiveness of CD on the performance of PA-DDS is

assessed in comparison with two alternative selection metrics and uniform random selection. The two selection metrics are two variants of hypervolume contribution, HVC1 as in Knowles et al. (2003) and HVC2 as in Bader and Zitzler (2011). Testing the effectiveness of alternative metrics means altering the metric selection type specified (RND, CD, HVC1 or HVC2) in the PA-DDS algorithm described in appendix A-Figure A- 6.

HVC1 measures the volume of the partition of objective space that is solely dominated by each solution of a set of non-dominated solutions. It has never been directly used as the primary mating selection metric of a MO algorithm before. However, MOCMA-ES ranks the solutions based on HVC1 and uses the rank instead of the HVC1 values in its mating selection (Igel et al. 2007). Calculation of HVC1 is shown in equation (4-1) for a solution $\boldsymbol{x}^p$ in the Pareto approximate set of solutions PS^a as the volume (VOL) of the partition of objective space $\boldsymbol{Z}$ that is solely dominated by $\boldsymbol{x}^p$. HVC1 was first introduced by Knowles et al. (2003) as an environmental selection strategy of MO algorithms that generate a single solution per generation in order to bound the archive size by identifying and discarding the solution with the least HVC1 value. The SMS-EMOA algorithm uses HVC1-based archiving strategy (Emmerich et al. 2005).

$$\mathrm{HVC1}(\boldsymbol{x}^p \in \mathrm{PS}^a) = \mathrm{VOL}(\boldsymbol{z} \in \boldsymbol{Z} | \, \boldsymbol{f}(\boldsymbol{x}^p) \leq \boldsymbol{z} \, , \nexists \boldsymbol{x}^q \in \mathrm{PS}^a : \boldsymbol{f}(\boldsymbol{x}^q) \leq \boldsymbol{z}) \qquad \textbf{(4-1)}$$

Igel et al. (2007) compared two versions of MO-CMA-ES: one with the NSGAII selection scheme and one with HVC1 as a secondary ranking scheme (in the event of ties in the primary dominance ranking scheme) for both environmental and mating selections. MO-CMA-ES generates a population of solutions (not necessarily one) per generation and ranks them based on the dominance rank. The latter version identifies the solution with the least HVC1 value, ranks it last, and recalculates HVC1 for remaining solutions to identify the second last rank solution and so on. The latter version of MO-CMA-ES primarily uses the dominance rank in both environmental and mating selections and ties are broken based on the secondary rank. The latter version showed superiority to the former version.

HVC2 was first proposed by Bader and Zitzler (2011) for mating and/or environmental selections. HVC2 shares the volume of each partition of the dominated objective space between solutions that dominate it. More precisely as in equation (4-2), the volume of the partition of objective space that is solely dominated by solution $\boldsymbol{x}^p$ is fully assigned to $\boldsymbol{x}^p$. Also, half volume of partition of the objective space that is dominated jointly by $\boldsymbol{x}^p$ and one more solution $(\boldsymbol{x}^1)$ is assigned to $\boldsymbol{x}^p$ and so on until all partitions of feasible objective space dominated jointly by $\boldsymbol{x}^p$ and at most $k-1$ other

solutions is considered. HVC2 calculation depends on parameter k that defines the maximum number of solutions including $\boldsymbol{x}^p$ that should be considered in the calculation of HVC2 for $\boldsymbol{x}^p$. For mating selection, k should be the number of all other solutions in PS^a and for environmental selection, k can be the number of solutions that must be discarded from the archive (Bader and Zitzler 2011).

$$HVC2(\boldsymbol{x}^p \in \mathrm{PS}^a, k, \boldsymbol{r}) = \sum_{i=1}^{k} \frac{\mathrm{VOL}(\boldsymbol{z}\in\boldsymbol{Z} \mid \exists \boldsymbol{A}\subseteq \mathrm{PS}^a\setminus\boldsymbol{x}^p : \|\boldsymbol{A}\|=i-1, \forall \boldsymbol{x}\in\{\boldsymbol{A}\cup\boldsymbol{x}^p\}: \boldsymbol{f}(\boldsymbol{x})\preceq\boldsymbol{z}\preceq\boldsymbol{r})}{i} \quad \textbf{(4-2)}$$

In Figure 4-1, $\mathrm{CD}(\boldsymbol{x}^p)$, $\mathrm{HVC1}(\boldsymbol{x}^p)$, and $\mathrm{HVC2}(\boldsymbol{x}^p, 5, \boldsymbol{r})$ are shown for solution $\boldsymbol{x}^p$ in the objective space of an example bi-objective optimization problem with five archived solutions. It should be noted that all these three metrics are calculated in the normalized objective space as in equation (3-5). Moreover, calculation of HVC2 requires a reference point $\boldsymbol{r}$ that is selected to be the maximum of all objectives, i.e. all coordinates of the reference point equal one in the normalized objective space.

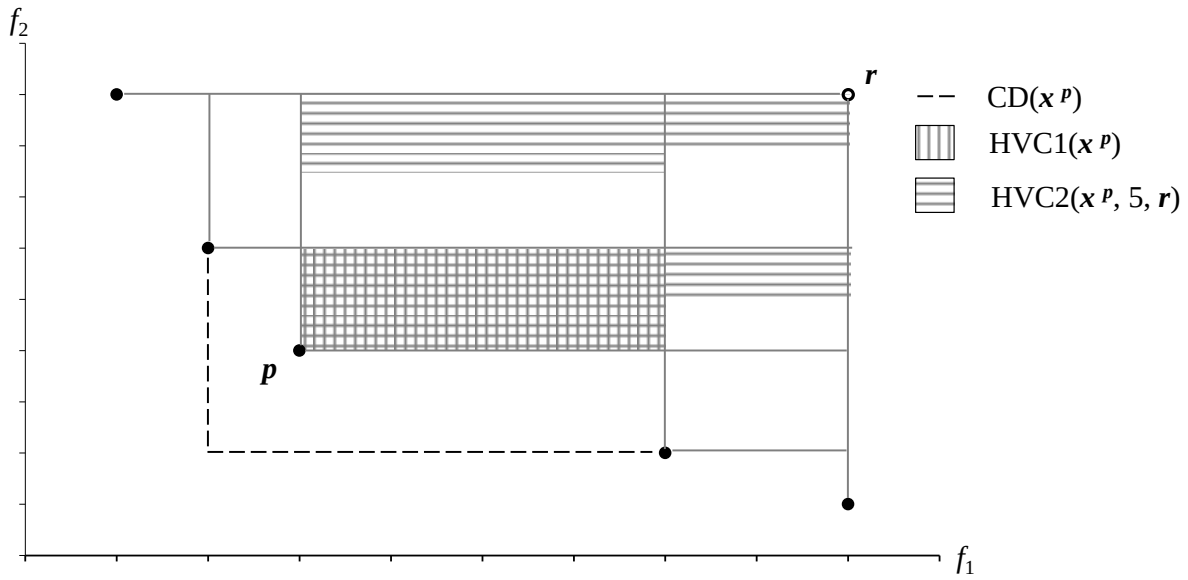

Figure 4-1 Comparison between CD (line segment lengths), HVC1 and HVC2 (both are areas) for solution $\boldsymbol{x}^p$ in an example bi-objective optimization problem.

The exact value of HVC1 is calculated by the fast algorithm introduced by Fonseca et al. (2006) and is publically available at http://iridia.ulb.ac.be/~manuel/hypervolume. Complexity of calculating HV is exponential in number of objectives, m (While 2005), so its calculation becomes too time-consuming in problems with more than four or five objectives; therefore, Monte Carlo (MC) sampling was proposed to estimate HVC1 (Bader et al. 2008) and HVC2 (Bader and Zitzler 2011). This approach samples a set of uniformly distributed random points in the objective space between the ideal point and the reference point. For each point on Pareto approximate front, HVC1 can be

estimated by the ratio of MC samples that are dominated only by the corresponding solution multiplied by the volume of the objective space between Utopia and Nadir points. Following equation (4-2), the ratio of dominated MC samples should be divided between all dominating solutions for calculating HVC2. For MO problems with more than three objectives, this approach is used in PA-DDS to estimate HVC1 and HVC2 with 10,000 samples as in (Bader et al. 2008; Bader and Zitzler 2011) rather than calculating their exact values.

4.3 Numerical Experiments

Three sets of numerical experiments are conducted. Experiment 1 in section 4.3.1 is designed to determine the most effective selection metric for PA-DDS when solving general multi-objective optimization problems. Based on this experiment, the most effective selection metric for PA-DDS is designated and utilized in Experiment 2 in section 4.3.2 and in Experiment 3 in section 4.3.3. These final two sets of experiments are designed to evaluate relative performance of PA-DDS against a wide variety of benchmark MO algorithms. Since PA-DDS is initially designed for solving water resources problems, in Experiment 2, PA-DDS is compared against two high-quality MO algorithms commonly applied to water resources MO problems with a somewhat limited computational budget. Experiment 3 is conducted with a computational budget that is an order of magnitude higher than that used in Experiment 2 and is designed to compare PA-DDS against results of 13 MO algorithms that participated in the IEEE 2009 Congress on Evolutionary Computation, CEC09 MO competition (Zhang et al. 2008). Test problems in Experiment 3 include 2-, 3- and 5-objective problems and thus provide the opportunity to observe how the number of objectives impacts the PA-DDS performance.

4.3.1 Experiment 1: Choosing the Selection Metric for PA-DDS

This experiment is designed to choose for PA-DDS the most effective selection metric from uniform random selection (RND), crowding distance (CD) and two versions of hypervolume contribution, HVC1 and HVC2. This experiment is conducted by solving three benchmark mathematical MO problems. The systematic approach in equation (4-3) was proposed by Deb (1999) to design scalable bi-objective optimization problems in the number of decision variables. h is a function that controls the PF shape (convexity or continuity), g is a function that controls the convergence difficulty and f is a function that controls the distribution of points on PF.

$$\begin{aligned} &\min f_1(\boldsymbol{x}) = f(x_1, \dots, x_m) \\ &\min f_2(\boldsymbol{x}) = g(x_{m+1}, \dots, x_n) \times h(f_1, g) \end{aligned} \quad \textbf{(4-3)}$$

Based on equation (4-3), bi-objective ZDT and multi-objective DTLZ test problems were designed by Zitzler et al. (2000) and Deb et al. (2001), respectively. ZDT4 with 10 decision variables has 21^9 local fronts and tests the ability of algorithms in escaping from local optima. Deb et al. (2001) solved the 3-objective DTLZ problems, and their results show that DTLZ6 is a relatively difficult problem. Huband et al. (2006) also reported the bi-objective DTLZ6 with 24 decision variables is a relatively difficult problem. So, ZDT4 with 10 decision variables, DTLZ6_2D and DTLZ6_3D (2- and 3-objective versions of DTLZ6) with 24 decision variables are included in Experiment 1.

Huband et al. (2006) studied the approach in equation (4-3) and noted that, ZDT and DTLZ test suites cannot test all features of MO algorithms mainly due to the lack of parameter dependencies. To address this weakness Huband et al. (2006) designed WFG suite of test problems. WFG1 and WFG8 are relatively difficult MO problems. Since WFG1 is used in Experiment 3 and experiments are designed to be independent, bi-objective WFG8 is selected as the last case study for Experiment 1.

In Experiment 1, each MO problem is solved 50 times with two computational budgets: the budget that is used in the original reference for each MO problem (25,000) and a budget that is one order of magnitude less (2,500). The purpose of using the limited budget is to investigate the dependency of results to the computational budget. This is important for PA-DDS that is designed to adjust the search strategy to the user-defined computational budget. Thus, if the relative effectiveness of selection metrics depends strongly on the computational budget, a heuristic might be required to switch between various selection metrics based on the budget and/or during the search. It should be noted that, in this experiment and for each computational budget, the only difference in the PA-DDS algorithm is the selection metric type in appendix A-Figure A- 6 and the random seed.

4.3.2 Experiment 2: Solving a Water Resources Multi-objective Optimization Problem

PA-DDS was originally designed for solving computationally relatively intensive MO problems. In Experiment 2, PA-DDS performance is assessed by solving a MO hydrologic model calibration problem introduced in the following section.

4.3.2.1 SWAT 2000: Town Brook Watershed

Town Brook is a 37km^2 sub-watershed located upstream of the Cannonsville watershed in upstate New York. It was modelled in the Soil and Water Assessment Tool version 2000 (SWAT 2000) and the 26 continuous parameters of the model were calibrated by Tolson and Shoemaker (2007) to adequately simulate flow (cms), total suspended sediment transport (kg), and total phosphorus

delivery (kg) measurements. The parameter description and ranges can be found in Tables 2 and 3 in Tolson and Shoemaker (2007). To measure the calibration quality, Tolson and Shoemaker (2007) used the reduced Nash-Sutcliffe (RNS) coefficient in equation (4-4) where m_i and s_i denote measured and simulated data (flow, sediment and/or phosphorus delivery) respectively at day i and t is the total number of days in the calibration period (considered in the calculation of objective functions). The simulation period was from Jan. 1st 1996 to Sept. 30th 2000 with a 639-day warm-up period resulting in a 1096-day calibration period (Oct. 1st 1997 to Sept. 30th 2000). Due to the data quality and availability, the calibration period for the water quality constituents (total phosphorus delivery and sediment transport) was from Oct. 1st 1998 to Sep. 30th 2000 (731 days).

$$\text{RNS} = 1 - \frac{\sum_{i=1}^{t}(m_i - s_i)^2}{\sum_{i=1}^{t}\left(m_i - \frac{\sum_{i=1}^{t} m_i}{t}\right)^2} - \max\{0, (|\text{Bias}| - b)\} \quad \textbf{(4-4)}$$

$$\text{Bias} = \frac{\left(\sum_{i=1}^{t} s_i - \sum_{i=1}^{t} m_i\right)}{\sum_{i=1}^{t} m_i} \quad \textbf{(4-5)}$$

Bias in equation (4-4) penalizes solutions that simulate the measured data with more bias than a specific threshold b. The RNS range is (-∞,1] and its higher value represents better fit between measured and simulated data (RNS $= 1$ is the perfect fit). Based on equations (4-4) and (4-5), for a simple model that results in the average observed values for the whole period t, RNS $= 0$. So, for a high quality model, RNS is expected to be higher than 0.

Tolson and Shoemaker (2007) aggregated RNS for the three measurements and solved the single objective optimization with a reasonably large computational budget of 10,000 solution evaluations. b was set to 10% for flow and 30% for sediment and phosphorus delivery based on the error in the measured data. In this chapter, the 3-objective version of this problem is solved to maximize RNS (minimize negative RNS) for each of the flow, total suspended sediment transport, and total phosphorus delivery measurements.

4.3.2.2 Benchmark Multi-Objective Optimization Algorithms

In order to comparatively assess the performance of PA-DDS, the above MO model calibration problem is also solved by two benchmark MO algorithms ε-NSGAII and AMALGAM in the field of water resources engineering. These two MO algorithms are briefly introduced as follows.

Epsilon Non-dominated Sorting Genetic Algorithm II (ε-NSGAII)

ε-NSGAII is a variant of NSGAII which uses the archiving strategy proposed by Laumanns et al. (2002). This archiving strategy discretizes the objective space into grid cells with epsilon as the grid size. It identifies non-dominated grid cells rather than non-dominated solutions and archives at most one solution in each non-dominated grid cell. This solution is the closest one to the best corner of the corresponding grid cell. ε-NSGAII uses crowding distance measure in its selection operator. Kollat and Reed (2005) and Kollat and Reed (2006) suggested that users set the epsilon value (grid size) for each objective to the desired publishable and meaningful precision level of objective functions rather than fine tune them for better algorithm performance. All three objectives of the hydrologic model calibration are unit-less and have the same desired range [0, 1]. The epsilon value of 0.01 is subjectively deemed appropriate for all three objective functions.

A Multi-Algorithm Genetically Adaptive Multiobjective (AMALGAM)

AMALGAM (Vrugt and Robinson 2007) is designed to benefit from the searching ability of multiple MO algorithms (sub-algorithms) simultaneously. In principle, AMALGAM initially divides the number of solutions in each generation between its sub-algorithms and after each generation collects the results from sub-algorithms and archives the result based on the dominance rank and crowding distance in a bounded archive. For new generations, AMALGAM divides the parent solutions between sub-algorithms based on their contribution to the number of archived solutions.

Both ε-NSGAII and AMALGAM have several common parameters with general Genetic Algorithm. Although these parameters can be tuned to seek the best performance of the algorithm, parameter tuning is not the purpose of this chapter. Instead, reasonable parameter values are selected from the literature for these MO algorithms. Table 4-1 shows parameter values that are set based on the studies by Tang et al. (2006) for ε-NSGAII and Zhang et al. (2010) for AMALGAM. Interested readers are referred to Kollat and Reed (2005) and Vrugt and Robinson (2007) for more details about ε-NSGAII and AMALGAM respectively.

Table 4-1 Algorithm parameter values for AMALGAM and ε-NSGAII for solving the MO hydrologic model calibration problem

Parameter name	Parameter Value
Population size	100
Probability of crossover	1.0
Probability of mutation*	1/(number of decision variables)
Crossover Distribution index	15
Mutation Distribution index	20

* This value is often recommended for Genetic Algorithm based optimization algorithm (Deb et al. 2002).

These two MO algorithms are not designed to solve MO problems with very limited computational budgets such as a limit of 1,000 solutions to be evaluated; however, the algorithms had been applied to MO hydrologic model calibration problems with budgets of 10,000 to 15,000; e.g. Tang et al. (2006) for ε-NSGAII and Zhang et al. (2010) for AMALGAM. Therefore, PA-DDS, ε-NSGAII, and AMALGAM are applied to the 3-objective hydrologic model calibration problem with the budget of 10,000 and their results are compared based on statistics and distribution from results of 10 independent trials.

4.3.3 Experiment 3: Solving CEC09 Problems with PA-DDS

In Experiment 3, PA-DDS is compared to all 13 MO algorithms that participated in CEC09 MO competition as described by Zhang et al. (2008) for solving 13 unconstrained MO problems called UF1 to UF13. The CEC09 MO problems have controllable shape Pareto front in the decision space and are made up of seven bi-objective optimization problems (UF1 through UF7), three 3-objective problems (UF8, UF9 and UF10) and three 5-objective problems (UF11, UF12, UF13). Based on CEC09 MO conditions, MO algorithms should keep their settings and parameter values constant for MO problems with the same number of objectives and with the budget of 300,000 and results should be reported for 30 independent trials for each MO problem. Thus, the value of the only parameter of PA-DDS is set to its default value ($r = 0.2$). As a requirement of CEC09, results of each trial must be presented by at most 100, 150 and 800 solutions for 2-, 3- and 5-objective problems, respectively. Since PA-DDS has an unbounded archive, in the Pareto approximate front of each PA-DDS trial, solutions that contribute least to the IGD are removed to respect this requirement. Interested readers are referred to Zhang et al. (2008) for more details about CEC09 MO competition.

4.3.4 Results Comparison Approach

All MO algorithms in this chapter are stochastic search algorithms that show variable performance as the initial random seed changes. Hence, any fair algorithm comparison must consider the distribution of performance metrics rather than a single performance metric assessed using a single trial. In Experiments 1 and 2, each MO algorithm is applied to each MO problem in multiple independent trials. The final result of each trial is assessed by the four performance metrics introduced in section 4.3.4.1. Then, the empirical Cumulative Distribution Function (CDF) of each performance metric is evaluated for each MO algorithm solving each MO problem. CDFs show the probability of equal or better performance at each level of the performance metric. The first degree stochastic dominance concept (Levy 1992) is used in the visual comparison of CDF plots and then the Wilcoxon

rank-sum test (Gibbons and Chakraborti 2003) is applied to measure the statistical significance of the differences in CDFs.

4.3.4.1 MO Performance Comparison Metrics

The solution of a MO problem is a Pareto approximate front in the objective space and the corresponding set of decision variable values. MO performance metrics that assess the quality of a Pareto approximate front by a single number are referred to as unary performance metrics. The unary performance metrics aim to measure the proximity and/or the diversity of points on the Pareto approximate front. Different unary performance metrics measure these two aspects of a Pareto approximate front differently. A detailed list of MO performance metrics can be found in Coello et al. (2007). Results of this chapter are evaluated using various unary performance metrics including Normalized Hypervolume (NHV), Additive epsilon Indicator (ε+ Indicator), Generational Distance (GD), and Inverse Generational Distance (IGD) performance metrics, which are detailed below.

Hypervolume (HV) by Zitzler and Thiele (1998) is a unary performance metric that measures both proximity and diversity of a Pareto approximate front. It measures the volume of the partition of the objective space that is bounded between the Pareto approximate front and a reference point. Based on the definition, HV is a complete unary performance metric in terms of the weak dominance relation, that is, a preferred solution by HV is not weakly dominated by its opponent (Zitzler et al. 2003). As suggested by Deb (2001), in this chapter, HV is calculated in the normalized objective space (equation 3-5) where all results are inside the unit hypercube $[0, 1]^m$ and normalized HV (NHV) values are less than or equal to 1. Higher values of NHV are desirable while lower values of the other three performance metrics are desirable. Therefore, 1-NHV is reported as the performance metric so that lower values indicate improved performance for all four metrics.

The unary performance metric additive epsilon indicator, ε+Indicator (Zitzler et al. 2003) also considers both the diversity and the proximity of a Pareto approximate front by measuring the smallest distance by which an approximate front must be shifted in the objective space to weakly dominate the reference set which is a subset of Pareto optimal front. So, ε+Indicator is a complete performance metric. Its value is sensitive to any point on the approximate front that is farthest from the reference set, and therefore, it can identify if a significant gap exists in the Pareto approximate front.

Generational Distance, GD by Van Veldhuizen and Lamont (1998), is a unary performance metric that measures the average Euclidean distance in the objective space between each point on Pareto

approximate front and its closest pre-identified point on a subset of Pareto optimal front. So, GD measures the proximity of a Pareto approximate front.

GD does not consider the diversity of points on the Pareto approximate front (Deb 2001). Inverse Generational Distance (IGD) was first proposed by Sato et al. (2004) as a performance metric that measures both diversity and proximity of points on the Pareto approximate front. IGD is a unary performance metric that measures the average Euclidean distance in the objective space between each pre-identified point on Pareto optimal front and its closest point on the Pareto approximate front.

The publically available code developed for the CEC09 MOEA competition is used to calculate IGD (see http://dces.essex.ac.uk/staff/qzhang/moeacompetition09.htm). Also provided at the above link are the sets of 1000, 10000, 5000 pre-identified point on Pareto optimal front points for 2-, 3- and 5-objective problems, respectively.

The calculation of all GD, IGD and ε+ Indicator requires a subset of Pareto optimal front that is unknown in Experiment 2. Okabe et al. (2004) noted that in this situation the most desired set of points should replace this set. So, for calculating these performance metrics in Experiment 2, the most desired set of points is created by merging results of all optimization trials of all three algorithms and finding all non-dominated points.

It should be noted that GD and IGD are not complete performance metrics as introduced in Zitzler et al. (2003) with respect to weak dominance relationship. Hence, the better GD or IGD value for a Pareto approximate front compared to an alternative front does not guarantee that the Pareto approximate front is not weakly dominated by the alternative front. However, IGD and GD are popular performance metrics since they measure the overall distance between the Pareto approximate front and a subset of Pareto optimal front; example applications include CEC09 Li and Zhang (2009) and Hadka and Reed (2012) and as such are utilized here along with other performance metrics.

4.3.4.2 Stochastic Dominance

Levy (1992) surveyed the history and application of stochastic dominance concept for comparing two random variables based on their CDFs. Although various degrees of stochastic dominance exist, first degree stochastic dominance has been utilized in some previous optimization algorithm comparison studies, e.g. Mugunthan et al. (2005) and Carrano et al. (2011). Comparing two algorithms, A and B, based on the CDFs, $F_A(x)$ and $F_B(x)$, of a performance metric (x) such that smaller values of x are preferred, A stochastically dominates B if and only if $F_A(x) \geq F_B(x)$ for all possible values of x

(Carrano et al. 2011). When the two CDFs cross each other, first degree stochastic dominance does not hold and cannot identify the preferred algorithm result.

4.3.4.3 Statistical Significance Test

The general form of the Wilcoxon rank-sum test (see Gibbons and Chakraborti, 1992) is used to quantify the significance of the difference in pairwise CDF comparisons. This test has been used for comparing optimization algorithms in Tang et al. (2007) and Hadka and Reed (2012). The null hypothesis of this test assumes that the two samples (i.e., performance metric values of two compared MO algorithms), A and B come from the same population such that $F_A(x) = F_B(x)$ (i.e., no significant difference between algorithms A and B). The two-sided alternative hypothesis only assumes the two samples come from different populations such that $F_A(x) \neq F_B(x)$.

Both the first degree stochastic dominance concept and the Wilcoxon rank-sum test are utilized to identify the preferable algorithm result. Algorithm A result is deemed clearly preferable to algorithm B result when algorithm A stochastically dominates algorithm B, and the corresponding P-value of the two-sided Wilcoxon rank-sum test is smaller than 0.05 (the significance level of the hypothesis test for 95% confidence level). However, when the Wilcoxon rank-sum test suggests that $F_A(x) \neq F_B(x)$ and neither algorithm stochastically dominates the other one (i.e., the CDFs exhibit some crossing behavior), the preferable algorithm result is not clear without additional subjective information (i.e., median metric value critical versus the avoidance of extremely poor metric values being critical).

4.4 Results

Results are presented in separate sub-sections for each of the three numerical experiments. In Experiment 1, the two most effective selection metrics of PA-DDS for solving each MO problem are designated based on the stochastic dominance analysis (visual comparison) of CDFs. Then the Wilcoxon rank-sum test is applied to determine if the differences in the CDFs are significant. In Experiment 2, pairwise comparisons between PA-DDS and each of the ε-NSGAII and AMALGAM algorithms are made based on the CDFs and the Wilcoxon rank-sum test. In Experiment 3, comparison between PA-DDS and CEC09 results is carried out based on the IGD metric. Further assessment of PA-DDS results in Experiment 3 is made based on all four performance metrics that represent the performance of PA-DDS relative to the quality of the reference set used in CEC09.

4.4.1 Results of Experiment 1

PA-DDS with four alternative mating selections (RND, CD, HVC1, and HVC2) is applied to mathematical problems, ZDT4, WFG8, DTLZ6_2D, and DTLZ6_3D as described in section 4.3.1. Each MO problem is solved with two computational budgets, 25,000 and 2,500 solution evaluations. Figure 4-2 shows the empirical CDF plots for additive epsilon indicator value based on 50 trials of ZDT4. Similar visual CDF comparison is made for all MO problems solved in Experiment 1 and based on all four MO performance metrics. The CDF plots are not provided here; instead, the CDF comparisons are summarized as follows. Based on the CDF plots alone, the HVC1 selection metric often generates empirical CDFs that stochastically dominate the other three empirical CDFs. Table 4-2 summarizes a statistical comparison of the empirical CDFs.

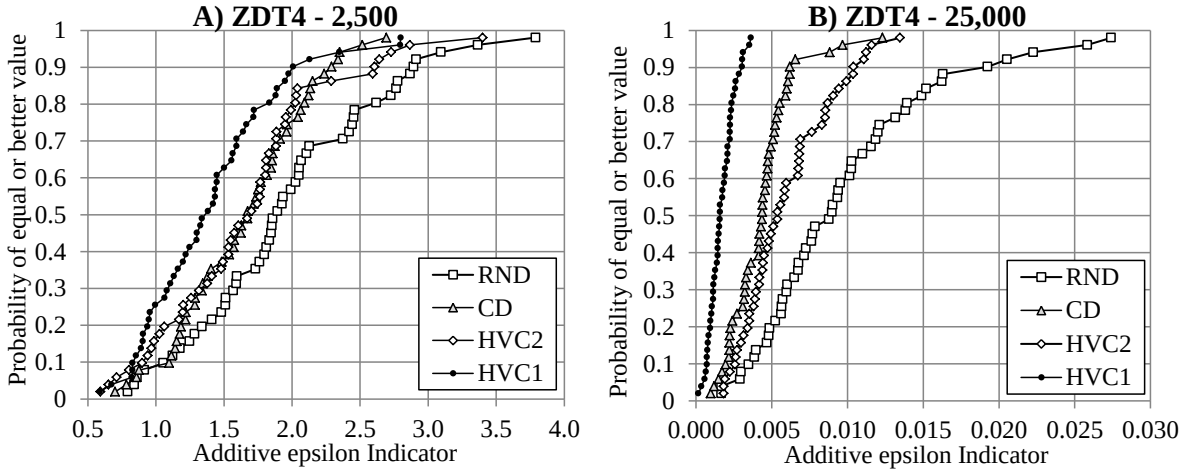

Figure 4-2 ZDT4, empirical CDF plots for Additive epsilon Indicator based on final results of 50 independent trials of the PA-DDS with RND, CD, HVC1 or HVC2 selection. (A) computational budget = 2,500, (B) computational budget = 25,000. Vertical line at 0 represents perfect result.

For each MO problem, Table 4-2 identifies and compares the two most promising selection metrics (referred to as the preferred and alternative metric). The preferred metric is identified as the metric generating the empirical CDF that stochastically dominates (first-degree) the three other empirical CDFs (e.g., HVC1 is preferred metric for ZDT4 based on additive epsilon indicator and budget of 25,000 in Figure 4-2-B). In cases where the preferred metric does not generate an empirical CDF that stochastically dominates all other empirical CDFs, the preferred metric is identified based on the smallest median performance metric value (e.g., HVC1 is preferred metric for ZDT4 based on additive epsilon indicator and budget of 2,500 in Figure 4-2-A). The alternative selection metric in Table 4-2 is essentially the next best alternative to the preferred selection metric as judged based on the empirical CDF comparisons. The alternative metric is identified as the metric generating the

empirical CDF that stochastically dominates (first-degree) two other empirical CDFs. In cases where the alternative metric does not stochastically dominate two other empirical CDFs, the alternative metric is identified as the one showing the highest Wilcoxon rank-sum P-value in comparison with the preferred selection metric empirical CDF. Thus, the P-values reported in Table 4-2 are an upper bound of the three pairwise comparison P-values (preferred metric in comparison with the other three selection metrics).

Table 4-2 Statistical comparison of selection metric performance from Experiment 1 based on the Wilcoxon rank-sum test. P-values are based on the sample size of 50 and compare the results of PA-DDS with two most effective selection metrics (referred to as Preferred and Alternative) designated based on stochastic dominance analysis (visual comparison) of CDF plots for all four performance metrics. **BOLD** names highlight selection metrics that are unambiguously preferred because the preferred metric yields a significantly different empirical CDF (P-value < 0.05) and stochastically dominates the alternative selection metric.

		Performance Metrics and computational budget							
		2500				25000			
MOP		NHV	IGD	GD	ε+ indicator	NHV	IGD	GD	ε_+ indicator
ZDT4	Preferred	HVC1*	HVC1*	HVC1*	HVC1*	**HVC1**	**HVC1**	**HVC1**	**HVC1**
	Alternative	HVC2	HVC2	HVC2	HVC2	CD	CD	CD	CD
	P-value	0.048	0.024	0.042	0.022	<0.001	<0.001	<0.001	<0.001
WFG8	Preferred	**HVC1**	**HVC1**	**HVC1**	**CD**	**HVC1**	**HVC1**	**HVC1**	**CD**
	Alternative	CD	CD	HVC2	HVC1	CD	CD	HVC2	HVC1
	P-value	<0.001	<0.001	<0.001	0.002	<0.001	<0.001	<0.001	0.001
DTLZ6 2D	Preferred	**HVC1**	**HVC1**	**HVC1**	HVC1	HVC1	HVC1	HVC1	HVC2*
	Alternative	HVC2	HVC2	HVC2	HVC2	HVC2	CD	HVC2	HVC1
	P-value	0.011	0.254	0.002	0.151	0.344	0.120	NA**	0.022
DTLZ6 3D	Preferred	**HVC1**	**HVC1**	**HVC1**	**HVC1**	**HVC1**	**HVC1**	HVC1	**HVC1**
	Alternative	HVC2	HVC2	RND	HVC2	HVC2	HVC2	HVC2	HVC2
	P-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	NA**	<0.001

* Despite the significantly different CDFs (P-value < 0.05) the two CDFs cross, so first degree stochastic dominance does not hold and visual CDF comparison required to subjectively identifying the preferred metric.

** GD values for preferred and alternative selection metrics are the same for all practical purposes (differences are all less than 10^{-5}) and hence statistical assessment unnecessary.

Table 4-2 shows that HVC1 is always among the two most promising selection metrics. The stochastic dominance analysis in conjunction with the P-values of the Wilcoxon rank-sum test show that HVC1 is the clearly preferred selection metric in 19 of the 32 pairwise comparisons in Table 4-2. For example, for the WFG8 problem and a budget of 2,500 assessed using the 1-NHV performance metric, the P-value <0.001 is strong evidence that the HVC1 selection metric CDF is significantly different from the CD selection metric CDF. Given that (for this example) the HVC1 selection metric CDF stochastically dominates the CD selection metric CDF, one can conclude that the HVC1 selection metric is unambiguously preferred over CD. Of the remaining 13 pairwise comparisons in

Table 4-2, CD was clearly preferred twice, and in the remaining eleven comparisons, there is no unambiguously preferred selection metric. Importantly, the above findings (favouring HVC1) remain the same regardless of the computational budget considered. A similar observation can be made with regards to the consistency of the findings regardless of the performance metric considered except that the results favouring HVC1 are slightly degraded when only considering the ε+ Indicator. So, HVC1 is designated as the most effective selection metrics for PA-DDS utilized in Experiment 2 and Experiment 3. In DTLZ6 problems solved with 25,000 solution evaluations, GD and IGD values are close to zero showing that PA-DDS performed very well and there is no practical difference between using CD, HVC1 or HVC2 selection metrics when the budget is 25,000 solution evaluations.

4.4.2 Results of Experiment 2

PA-DDS with HVC1-based selection is compared to two popular MO algorithms in the field of water resources management problems, ε-NSGAII and AMALGAM, for solving a hydrologic model calibration problem with three objective functions (introduced in section 4.3.2.1). Each of the three algorithms is applied to this problem 10 times independently, with 10,000 evaluations, and all four MO performance metrics. CDF plots in Figure 4-3 show that CDFs of PA-DDS stochastically dominate those of ε-NSGAII for all four performance metrics and those of AMALGAM for three performance metrics. Moreover, the small P-values in Table 4-3 (all less than 0.0028) are strong indications that these seven differences in the CDFs are statistically significant. Therefore, despite its simplicity compared to AMALGAM and ε-NSGAII, PA-DDS is promising for solving water resources MO problems with a somewhat limited computational budget.

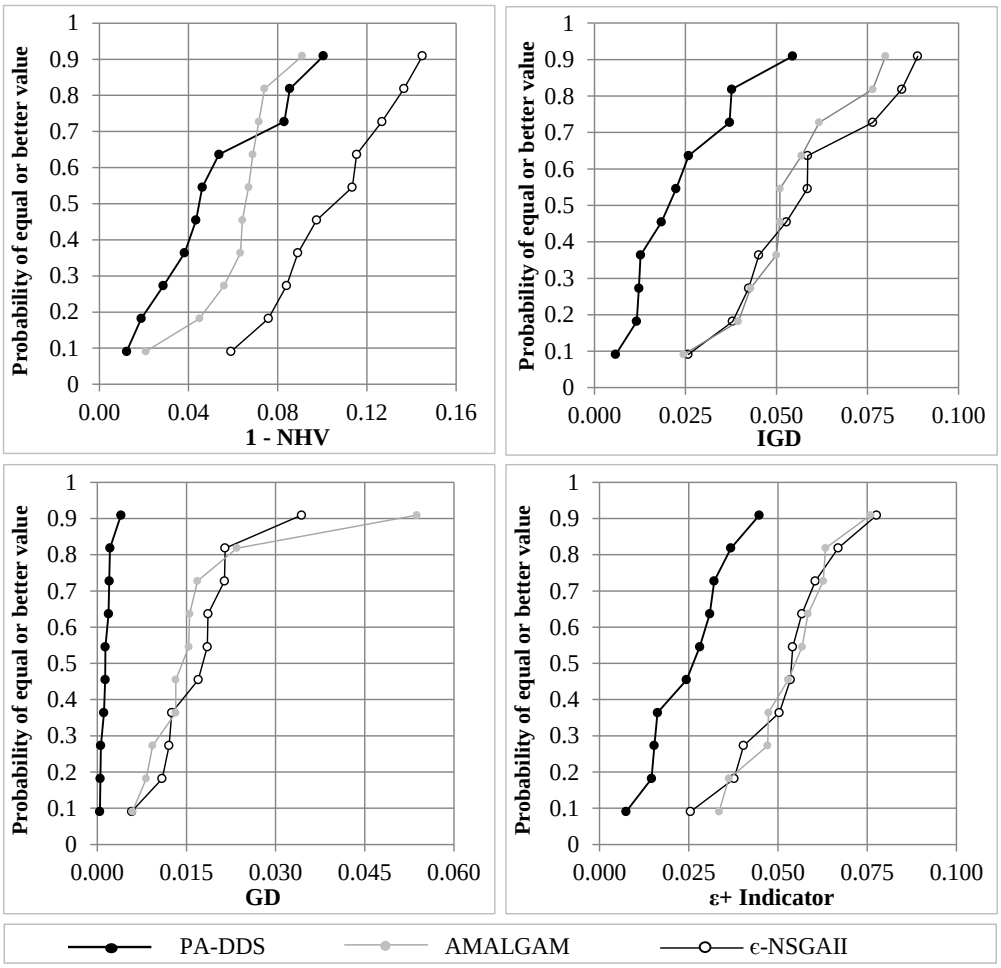

Figure 4-3 Empirical CDFs for 1-NHV, IGD, GD and ε+ Indicator performance metrics based for Experiment 2 (3-objective hydrologic model calibration). Vertical lines at 0 represent ideal results.

Table 4-3 Statistical comparison of PA-DDS, ε-NSGAII and AMALGAM in Experiment 2 based on the two-sided Wilcoxon rank-sum test. P-values are based on a sample size of 10 and compare of PA-DDS with HVC1 selection to each of the ε-NSGAII and AMALGAM algorithms. **BOLD** P-values highlight when PA-DDS is unambiguously preferred because it yields a significantly different empirical CDF (P-value < 0.05) and stochastically dominates the alternative selection metric.

MO Algorithms	Performance Metric			
	NHV	IGD	GD	ε+ indicator
ε-NSGAII	**0.0028**	**0.0017**	**0.0002**	**0.0013**
AMALGAM	0.2730	**0.0022**	**0.0002**	**0.0006**

4.4.3 Results of Experiment 3

Relative PA-DDS performance on CEC09 MO competition test problems is assessed using IGD that was the only performance metric used in CEC09 MO competition, see (Zhang and Suganthan 2008). Other performance metrics are not applied to the CEC09 results to avoid re-ranking original results. Table 4-4 and Table 4-5 are similar to the tabulated results reported by Zhang and Suganthan (2008) that rank all 13 MO algorithms and PA-DDS based on average IGD values over 30 independent trials.

Table 4-4 PA-DDS performance in comparison with 13 MO algorithms based on IGD for solving problems UF1 to UF10 from CEC09 competition. Algorithms sorted by average rank. IGD value is the average of 30 optimization trials each using a computational budget of 300,000. **BOLD** numbers highlight the best IGD value in each test problem.

MO Algorithms	MOP	UF1	UF2	UF3	UF4	UF5	UF6	UF7	UF8	UF9	UF10	Avg. Rank	# of Par	Var. Par?
	# of Obj	2	2	2	2	2	2	2	3	3	3			
MTS	IGD	0.00646	**0.00615**	0.05310	**0.02350**	**0.01489**	0.05917	0.04079	0.11251	0.11442	**0.15306**	4	6	N
	Rank	4	**1**	5	**1**	**1**	2	11	6	8	**1**			
DMOEADD	IGD	0.01038	0.00679	0.03337	0.04268	0.31454	0.06673	0.01032	0.06841	0.04896	0.32211	4.2	6	Y
	Rank	6	3	3	6	11	3	4	2	2	2			
MOEAD	IGD	**0.00435**	0.00679	**0.00742**	0.06385	0.18071	**0.00587**	**0.00444**	**0.05840**	0.07896	0.47415	4.4	8	Y
	Rank	**1**	4	**1**	14	7	**1**	**1**	**1**	4	10			
LiuLi Algorithm	IGD	0.00785	0.01230	0.01497	0.04350	0.16186	0.17555	0.00730	0.08235	0.09391	0.44691	5.5	3	Y
	Rank	5	7	2	8	4	9	2	3	6	9			
GDE3	IGD	0.00534	0.01195	0.10639	0.02650	0.03928	0.25091	0.02522	0.24855	0.08248	0.43326	7.1	3	N
	Rank	2	6	12	2	2	12	9	13	5	8			
AMGA	IGD	0.03588	0.01623	0.06998	0.04062	0.09405	0.12942	0.05707	0.17125	0.18861	0.32418	7.7	9	Y
	Rank	10	10	7	5	3	8	12	8	11	3			
PA-DDS	IGD	0.06240	0.01330	0.12880	0.03220	0.19150	0.22140	0.11048	0.12990	**0.04730**	0.35130	8.2	**1**	N
	Rank	12	9	13	3	8	11	14	7	**1**	4			
MOEADGM	IGD	0.00620	0.00640	0.04900	0.04760	1.79190	0.55630	0.00760	0.24460	0.18780	0.56460	8.3	6	N
	Rank	3	2	4	10	14	14	3	12	10	11			
DECMOSA SQP	IGD	0.07702	0.02834	0.09350	0.03392	0.16713	0.12604	0.02416	0.21583	0.14111	0.36985	8.3	4	Y*
	Rank	13	13	8	4	5	7	8	10	9	6			
NSGAIILS	IGD	0.01153	0.01237	0.10603	0.05840	0.56570	0.31032	0.02132	0.08630	0.07190	0.84468	9.1	8	Y
	Rank	7	8	11	12	13	13	6	4	3	14			
MOEP	IGD	0.05960	0.01890	0.09900	0.04270	0.22450	0.10310	0.01970	0.42300	0.34200	0.36210	9.1	2	Y
	Rank	11	11	9	7	9	6	5	14	14	5			
Clustering MOEA	IGD	0.02990	0.02280	0.05490	0.05850	0.24730	0.08710	0.02230	0.23830	0.29340	0.41110	9.3	5	N
	Rank	9	12	6	13	10	5	7	11	13	7			
OWMO SaDE	IGD	0.01220	0.00810	0.10300	0.05130	0.43030	0.19180	0.05850	0.09450	0.09830	0.74300	9.4	2	Y*
	Rank	8	5	10	11	12	10	13	5	7	13			
OMOEAII	IGD	0.08564	0.03057	0.27141	0.04624	0.16920	0.07338	0.03354	0.19200	0.23179	0.62754	10.4	5	N
	Rank	14	14	14	9	6	4	10	9	12	12			

* Self-adaptation (parameter control) is used

Based on Table 4-4, PA-DDS is ranked 1st in UF9 and 7th out of 14 overall based on average IGD values over 30 independent trials of solving the bi-objective optimization problems UF1 to UF7 and three-objective optimization problems UF8 to UF10. This result suggests that, PA-DDS has comparable performance to recently developed MO algorithms for solving MO problems with sufficient computational budget. The six MO algorithms ranked higher than PA-DDS in Table 4-4 have 3 to 9 algorithm parameters compared to only 1 parameter in PA-DDS. Furthermore, four of

these six MO algorithms use different parameter values for different problem whereas the PA-DDS parameter value is kept constant for MO problems in all dimensions.

Also, results in Table 4-4 show an overall increasing trend in the performance of PA-DDS with the increasing number of objectives. PA-DDS is ranked among the first half of algorithms for all three-objective optimization problems while this happened only once for the seven bi-objective optimization problems. This trend continues for solving five-objective optimization problems where PA-DDS achieved the overall second rank compared to ten other MO algorithms in Table 4-5. As mentioned in section 4.2.1, the exact calculation of HVC1 becomes very time consuming for solving MO problems with five or more objective functions; therefore, the Monte Carlo (MC) estimated version of HVC1 with 10,000 MC points is used in these cases. Based on the results of solving the five-objective optimization problems UF11 to UF13, approximating HVC1 for problems with many objectives has no ill effects on PA-DDS performance.

Table 4-5 PA-DDS performance in comparison with 13 MO algorithms based on IGD for solving problems UF11 to UF13 from CEC09 competition. Algorithms sorted by average rank. IGD value is the average of 30 optimization trials each using a computational budget of 300,000. **BOLD** numbers highlight the best IGD value in each test problem.

MO Algorithms	MOP	UF11	UF12	UF13	Avg. Rank
	# of Obj	5	5	5	
MOEAD	IGD	**0.11032**	**146.78**	1.8489	1.3
	Rank	**1**	**1**	2	
PA-DDS	IGD	0.14102	240.80	**1.7197**	2.6
	Rank	3	4	**1**	
LiuLiAlgorithm	IGD	0.13254	444.82	2.2884	5
	Rank	2	6	7	
NSGAIILS	IGD	0.17520	158.05	3.2323	5
	Rank	4	2	9	
GDE3	IGD	0.23425	202.12	3.2057	5.3
	Rank	5	3	8	
MTS	IGD	0.45505	305.20	1.9097	5.7
	Rank	9	5	3	
DECMOSA-SQP	IGD	0.38304	943.35	1.9178	6.7
	Rank	6	10	4	
DMOEADD	IGD	1.20328	477.65	1.9971	7.3
	Rank	10	7	5	
MOEP	IGD	0.43370	885.89	2.0145	7.7
	Rank	8	9	6	
OWMOSaDE	IGD	0.39510	734.56	3.2573	8.3
	Rank	7	8	10	

Table 4-6 summarizes PA-DDS results assessment for solving UF1 to UF13 by selected statistics of all four performance metrics. It should be noted that GD, IGD and ε+ Indicator assess the results against a subset of Pareto optimal front (reference set); hence, they represent the performance of PA-DDS relative to the ideal or best known set of points. To make NHV consistent with these metrics, it is calculated for the reference set ($NHV_{ref-set}$) and for each trial of PA-DDS the ratio $1-\frac{NHV_{PA-DDS}}{NHV_{ref-set}}$ is calculated. The perfect value of all the metrics is 0. Statistics in Table 4-6 provide future researchers a variety of baseline PA-DDS performance metrics with which to compare.

Table 4-6 Statistics of IGD, GD, ε+ Indicator, and $1-\frac{NHV_{PA-DDS}}{NHV_{ref-set}}$ performance metrics for solving problems UF1 to UF13 from CEC09 competition using PA-DDS. Minimum, average, maximum, and standard deviation based on 30 optimization trials each using 300,000 solution evaluations.

Metric	Stat.	MOP												
		UF1	UF2	UF3	UF4	UF5	UF6	UF7	UF8	UF9	UF10	UF11	UF12	UF13
IGD	Min	0.029	0.009	0.081	0.029	0.133	0.061	0.016	0.085	0.027	0.176	0.128	162.9	1.696
	Avg	0.062	0.013	0.129	0.032	0.192	0.221	0.110	0.130	0.047	0.351	0.141	240.8	1.720
	Max	0.106	0.019	0.224	0.036	0.217	0.414	0.410	0.208	0.159	0.745	0.156	340.4	1.740
	StD	0.022	0.002	0.032	0.002	0.017	0.099	0.139	0.039	0.030	0.205	0.007	45.3	0.011
GD	Min	0.000	0.000	0.003	0.003	0.035	0.004	0.000	0.000	0.001	0.000	0.002	45.9	0.019
	Avg	0.000	0.000	0.008	0.004	0.067	0.037	0.000	0.000	0.001	0.002	0.003	58.3	0.020
	Max	0.001	0.000	0.015	0.004	0.091	0.144	0.000	0.001	0.002	0.004	0.005	75.1	0.023
	StD	0.000	0.000	0.003	0.000	0.015	0.032	0.000	0.000	0.000	0.001	0.001	7.9	0.001
ε_+ Indicator	Min	0.072	0.034	0.137	0.043	0.213	0.139	0.064	0.238	0.063	0.710	0.128	157.8	1.344
	Avg	0.141	0.048	0.222	0.053	0.321	0.371	0.261	0.375	0.139	0.788	0.149	214.3	1.394
	Max	0.258	0.064	0.347	0.076	0.440	0.731	0.753	0.708	0.435	1.000	0.196	273.0	1.449
	StD	0.046	0.009	0.047	0.007	0.057	0.137	0.263	0.204	0.085	0.120	0.013	35.2	0.026
$1-\frac{NHV_{PA-DDS}}{NHV_{ref-set}}$ *	Min	0.057	0.014	0.169	0.131	0.264	0.121	0.046	0.035	0.013	0.507	0.002	0.000**	0.547
	Avg	0.120	0.019	0.216	0.139	0.401	0.263	0.183	0.101	0.034	0.585	0.002	0.000**	0.558
	Max	0.215	0.023	0.329	0.151	0.47	0.448	0.586	0.263	0.155	0.808	0.002	0.000**	0.554
	StD	0.044	0.002	0.034	0.004	0.039	0.069	0.186	0.100	0.030	0.112	0.000	0.000**	0.002

* This equation used to relate PA-DDS normalized hypervolume with the normalized hypervolume of the CEC09 distributed reference set. A value of 0 is best result.

** Reference point (dominated by all results) is too far from the reference set to yield different normalized hypervolumes even though PA-DDS results are much different than reference set. These values of 0 should be disregarded.

4.4.4 PA-DDS Algorithm Runtime Analysis and Limitations

PA-DDS archives all non-dominated solutions found during the search. Therefore, it may archive thousands of solutions when the computational budget (number of solution evaluations) is very large. In general, the number of non-dominated solutions found within a given computational budget is higher in problems with higher numbers of objective functions. As the number of archived solutions grows, the runtime of the dominance test grows. Also, PA-DDS uses a complex selection metric based on hypervolume. Therefore, the runtime of the dominance test and selection metric calculation

in PA-DDS could become a serious limitation of PA-DDS when solving many objective problems with a large number of solution evaluations. In this section, the PA-DDS (with HVC1 selection) algorithm runtime is analyzed across eight different MO test problems with 2, 3, and 5 objective functions and with 10^3, 10^4, 10^5, 10^6 computational budgets (problems ZDT4, DTLZ6_2D, DTLZ6_3D, WFG8 from Experiment 1 and problems UF8, UF11, UF12, UF13 from Experiment 3). In total, the analysis considers three 2-objective problems, two 3-objective problems and three 5-objective problems. Exact HVC1 was calculated for all 2-objective and 3-objective problems while the Monte Carlo-based HVC1 procedure was utilized for the 5-objective problems. A complete evaluation of all objectives for a single solution in these MO problems (referred to as a solution evaluation here) takes from 0.0002 to 0.002 seconds on average on a computer with Intel ® Core™ 2 Quad Q6600 @ 2.40 GHz CPU and with 4 GB of ram.

In some of the 3- and 5-objective test problems solved with the largest computational budget (10^6 solution evaluations), it is observed that PA-DDS runtimes were an order of magnitude (or more) longer in comparison with MO algorithms that use a bounded archive approach. For example, the most computationally expensive trial above required almost 13 days to solve the DTLZ6_3D (3-objective) problem with 10^6 solution evaluations. This relative computational inefficiency is clearly a limitation of PA-DDS for solving very quick to evaluate MO problems with 10^6 or more solution evaluations such that MO algorithms with a bounded archive would perform better than PA-DDS given the same computation time to solve these MO problems.

The two most computationally expensive optimization trials above are the 3-objective DTLZ6_3D and the 5-objective UF13 solved with 10^6 solution evaluations. A single solution evaluation of DTLZ6_3D is three times faster than that for UF13. However, UF13 is solved faster than DTLZ6_3D. The main difference between these two optimization trials is that PA-DDS uses the exact HVC1 calculation for DTLZ6_3D and the Monte-Carlo-based HVC1 for solving UF13. So, it can be hypothesized that in these two optimization trials, the HVC1 metric calculation controls the overall runtime of PA-DDS and that the Monte Carlo-based HVC1 could significantly reduce the computational burden of HVC1 calculation. This hypothesis is tested in a demonstrative example that measures the average runtime of the two most computationally expensive steps of the PA-DDS algorithm, the dominance test and HVC1 calculation on the abovementioned computer for a set of 5000 points on the Pareto optimal front of both UF13 and DTLZ6_3D (5000 is the size of the reference set for the 5-objective problems in CEC 09). PA-DDS checks the dominance for each new candidate solution (PA-DDS does not have the non-dominated sorting step as in AMALGAM and ε-

NSGAII) and the most time consuming dominance test occurs when the new perturbed solution is a new non-dominated solution because it must be compared against all archived solutions to confirm that it is non-dominated. On average, this test with a new non-dominated solution and the above reference sets takes 0.087 and 0.089 seconds for DTLZ6_3D and UF13, respectively. The selection metric calculation for all the above 5000 points takes 8.0 seconds for DTLZ6_3D and 0.6 seconds for UF13. The exact HVC1 calculation for UF13 is too time-consuming (e.g., it takes 28 seconds for 1 out of 5000 points in this 5-dimensional space). So, the Monte-Carlo-based HVC1 substantially reduces the computational burden of HVC1 calculation and the above result of 0.6 seconds for selection metric calculation suggests PA-DDS used in this way is a viable solution approach for MO problems with solution evaluation times of a few seconds or more.

The following analysis demonstrates that the computational efficiency limitation of PA-DDS almost disappears when dealing with MO problems that involve computationally intensive simulation/evaluation models – the precise types of problems PA-DDS was designed for. With the measured average solution evaluation runtimes, the portion of total algorithm runtime that is solely spent on solution evaluations of the abovementioned mathematical MO problems ranges from less than 0.1% to 37% (empty black circles in Figure 4-4). The solution evaluation runtime of these MO problems is artificially increased to assess how the portion of total algorithm runtime changes when more computationally intensive MO problems are solved (solid black dots in Figure 4-4). As can be seen in Figure 4-4, for solving MO problems with a 1-second average solution evaluation time, more than 50% of the total algorithm runtime is spent on solution evaluations. This is more than 90% when the average solution evaluation time is 10 seconds or more. Therefore, it can be concluded that as the solution evaluation runtime increases beyond 10 seconds, total PA-DDS runtime depends almost exclusively on solution evaluation runtime rather than the PA-DDS algorithm operators and procedures such as solution archiving, perturbation, and selection. This conclusion does not apply to problems with more than five objectives or problems that are solved with computational budgets exceeding 10^6 solutions. More analysis under these extreme conditions is needed to determine relative computational expense of PA-DDS operators and procedures.

Experiment 2 is a real example (rather than a test problem) that falls within the bounds of Figure 4-4 and demonstrates how solution evaluation time dominates total PA-DDS runtime. In Experiment 2, 10,000 solution evaluations take 5.6 hours on the abovementioned computer (2 seconds for each solution evaluation). PA-DDS spent 6.0 hours on average to solve the MO problem with the budget of 10,000 and archived 290 solutions. Therefore, PA-DDS operators only account for

7% of the total optimization runtime. This portion is 6% and 4% for ε-NSGAII and AMALGAM, respectively.

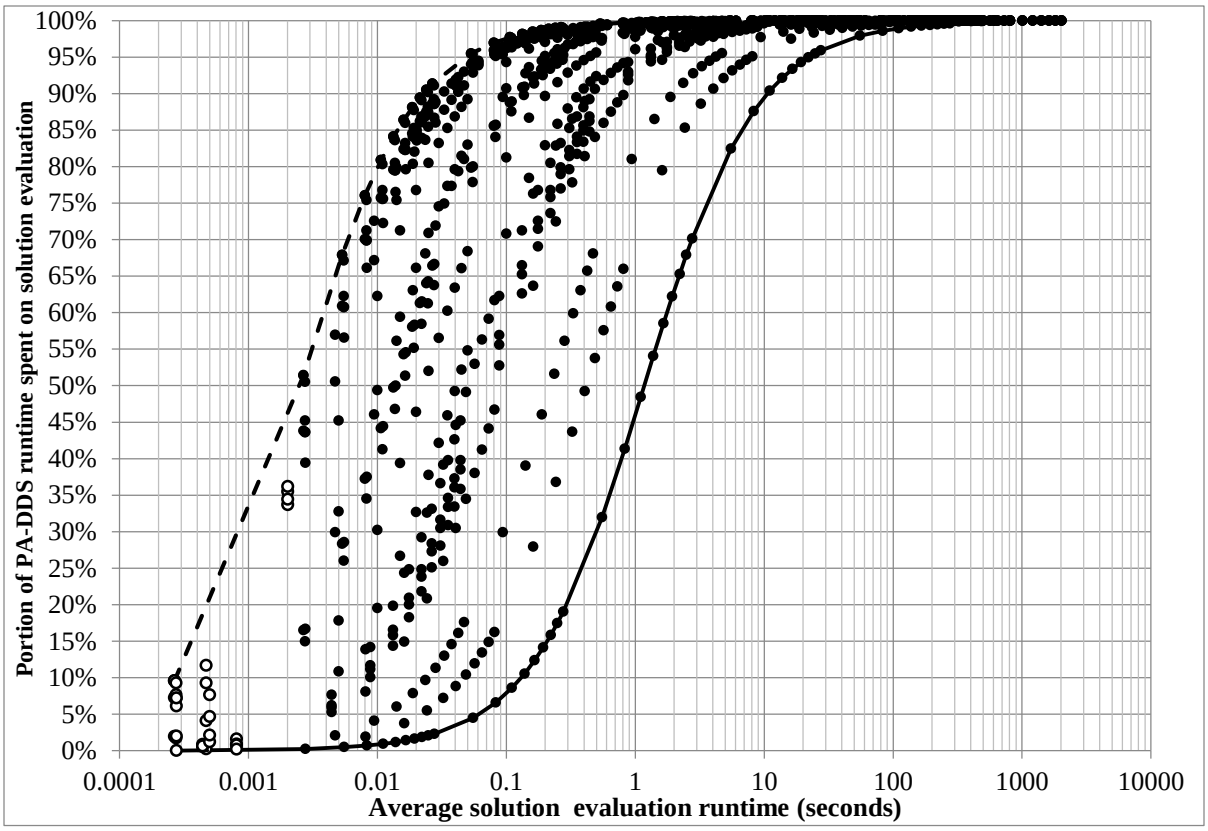

Figure 4-4 Portion of total PA-DDS runtime solely spent on evaluation of objective functions based on solving 10 different MO test problems with 2, 3 and 5 objectives and with four computational budgets of 10^3, 10^4, 10^5, 10^6. White dots represent timings based on actual optimization trials with observed average solution evaluation times while the solid black dots are based on the same optimization trials using artificially increased average solution evaluation times. The solid and dashed lines represent the most and the least algorithmically complex trials: a 3-objective MO problem solved with the budget of 10^6 and a BOP solved with the budget of 10^3, respectively.

4.5 Discussion and Conclusions

Hypervolume contribution (Knowles et al. 2003), referred to as HVC1 in this chapter, was successfully implemented as the selection metric for PA-DDS. Results show that this metric significantly improves the performance of PA-DDS. This improvement does not depend on the computational budget; therefore, this metric is designated as the most effective mating selection metric for PA-DDS for solving general multi-objective optimization problems. Also, for optimization problems with five objectives where exact calculation of this metric is very time consuming, the Monte Carlo approximation of HVC1 performs very well in PA-DDS.

PA-DDS with HVC1 selection was successfully applied to a hydrologic model calibration problem with three objectives. The comparison between PA-DDS and two well-known and recently developed MO algorithms, ε-NSGAII and AMALGAM, in the field of water resources shows that despite its simplicity and parsimony, PA-DDS results are promising relative to these algorithms for a somewhat limited computational budget for solving water resources management problems. PA-DDS with HVC1 selection was also applied to general mathematical MOPs (with up to five objectives) and results show comparable performance of PA-DDS to other recently developed MO algorithms.

The computational cost of PA-DDS operators (solution archiving, selection, and perturbation) with an unbounded archive was demonstrated to be negligible compared to the objective functions evaluation runtime when dealing with computationally intensive multi-objective optimization problems.

Chapter 5

Convex Hull Contribution, a Novel Selection Metric for Convex Multi-objective Optimization Problems: Application to Water Resources Calibration Problems

Summary

A novel selection metric called Convex Hull Contribution (CHC) is introduced multi-objective optimization algorithms applied to multi-objective optimization problems with convex Pareto fronts. CHC identifies the archived non-dominated solutions whose map in objective space form the convex approximation of the Pareto front. Hence, the optimization algorithm can sample solely from these solutions to approximate the convex shape of the Pareto front more accurately.

It is empirically demonstrated that CHC improves the performance of PA-DDS when solving multi-objective optimization problems with convex Pareto fronts. This conclusion is based on the results of solving several benchmark mathematical problems and several water resources model calibration problems with two or three objective functions. The impact of CHC on PA-DDS performance is most evident when the computational budget is somewhat limited. Also, when the computational budget is relatively large, PA-DDS with CHC-based selection for solving several multi-objective model calibration problems performs comparable with two multi-objective optimization algorithms called AMALGAM and ε-NSGAII commonly applied to water resources problems. It is also demonstrated that PA-DDS with CHC-based selection can find acceptable calibrated models by 1,000 solution evaluations (limited budget in this chapter) relative to the results by PA-DDS, AMALGAM and ε-NSGAII with 10,000 solution evaluations.

This chapter is organized as follows. In Section 5.1, the scientific literature of model calibration problems is reviewed to show the convexity of Pareto front in such problems. In section 5.2, CHC is introduced with detailed formulation. In Section 5.3, the numerical experiment to assess the performance PA-DDS with CHC is explained. Finally, results are presented and discussed in Section 5.4.

5.1 Introduction

Simulation is a common practise among engineers to study the behavior of water resources systems such as rainfall-runoff, surface and ground water quantity and quality, and water distribution. Models often have several parameters that need to be calibrated to have simulation results approximate the observed system behavior. For example in hydrologic models the system behaviour can be an output flux such as streamflow or a state variable such as snow water equivalent. In a water distribution systems the output flux can be nodal pressure and state variable can be pipe roughness coefficient in water distribution systems. The calibration task can be conducted in a trial-and-error experiment (manual calibration) or in a simulation-optimization framework (automatic calibration).

Various metrics including Nash Sutcliffe coefficient (Nash and Sutcliffe 1970), absolute relative error, root mean squared error, and bias have been utilized to evaluate the calibration effectiveness as a function of the difference between simulated and observed data. A hydrologic model calibrated to a single output flux (e.g., streamflow) that perfectly simulates the observations, optimizes all simulation error metrics or objective functions simultaneously regardless of how the metric defines the best fit between simulated and observed data. However, when simulations do not perfectly match the observations, the simulation error metrics become conflicting due to the difference in their definition of the best fit between simulated and observed data. Gupta et al. (1998) empirically demonstrated this conflict by calibrating a hydrologic model with a single output flux, streamflow, in two independent single objective optimization problems with two different objective functions. The two problems converged to two different solutions each of which performed well in a portion of streamflow time series that the other solution could not simulate well. Gupta et al. (1998) also noted that conflicting objectives are very likely when different calibration objectives are based on different output fluxes (e.g., streamflow in two different monitoring stations) and/or different state variables (e.g., Nash Sutcliffe for streamflow versus Snow Water Equivalent). So, in general, single objective model calibration to optimize a single simulation error metric would not result in a calibrated model that adequately simulates even a single output flux, and model calibration is inherently a multi-objective optimization task to balance between various simulation error metrics.

It is often desired to achieve a good balance between multiple calibration objectives by marginally sacrificing all objective functions from their individual optima. This desired behavior is easy to achieve for Pareto fronts with a convex shape that is bent in its central region close to the ideal point and generally extends in the tails towards the individual optima. Figure 5-1-A represents a set of non-

dominated points in the objective space of an example bi-objective minimization problem with this desired behaviour. As shown in Figure 5-1-A, this type of Pareto front can be accurately approximated by a convex curve, piecewise linear curve connecting a portion of non-dominated solutions. However, in a calibration problem without this desired behaviour, the convex approximation of the Pareto front may partially or totally miss-represent the Pareto front as in Figure 5-1-B and C respectively.

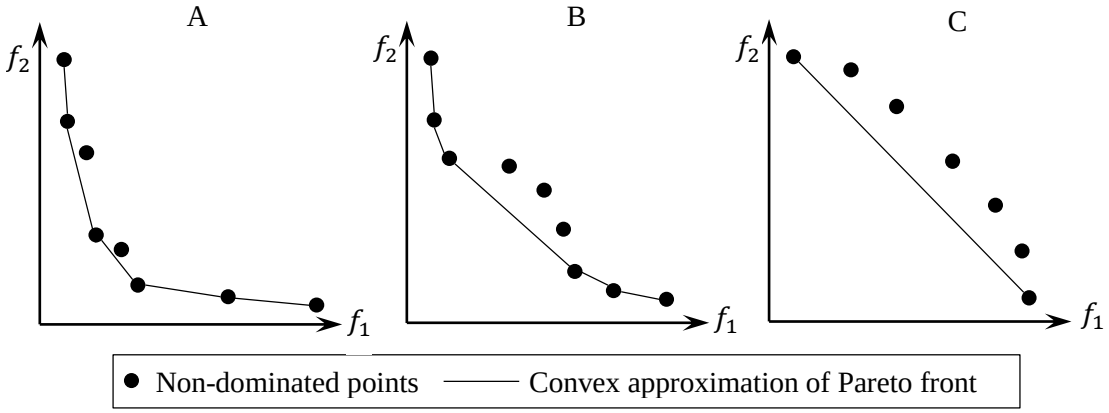

Figure 5-1 a set of non-dominated points in the objective space of three example bi-objective minimization problems and the Convex approximation of Pareto front based on these points. A) the convex curve accurately approximates the Pareto front, B) the convex curve does not some represent a portion of the Pareto front, C) the convex curve does not represent the whole Pareto front

The majority of bi-objective model calibration studies in the literature of water resources result in Pareto fronts that can be accurately approximate by a convex curve, similar to Figure 5-1-A. Representative studies from hydrologic model calibration include Yapo et al. (1998), Madsen (2000), Madsen (2003), Vrugt et al. (2003), Bekele and Nicklow (2007), Confesor and Whittaker (2007), Zhang et al. (2010), and Lee et al. (2011) that showed the Pareto front in the 2-dimensional space. However, Multi-objective model calibration problems are often solved by general multi-objective optimization algorithms that are designed to be insensitive to the shape of the Pareto front. Some studies including Feng et al. (1997), Cococcioni et al. (2007), Chang et al. (2010), empirically demonstrated that designing MO algorithms with the knowledge that the given MO problem has a known, expected, or desired convex PF can be beneficial for solving such problems.

Furthermore, the Pareto front shape represents the quality of a calibrated model. Xia et al. (2002) empirically demonstrated that more advanced or complex models result in smaller Pareto front closer to the ideal point (no error) in the objective space. Also Fenicia et al. (2007) noted that hydrologic model improvement can be identified as Pareto front progressively moves towards the ideal point.

They also showed that when the model simulates the reality very precisely, the Pareto front between the simulation error metrics shrinks and therefore the final solution becomes less dependent to the error metric selection (objective function). Likewise, Lee et al. (2011) studied the impact of using more advanced models on modelling hydrologic events by calibrating a simple and an advanced model to minimize two different simulation error metrics for streamflow. They showed that using the more advanced model shrinks the Pareto front and moves it significantly towards the ideal point.

Kollat et al. (2012) argued that meaningful conflict between error metrics happens in the presence of structural deficiencies in hydrologic models. They calibrated two different hydrologic models (two different levels of model complexity) for 392 catchments across the United States in a four-objective automatic calibration framework. They showed that, with proper precision levels of the objective function values, 55% of MO model calibrations of the less complex model and 80% for the more complex model collapsed to 10 or fewer solutions on the final Pareto front meaning that the conflict between the four objective functions with appropriate precision levels almost disappeared. They also showed that conflicting objectives in other calibration tasks are a sign of some structural error in the hydrologic model. So, even if objective functions with full precision are implemented, good quality hydrologic model calibrations are expected to have a Pareto front with a sharp bend in the middle and extended tails that would disappear if proper precision levels of objective functions are considered.

Popular MO algorithms such as NSGAII (Deb et al. 2002) and SPEA-2 (Zitzler et al. 2001) are designed to be insensitive to the Pareto front shape. Coello (2001) highlighted this characteristic as an advantage of a MO algorithm because the Pareto front shape is often unknown before solving the problem. However, some studies including Feng et al. (1997), Cococcioni et al. (2007), Chang et al. (2010), empirically demonstrated that designing MO algorithms with the knowledge that the given MO problem has a known, expected, or desired convex Pareto front can be beneficial for solving such problems. Feng et al. (1997) introduced a selection operator for MO algorithms specialized for solving bi-objective optimization problems with convex Pareto front and solved the Time-Cost Tradeoff Problem of a construction project that is expected to have a convex Pareto front. After each generation, the best piecewise linear convex polyline (connecting curve) of the population of solutions is formed in the objective space and solutions closer to this convex polyline are given more chance to be selected as the parents for the new generation. Chang et al. (2010) applied this MO algorithm to solve a bi-objective water allocation problem for minimizing water deficit in irrigation and public water sectors and showed that the final Pareto front has a convex shape. Also, Cococcioni et al. (2007) introduced Convex Hull Evolutionary Algorithm (CHEA) for solving a bi-objective

optimization problem where only solutions whose map to the objective space lie on the convex polyline (connecting curve) of non-dominated solutions are desirable while the other non-dominated solutions can be considered less valuable. CHEA replaces the non-dominated sorting operator of NSGA-II with a convex hull sorting operator. After each generation, CHEA starts filling up the parent population with the first ranked solutions from the combined parents and mated solutions. The population size is fixed; therefore, if the number of first ranked solutions is more than the population size, some of them will randomly be removed from the new population. But if the first ranked solutions are fewer than the population size, some second ranked solutions will randomly be added to the parent population. This process continues until the parent population is filled up. Cococcioni et al. (2007) demonstrated that CHEA performs better than NSGA-II for solving a bi-objective optimization problem with convex Pareto front.

In this chapter, a novel selection metric called Convex Hull Contribution (CHC) is introduced for solving MO problems with a known or expected convex Pareto front. CHC is utilized as the selection metric for PA-DDS. It is empirically demonstrated that CHC improves the performance of PA-DDS for solving MO hydrologic model calibration problems that are expected to have convex Pareto front.

5.2 Methodology

In any MO algorithm, selecting and perturbing a previously evaluated solution aims to find a new solution in its neighborhood that can improve the current approximation of the Pareto front. The following discussion shows that perturbing the non-dominated solutions corresponding to the vertices of the convex approximation of Pareto front is likely to be more beneficial than perturbing other non-dominated archived solutions for solving MO problems with a convex Pareto front.

5.2.1 Convex Hull Background

Based on the definition, convex hull of a finite set of given points ($\boldsymbol{Y} \subset \mathbb{R}^{\mathrm{m}}$) in an m-dimensional space is the smallest convex set that contains all the given points Barber et al. (1996). A set is called convex if for each two points inside the set, all points on the line segment between them are inside the set. Figure 5-2 shows the convex hull (shaded area) of a set of given points (empty circles) in a two-dimensional space as the minimal (in area) convex polygon that contains all the given points. In a higher dimensional space, the convex hull forms a minimal (in volume) convex hypervolume that contains all the given points.

Convex hull is bounded by its facets (dashed line segments in Figure 5-2) and all the given points inside the convex hull and on its facets satisfy the inequality (5-1) where $\boldsymbol{A}$ is an $l \times m$ matrix, l is the number of convex hull facets, and m is the number of dimensions.

$$\boldsymbol{Ax} + \boldsymbol{b} \leq \boldsymbol{0} \quad \textbf{(5-1)}$$

Each facet is a portion of a hyper-plane that divides the space into two half-spaces (two sides of the facet), one half-space contains the convex hull and the other one does not. Each row $\boldsymbol{a}_i$ of matrix $\boldsymbol{A}$ is the unit normal vector of facet i and points outward from the convex hull (i.e., arrows in Figure 5-2 that point to the half-space that does not contain the convex hull). The corresponding component of vector $\boldsymbol{b}$ is the facet's offset from origin. Points on each facet i satisfy equation (5-2), points inside the convex hull satisfy inequality (5-3) for all l facets and points outside the convex hull satisfy equation (5-4).

$$\sum_{j=1}^{m} a_{i,j} y_j + b_i = 0 \quad \textbf{(5-2)}$$

$$\sum_{j=1}^{m} a_{i,j} y_j + b_i < 0 \quad \forall i = 1, \dots, l \quad \textbf{(5-3)}$$

$$\sum_{j=1}^{m} a_{i,j} y_j + b_i > 0 \quad \textbf{(5-4)}$$

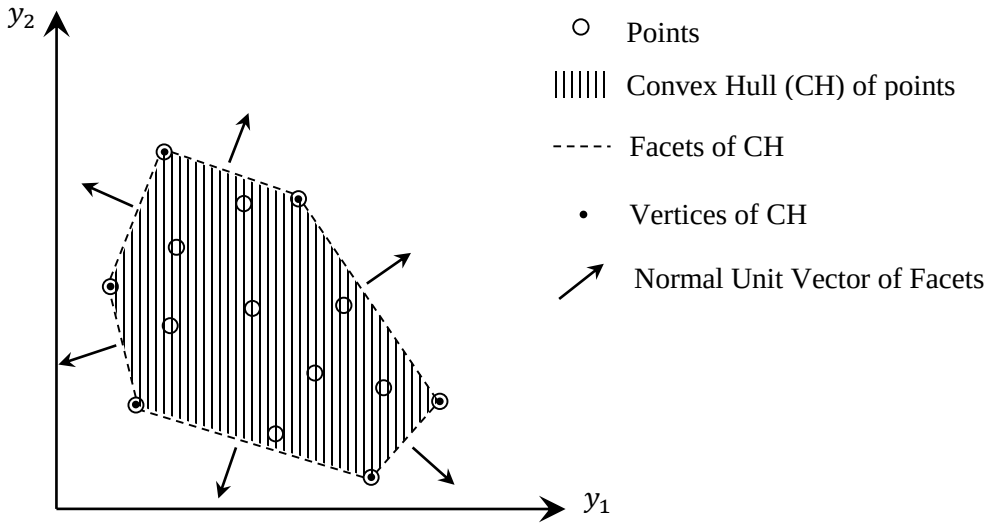

Figure 5-2 Convex Hull of a set of points in a two dimensional space

Some of the given points form the vertices of the convex hull. If any of these points is removed from the set of the given points, the area (hypervolume in more than two-dimensional space) of the convex hull would decrease. However, removing any given point from inside the convex hull would

not change the convex hull. Figure 5-3 shows a set of given non-dominated points (intermediate solutions) in the objective space of an example bi-objective optimization problem with a convex Pareto front and the convex hull of these points (the polygon shaded with vertical lines). The convex hull facets in this figure can be divided into the following two groups:

- Facets whose outward unit normal vector has only non-positive components (solid black line segments in Figure 5-3)
- Other facets (dashed line segments in Figure 5-3).

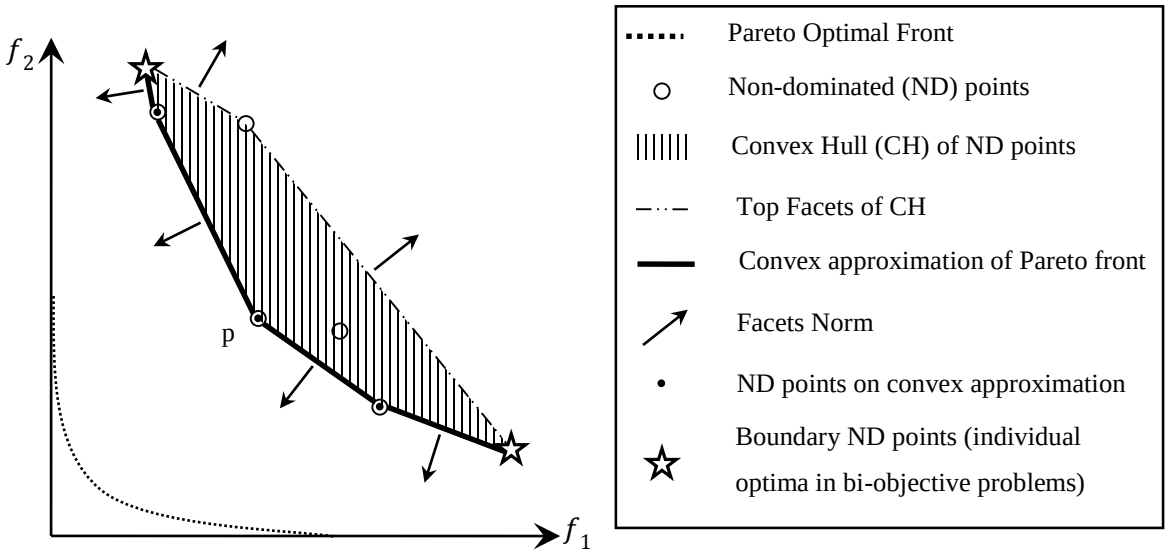

Figure 5-3 Convex Hull of a set of non-dominated points in a normalized bi-objective minimization problem

The former facets construct a bottom boundary for the convex hull while the latter facets form a top boundary for it. As shown in Figure 5-3, the bottom facets form a convex curve or polyline of the given set of non-dominated points. Based on the definition a curve or a function is called convex if it lies below the line segment joining any two points of it. Obviously, none of the given points could be below these facets; otherwise, the convex hull would not contain that point, which is against the definition of convex hull. This curve is called the convex approximation of Pareto front.

Vertices of convex approximation of Pareto front are some of the given non-dominated points. Any new solution that dominates each of these vertices would definitely change the convex approximation of Pareto front towards the Pareto optimal front. Assuming that solution $\boldsymbol{x}^p$ corresponds to one of the vertices of the convex approximation of Pareto front, $\boldsymbol{f}(\boldsymbol{x}^p)$ lies on at least one of the facets (i) so the

statement (5-5) holds. Note that i denotes the i^{th} row of matrix $\boldsymbol{A}$ in (5-5). Also for any new solution $\boldsymbol{x}^q$ that weakly dominates $\boldsymbol{x}^p$, statement (5-6) holds.

$$\exists i\colon \sum_{j=1}^{m} a_{i,j} f_j(\boldsymbol{x}^p) + b_i = 0 \qquad \textbf{(5-5)}$$

$$\boldsymbol{x}^q \preccurlyeq \boldsymbol{x}^p\colon \quad \forall j = 1, \dots, m\colon\ f_j(\boldsymbol{x}^q) \le f_j(\boldsymbol{x}^p) \qquad \textbf{(5-6)}$$

Combining the statements (5-5) and (5-6) and recalling that $\forall j = 1, \dots, m\colon a_{i,j} \le 0$ for $\boldsymbol{f}(\boldsymbol{x}^p)$ that is on a bottom facet, statement (5-9) can be concluded.

$$\exists i\colon a_{i,j} f_j(\boldsymbol{x}^q) \ge a_{i,j} f_j(\boldsymbol{x}^p) \qquad \textbf{(5-7)}$$

$$\Longrightarrow \exists i\colon \sum_{j=1}^{m} a_{i,j} f_j(\boldsymbol{x}^q) + b_i \ge \sum_{j=1}^{m} a_{i,j} f_j(\boldsymbol{x}^p) + b_i = 0 \qquad \textbf{(5-8)}$$

$$\Longrightarrow \exists i\colon \sum_{j=1}^{m} a_{i,j} f_j(\boldsymbol{x}^q) + b_i \ge 0 \qquad \textbf{(5-9)}$$

The equality case can only happen in the situations shown in equation (5-10). Assuming that these are unlikely situations, inequality (5-11) holds for any new solution $\boldsymbol{x}^q$ that weakly dominates a vertex solution $\boldsymbol{x}^p$ of a bottom facet. This means that $\boldsymbol{x}^q$ is in the half-space of facet i that does not contain the convex hull of the current set of non-dominated points. Hence, if such a solution is found, the convex approximation of Pareto front and the convex hull of non-dominated solutions will definitely change to contain the new point. This means that the convex approximation of Pareto front moves towards the Pareto optimal front. However, it is possible that a new solution dominates one of the current non-dominated solutions inside the convex hull or on one of the top facets but itself lies inside the convex hull of the current non-dominated points. In this case, the convex approximation of Pareto front would not change.

$$\begin{aligned} &\forall j = 1, \dots, m\colon f_j(\boldsymbol{x}^q) = f_j(\boldsymbol{x}^p) \Longrightarrow \sum_{j=1}^{m} a_{i,j} f_j(\boldsymbol{x}^q) + b_i = \sum_{j=1}^{m} a_{i,j} f_j(\boldsymbol{x}^p) + b_i = 0 \\ &\text{OR} \\ &\forall j = 1, \dots, m\colon a_{i,j} = 0 \Longrightarrow \sum_{j=1}^{m} a_{i,j} f_j(\boldsymbol{x}^q) + b_i = \sum_{j=1}^{m} a_{i,j} f_j(\boldsymbol{x}^p) + b_i = 0 \end{aligned} \qquad \textbf{(5-10)}$$

$$\exists i\colon \sum_{j=1}^{m} a_{i,j} f_j(\boldsymbol{x}^q) + b_i > 0 \qquad \textbf{(5-11)}$$

Feng et al. (1997) designed a MO algorithm with a fitness assignment scheme that gives the highest selection priority to the vertices of the convex approximation of Pareto front for generating new solutions. This priority decreases as a function of solutions' distance in the objective space to the convex approximation of Pareto front. Also, Cococcioni et al. (2007) introduced the Convex Hull Evolutionary Algorithm (CHEA) that replaces the non-dominated sorting operator of NSGA-II (Deb et al. 2002) with the convex hull sorting operator. The first rank solutions are those that form the convex approximation of Pareto front of the whole population of solutions and the second rank solutions are the ones that form the convex approximation of Pareto front of the remaining solutions and so forth. Therefore, CHEA gives the highest selection chance to vertices of the convex approximation of Pareto front. Both of these studies applied the proposed algorithms only to bi-objective optimization problems and Cococcioni et al. (2007) empirically showed better performance of CHEA compared to NSGA-II.

5.2.2 Convex Hull Contribution (CHC) Selection Metric

In this chapter, a novel selection metric based on the Convex Hull concept is introduced and applied to guide the search of PA-DDS when solving MO problems with a known or expected convex Pareto front. This metric is called Convex Hull Contribution (CHC). PA-DDS utilizes CHC by selecting and perturbing only archived non-dominated solutions that form the convex approximation of Pareto front. So, these solutions are called active archived solutions in contrast to inactive archived solutions (all other archived solutions that lie on the top facets of the convex hull or inside the convex hull). The calculation procedure for CHC is outlined here.

In the first step, the objective space is normalized by equation (3-5) to make CHC unbiased to various scales of objective functions. In Figure 5-4, a set of non-dominated points in the normalized objective space of an example bi-objective optimization problem is shown. The convex hull of all archived solutions, shaded area with vertical lines in Figure 5-4, in the normalized space is formed using the "qhull" code, available at: http://www.qhull.org/ based on the work by Barber et al. (1996). This code can measure the convex hull area (hypervolume in problems with more than two objective functions) and identify the convex hull facets, vertices, and the unit normal outward vector for each facet. Using this information, all archived non-dominated solutions are divided into the following four mutually exclusive groups based on their position in the objective space:

i. Points inside the convex hull

ii. Vertices of top facet only

iii. Vertices of bottom facets only

iv. Vertices in the intersection of top and bottom facets

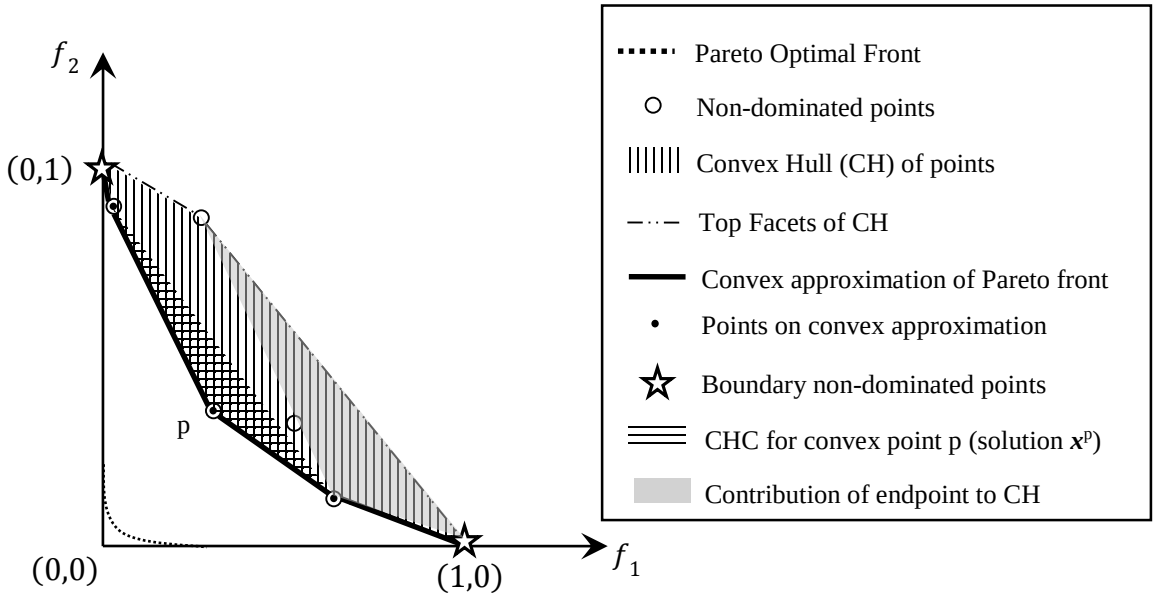

Figure 5-4 Convex Hull Contribution (CHC) calculations for an active non-dominated solution in the normalized objective space of an example bi-objective optimization problem

Solutions in group (i) do not have any contribution to the convex hull (i.e., removing them from the set of non-dominated solutions does not change the convex hull). $\mathrm{CHC} = 0$ is assigned to these solutions and they are removed from the process of calculating CHC for other solutions. Also, $\mathrm{CHC} = 0$ is assigned to solutions in group (ii) but they are not removed from the set of non-dominated solutions in the process of calculating CHC for other solutions because removing them will change the size of convex hull. So, solutions in (i) and (ii) become inactive archived solutions based on CHC. CHC is only calculated for solutions in group (iii). The calculation of CHC for an active solution such as $\boldsymbol{x}^p$ in Figure 5-4 is shown in equation (5-11) where CH is the convex hull area (hypervolume) of a set of non-dominated points ND calculated by the "qhull" code mentioned above. So, CHC is calculated as the difference in the convex hull area (hypervolume) with or without an active solution such as $\boldsymbol{x}^p$ in Figure 5-4.

$$\mathrm{CHC}(\boldsymbol{x}^p) = \mathrm{CH}(ND) - \mathrm{CH}(ND \backslash \{\boldsymbol{f}(\boldsymbol{x}^p)\}) \quad \textbf{(5-12)}$$

In bi-objective optimization problems, solutions in group (iv) are the endpoints of the convex approximation of Pareto front (stars in Figure 5-4). In higher dimensions though, more than just endpoints of the convex approximation of Pareto front are at the intersection of top and bottom facets. These solutions are important for defining the convex approximation of Pareto front; however, their contribution to the size of the convex hull (shaded gray area in Figure 5-4) is not consistent with that for solutions in group (iii). So, instead of calculating CHC for each solution in group (iv) above, the CHC value of its closest active solution in the objective space from group (iii) is assigned to it.

A relative high CHC value for a solution indicates two things. First of all, the solution can be far from its neighboring active solutions in objective space. Therefore, CHC measures the diversity of active solutions. Second, the solution can be far from the convex approximation of the Pareto front based on neighbouring solutions in the direction closer to the ideal point. Therefore, CHC also measures the proximity of active solutions. A good characteristic of CHC is that it avoids extensively sampling from solutions on the extended tails of the Pareto front (i.e., near vertical or near horizontal lines in bi-objective space) because these solutions have negligible contribution to the convex hull area (hypervolume in problems with more than two objective functions). The archiving strategy of ε-NSGAII has a similar characteristic by defining grid cells in the objective space and then archiving at most a single non-dominated solution in each grid cell. If the grid cell sizes (epsilon values) are set properly in ε-NSGAII, the algorithm avoids extensively sampling from the extended tails by approximating these tails with only non-dominated grid cells and hence is capable of approximating an extended tail with only a single grid cell and therefore a single non-dominated solution.

5.3 Numerical Experiments

All experiments to assess the performance of PA-DDS with CHC selection are based on solving MO problems with known or expected convex Pareto front. Asadzadeh and Tolson (accepted) empirically demonstrated that hypervolume contribution (HVC) as in Knowles et al. (2003) is the most effective selection metric for PA-DDS when solving general MO problems. In the first phase of the numerical experiment, PA-DDS_CHC and PA-DDS_HVC are compared for solving seven benchmark mathematical MO problems with known convex Pareto optimal fronts (section 5.3.1(. The only difference between these two versions of PA-DDS is the selection metric. This comparison is based on 50 independent trials of each algorithm and at a low and a high computational budget. Similar to the previous chapter, the budget of 25,000 solution evaluations is considered large enough to achieve high quality results for these MO problems. Also, PA-DDS_CHC is evaluated at a low computational

budget of 2,500 solution evaluations. The comparison of PA-DDS_CHC at the low and high computational budgets would show the dependency of results to the computational budget. This is important for PA-DDS since it is designed to adjust the search strategy to the user-defined computational budget. Thus, if the relative effectiveness of selection metrics depends strongly on the computational budget, a heuristic might be required to switch between various selection metrics based on the budget and/or during the search. Results of this phase are assessed by three MO performance metrics called normalized hypervolume, additive epsilon indicator, and inverse generational distance as described in section 4.3.4.1. The assessment is by stochastic dominance concept as in section 4.3.4.2 and Wilcoxon rank-sum statistical test introduced in section 4.3.4.3.

In the second phase of this numerical experiment, PA-DDS_CHC is applied to six water resources MO model calibration problems that are introduced in section 5.3.2. Again, a low and a high budget of 1,000 and 10,000 solution evaluations are considered. The budget of 10,000 solution evaluations is deemed large enough for these problems based on the hydrologic model calibration literature including Tang et al. (2006), Tolson and Shoemaker (2007) and Zhang et al. (2010) who solved similar problems with this budget. Also, the performance of PA-DDS_CHC with limited budget is studied by solving these problems with 1,000 solution evaluations. At the high budget of 10,000 solution evaluations, PA-DDS_CHC is compared with two popular MO algorithms in the water resources research community: ε-NSGAII by Kollat and Reed (2005) and AMALGAM by Vrugt and Robinson (2007). These two algorithms are discussed in section 4.3.2.2. This comparison is based on 10 independent trials of each of the algorithms. ε-NSGAII and AMALGAM are not designed for and have never been applied with the computational budget as low as 1,000 solution evaluations.

Comparisons with some state-of-the-art meta-modelling based MO algorithms including ParEGO (Knowles 2006) and SMS-EGO (Ponweiser et al. 2008) at a low 1,000 solution evaluations were initially considered. However, these algorithms are specifically designed for solving MO problems with even more strictly limited computational budgets and their algorithm runtime increases drastically with the number of decision variables. For example, Knowles (2006) solve MO test problems with a maximum of 10 decision variables and 260 solution evaluations while Ponweiser et al. (2008) solve MO test problems with a maximum of 6 decision variables and 130 solution evaluations. Initial attempts to apply these MO algorithms (source codes are obtained from the developers of SMS-EGO) to a 5-dimensional MO model calibration case study were unsuccessful due to the computational burden of these MO algorithms (i.e., one optimization trial using only 500 solution evaluations required roughly 4.5 days of serial computation time even though 1000 solutions

could be evaluated independent of the algorithm in less than a minute). So at the low budget of 1,000 solution evaluations, only PA-DDS_CHC and PA-DDS_HVC results are compared. Results at this phase are compared by the approach explained in section 5.3.3.

5.3.1 Mathematical Multi-objective Optimization Problems

Mathematical test problems are designed with special characteristics to challenge MO algorithms. ZDT1 and ZDT4 from Zitzler et al. (2000) are bi-objective optimization problems with controllable number of decision variables. They have convex Pareto optimal fronts and are included in the numerical experiment. Also, Deb et al. (2001) designed a suite of seven MO problems known as DTLZ with user ability to control on both the number of decision variables and the number of objective functions. Zhang et al. (2008) added some difficult features to DTLZ2 by rotating its decision variable space and extending its search space while adding a penalty to solutions outside the original search space. The modified 2- and 3-objective versions of this problem (R2_DTLZ2_M2 and R2_DTLZ2_M3 respectively) with 30 decision variables as in Zhang et al. (2008) are included in the numerical experiment. Original DTLZ2 and its modified versions R2_DTLZ2_M2 and R2_DTLZ2_M3 have non-convex Pareto optimal fronts. Decision variable ranges of R2_DTLZ2_M2 and R2_DTLZ2_M3 are slightly modified in this chapter to make their Pareto fronts convex.

DTLZ2_M2 with 30 decision variables is given by equation (5-13) with the Pareto optimal front at $x_i = 0.5$ for all $i = 2, \dots, 30$ and therefore $1 + g(x_m) = 1$ and $f_1^2 + f_2^2 = 1$ (Deb et al. 2001). As shown in Figure 5-5-A, this Pareto optimal front is one-quarter of the unit circle cantered at the origin of the two-dimensional space (the objective functions in equation (5-13) are the formula of this circle in the polar coordinate system). Therefore, the Pareto optimal front of DTLZ2_M2 is a non-convex curve; however, as shown in Figure 5-5-B, if the range of x_1 is changed from [0, 1] to [2, 3] and the objective functions are moved so that the circle is centered around [1, 1], the Pareto optimal front of the modified R2_DTLZ2_M2 problem becomes convex.

$$
\begin{aligned}
&f_1(\boldsymbol{x}) = (1+g)\cos\left({x_1\pi}/{2}\right)\\
&f_2(\boldsymbol{x}) = (1+g)\sin\left({x_1\pi}/{2}\right)\\
&g = \textstyle\sum_{i=2}^{30}(x_i - 0.5)^2\\
&0 \le x_i \le 1 \ \ for\ all\ \ i = 1, \dots, 30
\end{aligned}
\qquad \textbf{(5-13)}
$$

Similar changes are applied to R2_DTLZ2_M3 with the problem formulation given in equation (5-14). The optimal Pareto front of DTLZ2_M3 corresponds to $x_i = 0.5$ for all $i = 3, \dots, 30$ and

therefore $1 + g(x_m) = 1$ and $f_1^2 + f_2^2 + f_3^2 = 1$ (Deb et al. 2001). With the same logic above, R2_DTLZ2_M3 is modified to have a convex Pareto optimal front by modifying the range of decision variables x_1 from [0, 1] to [-1, 0] and x_2 from [0, 1] to [2, 3].

Li and Zhang (2009) proposed a systematic approach for designing MO problems with controllable difficulty of the PS shape (in the decision space). Three of these test problems that were used in the CEC09 MO competition in Zhang et al. (2008), referred to as UF1, UF2, and UF3, have convex optimal Pareto front and are included in the numerical experiment.

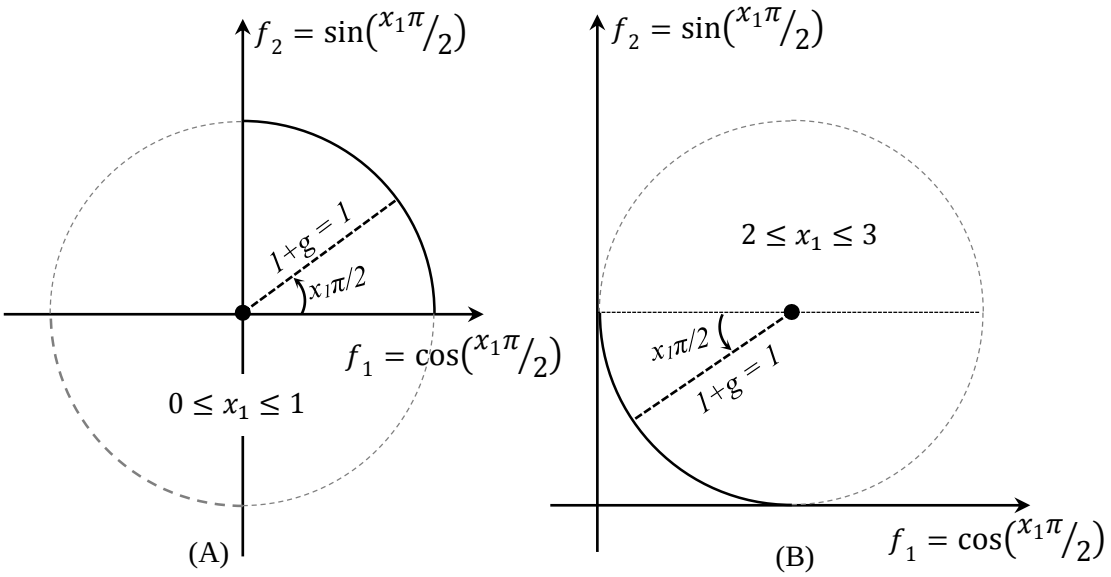

Figure 5-5 The true Pareto front of DTLZ2 with two objective functions

$$f_1(\boldsymbol{x}) = (1+g)\cos(x_1\pi/2)\cos(x_2\pi/2)$$
$$f_2(\boldsymbol{x}) = (1+g)\cos(x_1\pi/2)\sin(x_2\pi/2)$$
$$f_3(\boldsymbol{x}) = (1+g)\sin(x_1\pi/2) \quad \textbf{(5-14)}$$
$$g = \sum_{i=3}^{30}(x_i - 0.5)^2$$
$$0 \le x_i \le 1 \ \ for\ all\ \ i = 1, \ldots, 30$$

5.3.2 MO Water Resources Model Calibration

5.3.2.1 EPANET2, Battle of Water Calibration Networks (BWCN) for C-Town

A special session of the 12th Water Distribution Systems Analysis Symposium in Tucson, Arizona, September 2010 was devoted to a competition on the calibration of a real water distribution system

model for an area called C-Town. The distribution system was modelled using EPANET2 (USEPA, 2002, version 2.00.12) which can simulate hydraulics (pipe flows and nodal pressures) and water quality in pressurized pipe networks. The layout of this water distribution system is shown in Figure 5-6 and it consists of 429 pipes that connect 388 demand nodes, seven tanks and five pumping stations. C-Town is divided into five parts based on the differences in their hourly demand patterns.

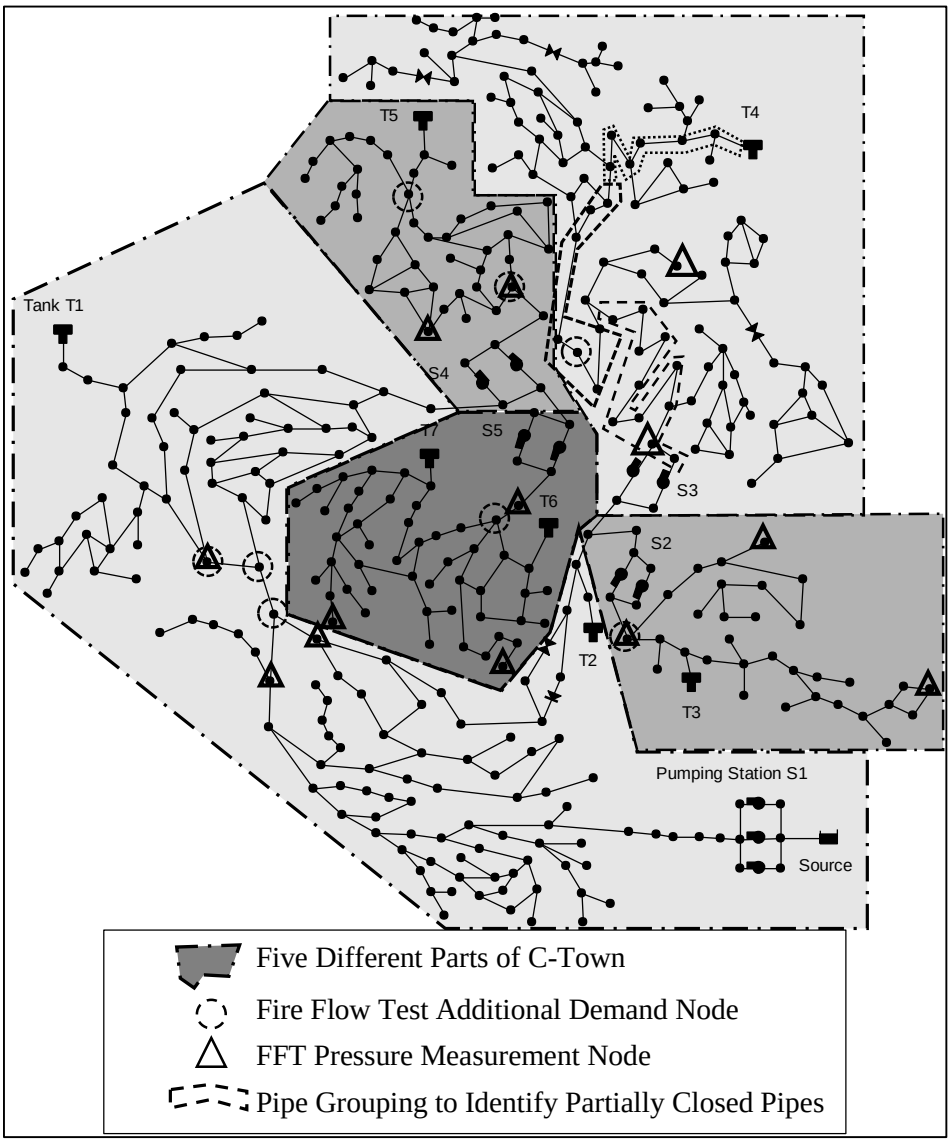

Figure 5-6 C-Town water distribution system layout.

Observations for the calibration of the model consisted of hourly tank levels and pumping station flow rates recorded over one week (168 hours) and fire flow test data. A total of five fire flow tests were conducted (one test in each of the five parts of the network separately) in an evening when the

basic system demands (excluding the observed hydrant demands) are the same as the demands during hour 1 in the one-week operation. Fire flow test (FFT) observations include nodal static pressure (before test being conducted), fire flow test pressure at the same nodes and pumping flow rates at the pumping station corresponding to the area where FFT is conducted. Interested readers are referred to Ostfeld et al. (In press) for more information about the BWCN problem.

Although the EPANET2 model input file and the data provided for BWCN 2010 had some artificial deficiencies included to challenge competitors, the perfect EPANET2 model input file for this network and the perfect one-week observed data and fire flow test observations are known and archived as auxiliary materials for Ostfeld et al. (In press). This perfect input file and observed data set were utilized to formulate a specific multi-objective calibration problem for this thesis. Decision variables in the calibration problem are demand pattern multipliers and pipe roughness coefficients while the objectives are to simulate observed FFT and hour 1 (of 168 hours) tank and flow rate data. Note that this calibration formulation only solves a sub-problem of the original BWCN 2010 based on the approach in Asadzadeh et al. (2010).

Five demand pattern multipliers corresponding to the five parts of C-Town shown in Figure 5-6 at hour 1 of the one-week simulation were calibrated and were assumed to range from 0 to 1. A total of 27 pipe roughness coefficients describing the resistance to flow of all pipes throughout the network were also calibrated. Roughness coefficient ranges were assigned based on guidelines in Table 2.3 in Walski et al. (2003) from Lamont (1981). For complete details on the logic and final ranges of the 27 pipe roughness coefficients see Asadzadeh et al. (2010).

When the 32 decision variable values of a solution are set in the EPANET2 input file, the model is run from hour 0 to 1 in an extended period simulation to simulate the hour 1 data. Then a single period simulation is run to simulate the static pressure measurements. Finally, five independent single period simulations are run with the FFT demands applied to the correct nodes. The objective functions are the simulation quality as measured by sum of squared relative error, SSRE in equation (5-15) as used in BWCN (Ostfeld et al. In press). In this equation, m_i and s_i denote measured and simulated data (flow, pressure, or water level in a tank) respectively for each data point i. The bi-objective optimization problem is to minimize SSRE for hour 1 tank level and flow data and minimize SSRE for FFT data.

$$\text{SSRE} = \sum_{\text{i=1}}^{\text{\#data points}} \left(\frac{s_\text{i} - m_\text{i}}{m_\text{i}}\right)^2 \quad \textbf{(5-15)}$$

5.3.2.2 SAC-SMA: Leaf River

Leaf River is a 1944 km^2 watershed located north of Collins, Mississippi and has been investigated in many single- and multi-objective calibration studies including Sorooshian et al. (1993), Vrugt et al. (2003), Tang et al. (2006), and Kollat et al. (2012). Sorooshian et al. (1993) modelled this watershed in the Sacramento Soil Moisture Accounting (SAC-SMA) to simulate the streamflow (cms) and suggested the calibration of 13 model parameters from the ranges reported in Table 5-1 from Table 2 in Sorooshian et al. (1993). Tang et al. (2006) used the data from July 28th 1952 to September 30th 1954 with the warm-up period from July 28th to September 30th 1952 to calibrate these parameters. This warm-up period appears to be enough for minimizing the impact of the initial values of state variables on the calibration quality (Vrugt et al. 2003). So the calibration objective functions in Tang et al. (2006) represent the streamflow simulation quality from October 1st 1952 to September 30th 1954 in the Leaf River watershed.

The 13 parameters of SAC-SMA are calibrated in a bi-objective optimization formulation. The first objective is to maximize the Nash Sutcliffe coefficient in equation (5-16) that emphasizes peak flows (Gupta et al. 2009). The second objective is to minimize daily root mean squared error, DRMSE in equation (5-17) calculated for Box-Cox transformed flows shown in equation (5-18) where index m_i and s_i denote the measured and simulated streamflow at day number i, z is the transformed value of y and λ is the power parameter of transformation. $\lambda = 0.3$ is used following Tang et al. (2006). The Box-Cox transformation increases the influence of low flow periods on DRMSE (Tang et al. 2006).

$$\mathrm{NS} = 1 - \frac{\sum_{i=1}^{t}(m_i - s_i)^2}{\sum_{i=1}^{t}\left(m_i - \frac{\sum_{i=1}^{t} m_i}{t}\right)^2} \quad \textbf{(5-16)}$$

$$\mathrm{DRMSE} = \sqrt{\sum_{i=1}^{t}(m_i - s_i)^2} \quad \textbf{(5-17)}$$

$$\mathrm{z} = \left((\mathrm{y}+1)^{\lambda} - 1\right)/\lambda \quad \textbf{(5-18)}$$

5.3.2.3 HYMOD: Leaf River

Vrugt et al. (2003) modelled the Leaf River watershed in the HYMOD model with five model parameters. HYMOD is a conceptual rainfall-runoff model that was first introduced by Boyle et al. (2000). The five parameters of HYMOD are calibrated from the ranges reported in Table 5-1. Note that the lower bound of parameter 4 is modified compared to Table 3 in Vrugt et al. (2003). The two calibration objectives for HYMOD are the same as in SAC-SMA calibration problem.

5.3.2.4 SWAT 2000: Town Brook Watershed

The Town Brook case study setup in SWAT 2000 introduced in section 4.3.2.1 is another MO model calibration case study of this chapter. A bi-objective version of this problem is solved to be able to more clearly visualize the comparison of the results (Pareto approximate fronts). Results of the three-objective version of the problem show the objective functions for total phosphorus and for total suspended sediment are highly correlated (correlation coefficient of 77%). So, a bi-objective version of this problem is solved in this chapter by disregarding the objective function for total sediment transport and optimizing the objective function for flow and total phosphorus delivery. Parameter names and ranges can be found in Tables 2 and 3 in Tolson and Shoemaker (2007).

5.3.2.5 SWAT2003: Mahantango Creek Experimental Watershed (MCEW)

MCEW is a 420km^2 benchmark research watershed located in central Pennsylvania. Zhang et al. (2010) calibrated 16 parameters (reported in Table 5-1 from table 2 in Zhang et al. 2010) of the SWAT2003 model for this watershed in a bi-objective optimization problem to maximize the Nash Sutcliffe coefficient for the daily streamflow at two monitoring stations named FD36 and WE38. The simulation period is from Jan. 1st 1995 to Dec. 31st 1998. The calibration period (for objective function calculation) is from Jan. 1st 1997 to Dec. 31st 1998 (personal communication, Zhang, X.).

Table 5-1 Parameter name and range, lower bound (lb) to upper bound (ub) used for the hydrologic model calibration problems SAC-SMA, HYMOD, and SWAT2003 case studies.

Model Name	Name	lb	ub	Name	lb	ub
SAC-SMA	1- $UZTWM$	1.00	150.00	8- $LZTWM$	1.00	500.00
	2- $UZFWM$	1.00	150.00	9- $LZFSM$	1.00	1000.00
	3- UZK	0.10	0.50	10- $LZFPM$	1.00	1000.00
	4- $PCTIM$	0.00	0.10	11- $LZSK$	0.01	0.25
	5- $ADIMP$	0.00	0.40	12- $LZPK$	0.0001	0.025
	6- $ZPERC$	1.00	250.00	13- $PFREE$	0.00	0.10
	7- $REXP$	0.00	5.00			
HYMOD	1- $C_{max}\ (mm)$	0.00	500.00	4- $R_s\ (day)$	0.00001*	0.10
	2- b_{exp}	0.10	2.00	5- $R_q\ (day)$	0.10	0.99
	3- $Alpha$	0.10	0.99	-	-	-
MCEW	1- $SFTMP$	0.0	5.0	9- $ALPHA\_BF$	0.1	1.0
	2- $SMTMP$	0.0	5.0	10- $GWQMN$	0.0	5000.0
	3- $SMFMX$	0.0	10.0	11- GW_REVAP	0.02	0.2
	4- $SMFMN$	0.0	10.0	12- $REVAPMN$	0.0	500.0
	5- $TIMP$	0.01	1.0	13- $RCHRG\_DP$	0.0	1.0
	6- $ESCO$	0.01	0.99	14- $CN2\_f$**	80%	120%
	7- $SURLAG$	0.01	9.99	15- CH_K2	-0.001	150
	8- GW_DELAY	1.0	50.0	16- AWC_f**	80%	120%

* In Vrugt et al. (2003) the lower bound of R_s=0; however, R_s=0 is infeasible in HYMOD.
** Decision variable value is multiplied by the default value of the corresponding SWAT parameter.

5.3.3 Results Comparison Approach

Results of the numerical experiments in this work are assessed by the MO performance metrics normalized hypervolume $(1 - \mathrm{NHV})$, additive epsilon indicator $(\varepsilon + \mathrm{Indicator})$ and inverse generational distance (IGD) as introduced in section 4.3.4. The stochastic dominance concept introduced in section 4.3.4.2 and the Wilcoxon rank-sum test introduced in section 4.3.4.3 are then used to statistically compare performance metrics for each MO algorithm. Furthermore, as explained in section 5.3.3.1, results of the second phase of the numerical experiment are analyzed further by identifying and comparing the best and the worst performed trial of each algorithm. Also, results of the second phase are compared for detecting practical differences between calibrated models by different MO algorithms (see section 5.3.3.2).

5.3.3.1 The Best and Worst Pareto Approximate Front Comparison

In the second phase of the numerical experiments, multiple trials (10 or 50) of each MO algorithm applied to each MO calibration problem are ranked based on each of the three MO performance metrics. The sum of the three ranks is calculated for each MO algorithm and each calibration problem and the trials with the highest and the lowest sum of ranks are selected as the best and worst trials for each MO algorithm, respectively. Pareto approximate fronts of the best trials and the worst in each MO problem are visually compared.

5.3.3.2 Calibrated Model Simulations Comparison

For each MO model calibration problem solved in the second phase of the numerical experiments, the closest solution on the Pareto approximate front to the ideal point in objective space is selected as a representative from the best and worst trials of each MO algorithm. Selected solutions from the best performed MO trials are compared to each other and those from the worst performed MO trials are compared to each other based on their objective function values. This comparison demonstrates the performance range one would expect from all MO algorithms implemented in this chapter. Also, the maximum absolute simulation error and corresponding absolute relative error (based on a daily time step) are compared for the selected solution from PA-DDS_CHC_1000 and the overall closest solution to the ideal point from each of the best and worst performed optimization trials. This comparison provides a meaningful measure for the benefits one could expect by solving these calibration problems with an optimization algorithm other than PA-DDS_CHC_1000 and with an order of magnitude larger budget.

5.4 Results

Results are presented in separate sub-sections for each of the two phases of the numerical experiment. In section 5.4.1, the performance of PA-DDS with Convex Hull Contribution selection (PA-DDS_CHC) and Hypervolume Contribution (PA-DDS_HVC) selection metrics are compared in solving seven mathematical MO problems. This comparison is based on the first degree stochastic dominance and Wilcoxon rank-sum test. In section 5.4.2, pair-wise comparisons are made between PA-DDS_CHC and each of the ε-NSGAII and AMALGAM algorithms applied to six MO model calibration problems with 10,000 solution evaluations. Also, the comparison between PA-DDS_CHC and PA-DDS_HVC at the budget of 1,000 solution evaluations is presented here. In section 5.4.3, the selected solution from PA-DDS_CHC_1000 and an alternative solution which is deemed to be a representative of calibrated models by 10,000 solution evaluations are compared based on some statistics of the simulation error (difference between simulated and observed data).

5.4.1 CHC versus HVC for Mathematical Problems with Convex Pareto Front

PA-DDS_CHC and PA-DDS_HVC are applied to ZDT1, ZDT4, UF1, UF2, UF3, R2_DTLZ2_2D bi-objective problems and R2_DTLZ2_3D with three objective functions introduced in section 5.3.1. Figure 5-7 shows a sample empirical CDF plot for ε+ Indicator values, based on 50 independent trials for solving UF3. Based on this figure, at the limited budget of 2,500 solution evaluations, PA-DDS_CHC stochastically dominates PA-DDS_HVC since at any level of ε+ Indicator, CHC_2500 has higher probability of achieving better ε+ Indicator values. But PA-DDS_CHC_25000 does not stochastically dominate PA-DDS_HVC_25000 due to the crossing behavior of the CDF plots. In this case, the preferred selection metric (CHC) is identified based on the smallest median of the performance metric values.

Similar visual CDF comparison is made for all MO test problems and based on all three MO performance metrics. CDF plots are not provided here; instead, a more concise and informative analysis of the results is provided in Table 5-2 by the two-sided Wilcoxon rank-sum test that statistically measures the significance of the difference in the pairwise comparisons of PA-DDS_CHC and PA-DDS_HVC for each MO test problem. Table 5-2 shows that, PA-DDS_CHC is preferred to PA-DDS_HVC in 19 out of 21 comparisons at the low budget and 13 out of 21 comparisons at the high budget based only on a better median of the performance metric (**BOLD** numbers in Table 5-2). The first degree stochastic dominance (asterisks in Table 5-2) analysis in conjunction with the P-value<0.05 in the Wilcoxon rank-sum test show that at the low budget, PA-DDS_CHC is clearly

preferred to PA-DDS_HVC in 10 out of 21 pairwise comparisons but no clear preference for PA-DDS_HVC. At the high budget, PA-DDS_CHC is clearly preferred in 7 out of 21 comparisons and PA-DDS_HVC is clearly preferred in 5 out of 21 comparisons. Hence, CHC is preferred to HVC as the most effective selection metric for PA-DDS when the computational budget is relatively limited. However, this preference is not very clear when the computational budget is relatively high.

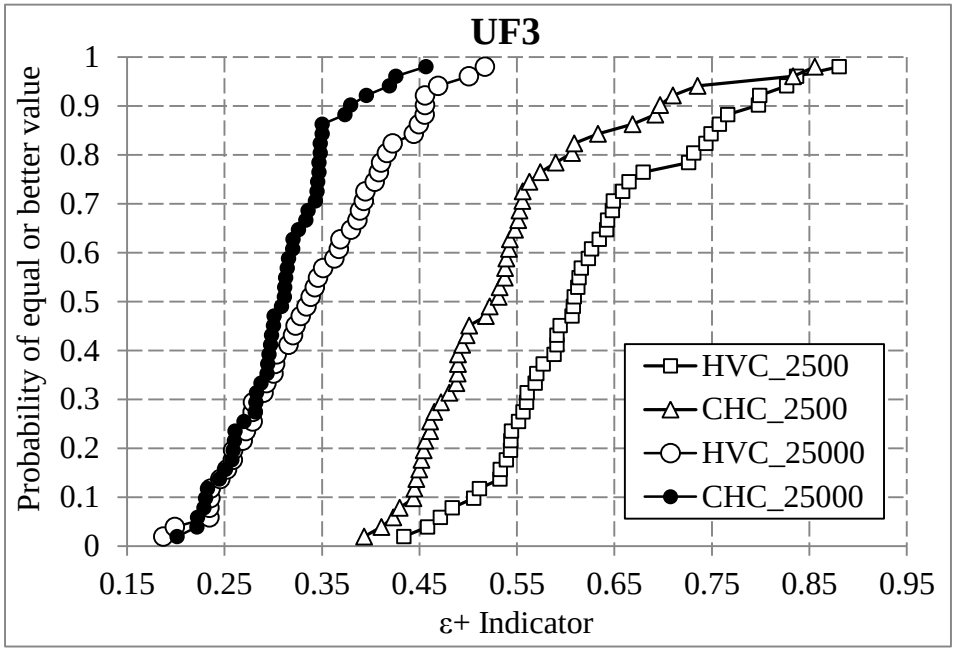

Figure 5-7 UF3, empirical CDF plot based on ε+ Indicator and final results (Pareto approximate fronts) of 50 independent trials of PA-DDS_CHC and PA-DDS_HVC and with computational budget of 2,500 and 25,000 solution evaluations. Vertical line at 0 represents perfect result.

Table 5-2 Statistical comparison of selection metric performance for solving seven mathematical problems based on the two-sided Wilcoxon rank-sum test. Numbers are P-values based on the sample size of 50 and compare the results of PA-DDS_CHC to PA-DDS_HVC. P-value<0.05 with an asterisk shows clear preference.

	Performance Metrics and computational budget					
	2,500			25,000		
MOP	NHV	ε+ indicator	IGD	NHV	ε_+ indicator	IGD
ZDT1	**<0.001***	**<0.001***	**<0.001***	**0.006**	<0.001*	0.002
ZDT4	**0.248**	**0.303**	**0.290**	**<0.001***	**0.003***	**<0.001***
UF1	**0.141**	0.738	0.224	0.053	0.012*	0.069
UF2	**<0.001***	**0.240**	**0.904**	<0.001*	<0.001*	<0.001*
UF3	**<0.001***	**<0.001***	**<0.001***	**0.118**	**0.053**	**0.034**
R2_DTLZ2_2D	**<0.001***	**<0.001***	**<0.001***	**<0.001***	**0.057**	**<0.001***
R2_DTLZ2_3D	**0.035**	**0.584**	**0.234**	**<0.001***	**0.299**	**<0.001***

BOLD P-values highlight the incidents that PA-DDS_CHC is preferred to PA-DDS_HVC based only on better median of the performance metric.
* Preferred algorithm stochastically dominates the alternative algorithm.

5.4.2 PA-DDS Performance Assessment in Water Resources Calibration Problems

PA-DDS_CHC is compared to AMALGAM and ε-NSGAII for solving six MO hydrologic model calibration problems introduced in sections 5.3.2 with 10,000 solution evaluations. The epsilon value for ε-NSGAII is set to 0.0001 for the C-town problem and 0.01 for other five calibration problems in this chapter. These epsilon values are selected based on the meaningful precision level of the objective function (error metrics). Table 5-3 represents the result assessment summary of this comparison based on the first degree stochastic dominance concept and the Wilcoxon rank-sum test. Bold numbers show that, PA-DDS_CHC is preferred (based only on a better median value of the performance metric, regardless of P-value) to ε-NSGAII in all 17 out of 18 comparisons 15 of which are clear preference (stochastic dominance and statistically different performance metric values). Also PA-DDS_CHC is preferred to AMALGAM in 10 out of 18 comparisons 1 of which is clear preference and AMALGAM is clearly preferred to PA-DDS_CHC in 6 out of 18 comparisons.

Table 5-3 Statistical analysis of results comparing PA-DDS with CHC selection to ε-NSGAII and AMALGAM at the budget of 10,000. P-values from the two-sided Wilcoxon rank-sum test are based on a sample size of 10. P-value<0.05 with an asterisk shows clear preference.

		Performance Metrics		
MOP	MO Algorithm	1-NHV	ε_+ indicator	IGD
EPANET2, C-Town	AMALGAM	0.009*	0.017*	0.025*
	ε-NSGAII	**0.002***	**0.038***	**0.045***
SWAT 2000, Town Brook 3D	AMALGAM	**0.076**	**<0.001***	**0.212**
	ε-NSGAII	**<0.001***	**0.001***	**<0.001***
SWAT 2000, Town Brook 2D	AMALGAM	**0.064***	**0.054***	**0.212**
	ε-NSGAII	**0.002***	**0.006***	**0.004***
HYMOD, Leaf River	AMALGAM	<0.001*	<0.001*	<0.001*
	ε-NSGAII	**0.031**	1.000	**<0.001***
SAC-SMA, Leaf River	AMALGAM	**0.910**	**0.791**	**0.970**
	ε-NSGAII	**0.026***	**0.162**	**0.011***
SWAT2003, MCEW	AMALGAM	0.472	0.0385	**0.734**
	ε-NSGAII	**<0.001***	**<0.001***	**<0.001***

BOLD P-values highlight comparisons where PA-DDS with CHC is preferred to alternative algorithm based only on better median of the performance metric.
* Preferred algorithm stochastically dominates the alternative algorithm.

Also, PA-DDS_CHC and PA-DDS_HVC are compared by solving these problems with 1,000 solution evaluations. Table 5-4 represents the summary of this comparison based on the first degree stochastic dominance concept and the Wilcoxon rank-sum test. Results show that, PA-DDS_CHC is

preferred to PA-DDS_HVC in 17 out of 18 comparisons 4 of which represent clear preference, i.e. P-value<0.05 and stochastic dominance.

Table 5-4 Statistical analysis of results comparing CHC and HVC selections for PA-DDS at the budget of 1,000. P-values from the two-sided Wilcoxon rank-sum test are based on a sample size of 50 for C-Town and Town Brook case studies and 10 for other problems. P-value<0.05 with an asterisk shows clear preference.

	Performance Metrics		
MOP	1-NHV	ε+ indicator	IGD
EPANET2 C-Town	**0.182***	**0.863**	**0.780**
SWAT 2000 Town Brook 3D	**0.354**	0.497	**0.060**
SWAT 2000 Town Brook 2D	**<0.001***	**0.004**	**0.015***
HYMOD Leaf River	**0.140**	**0.121***	**0.064**
SAC-SMA Leaf River	**0.521***	**0.910**	**0.910**
SWAT2003 MCEW	**0.026***	**0.054***	**0.045***

BOLD P-values highlight the incidents that PA-DDS_CHC is preferred to PA-DDS_HVC based on stochastic dominance or better median of the performance metric.
* Preferred algorithm stochastically dominates the alternative algorithm.

Figure 5-8-A visually compares the Pareto approximate fronts of the best case performance (see section 5.3.3.1) of PA-DDS_CHC and PA-DDS_HVC selections at the budget of 1,000 solution evaluations and PA-DDS_CHC, AMALGAM and ε-NSGAII at the budget of 10,000 solution evaluations in the C-Town problem. The Pareto approximate fronts of the worst case performance are presented in Figure 5-8-B. At the high budget, all MO algorithms perform very well in identifying solutions very close to the ideal solution (no error) in their best and worst trials. However, at the low budget, PA-DDS_CHC performs better than PA-DDS_HVC since its best and worst trials dominate and almost dominate the ones by PA-DDS_HVC, respectively. Similar comparisons are shown for other four bi-objective calibration problems in Figure 5-9 to Figure 5-12. The visual comparison for Town Brook with three objectives is not presented due to the difficulty in interpreting multiple sets of points in a 3D figure.

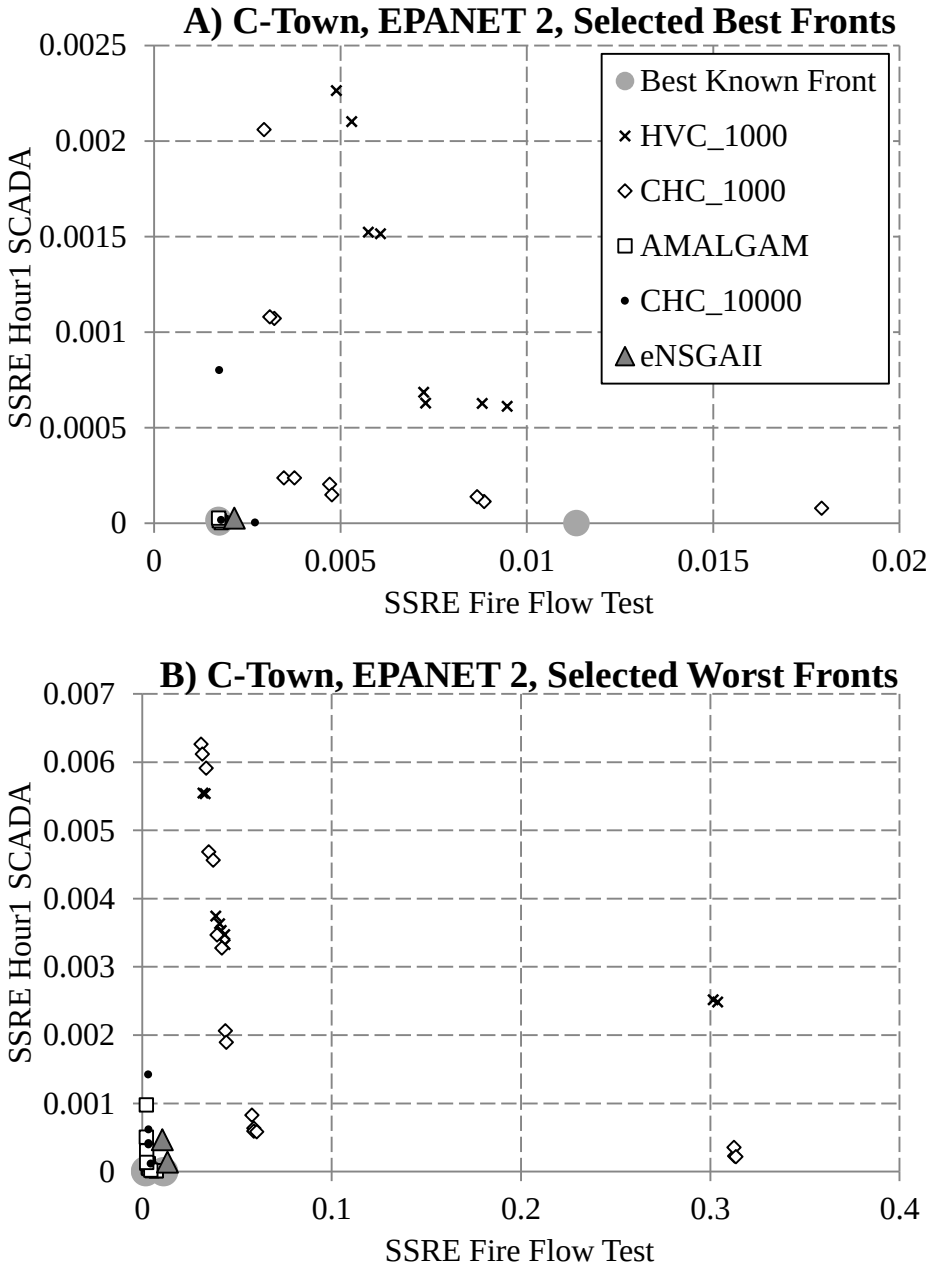

Figure 5-8 EPANET2 calibration problem, Pareto approximate fronts of A) selected best trial and B) selected worst trial of PA-DDS with CHC and HVC selections with 1,000 solution evaluations and PA-DDS with CHC selection, ε-NSGAII and AMALGAM with 10,000 solution evaluations. The different scale of axes is due to the significant difference in the quality of results in A and B

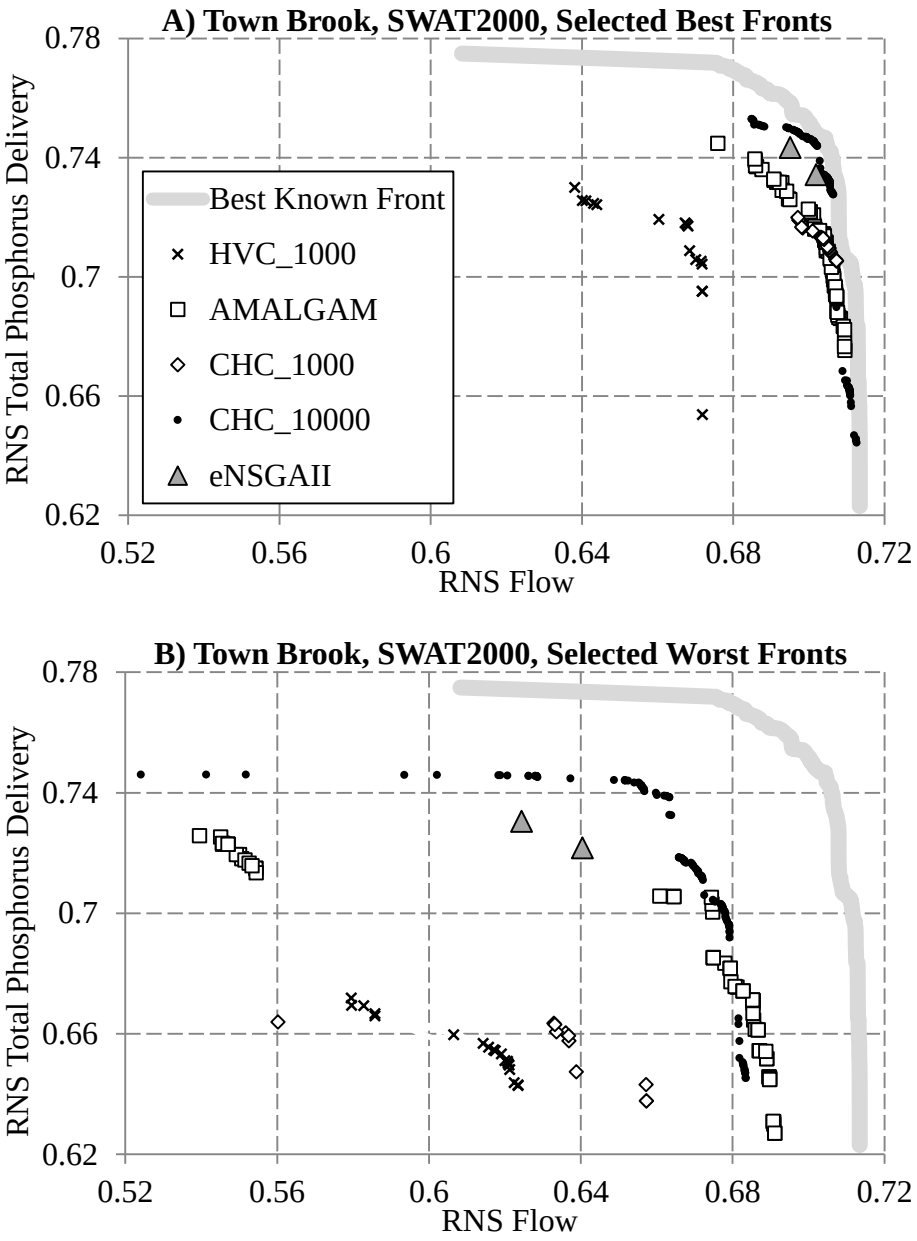

Figure 5-9 SWAT 2000 calibration problem, Pareto approximate fronts of A) selected best trial and B) selected worst trial of PA-DDS with CHC and HVC selections with 1,000 solution evaluations and PA-DDS with CHC selection, ε-NSGAII and AMALGAM with 10,000 solution evaluations.

As shown in Figure 5-9-A (the best case performance), in the Town Brook bi-objective calibration problem, all MO algorithms perform very well and very similar except for PA-DDS_HVC_1000 that results in a significantly worse front. In this case chapter, PA-DDS_CHC can reach very high quality results with 1,000 solution evaluations compared to other MO algorithms with 10,000 solution

evaluations. In Figure 5-9-B (the worst case performance) PA-DDS_CHC and AMALGAM at the high budget perform similar and better than others. ε-NSGAII results in noticeably few solutions. This behavior is expected by ε-NSGAII that archives only a single solution in each grid (epsilon) cell in the objective space (see section 4.3.2.2).

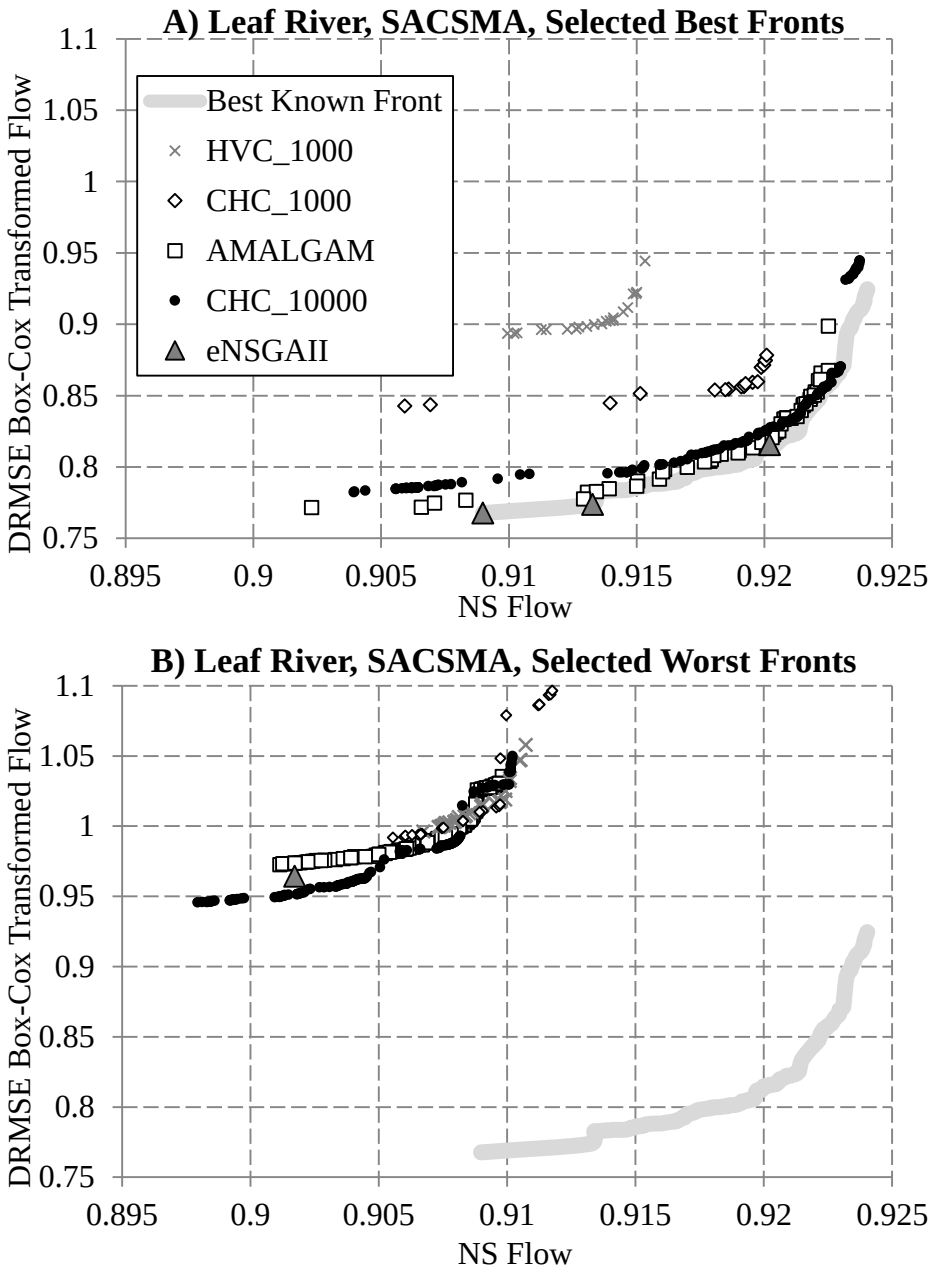

Figure 5-10 SAC-SMA calibration problem, Pareto approximate fronts of A) selected best trial and B) selected worst trial of PA-DDS with CHC and HVC selections with 1,000 solution evaluations and PA-DDS with CHC selection, ε-NSGAII and AMALGAM with 10,000 solution evaluations.

In Figure 5-10 and at the high budget, PA-DDS_CHC performs comparable to AMALGAM in each of the best and worst performed trials of solving SAC-SMA problem. In Figure 5-10-A, the best case performance of PA-DDS_CHC dominates that of PA-DDS_HVC at the low budget; although, this difference is practically negligible (see section 5.4.3). In Figure 5-10-B, all algorithms perform similar but ε-NSGAII results in only a single solution.

In HYMOD (Figure 5-11), all trials at the low and high budget result in Pareto approximate front very close to the best known Pareto front. However, in the worst case performance of PA-DDS, it only converges to the upper region of the Pareto front. This behavior is likely explained by the search scheme of DDS that stochastically reduces the number of perturbed decision variables per iteration. HYMOD has only five decision variables, so the expected number of decision variables perturbed per iteration in DDS becomes one as soon as the iteration count becomes $NFE^{4/5}$, where NFE is the total number of solution evaluations. For example, when the computational budget is 10,000 solution evaluations, DDS perturbs on average only a single decision variable of the selected solution after 1,585 solution evaluations (iteration 1,585 to 10,000). At this stage, DDS searches locally; see Tolson et al. (2009) and Asadzadeh and Tolson (2012), and therefore, PA-DDS only refines the currently archived solutions and would hardly be able to explore other parts of the decision space. This behavior is not observed from PA-DDS in SAC-SMA model that is set up for the exact same case study (Leaf River) or in any other MO problems in this study. SAC-SMA has 13 parameters and the expected number of perturbed decision variables per iteration becomes one after 4,924 solution evaluations. Therefore, it is concluded that this PA-DDS behavior is likely due to the low dimensionality of HYMOD.

In MCEW (Figure 5-12), comparing the best case performance, PA-DDS_CHC and AMALGAM perform similar, and ε-NSGAII results in a comparable front but with only a few solutions. Comparing the worst case performance (Figure 5-12-B), ε-NSGAII performs relatively very poorly compared to both of the AMALGAM and PA-DDS_CHC that almost converge to the best known Pareto front. Also at the low budget, PA-DDS_HVC performs very poorly compared to PA-DDS_CHC.

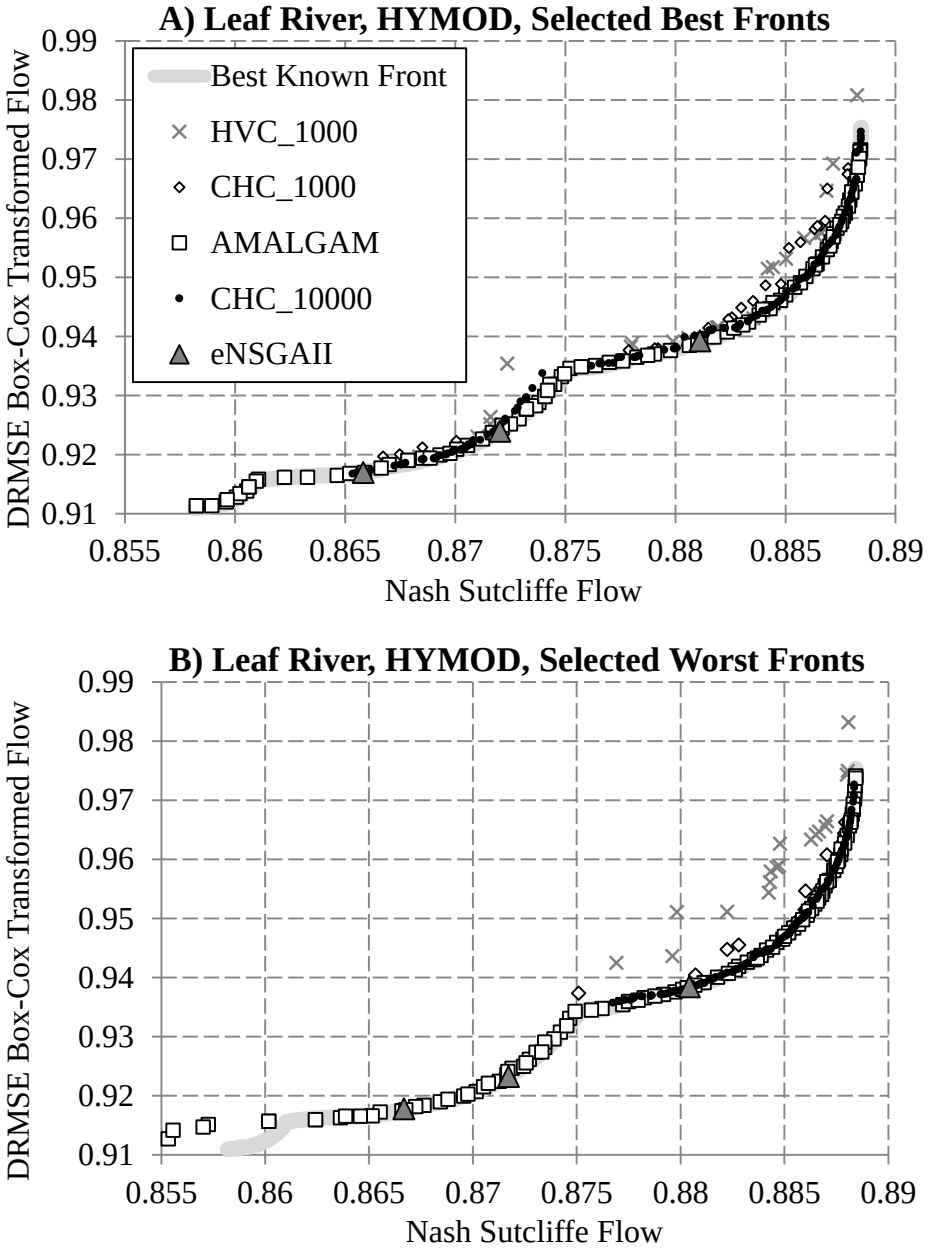

Figure 5-11 HYMOD calibration problem, Pareto approximate fronts of A) selected best trial and B) selected worst trial of PA-DDS with CHC and HVC selections with 1,000 solution evaluations and PA-DDS with CHC selection, ε-NSGAII and AMALGAM with 10,000 solution evaluations.

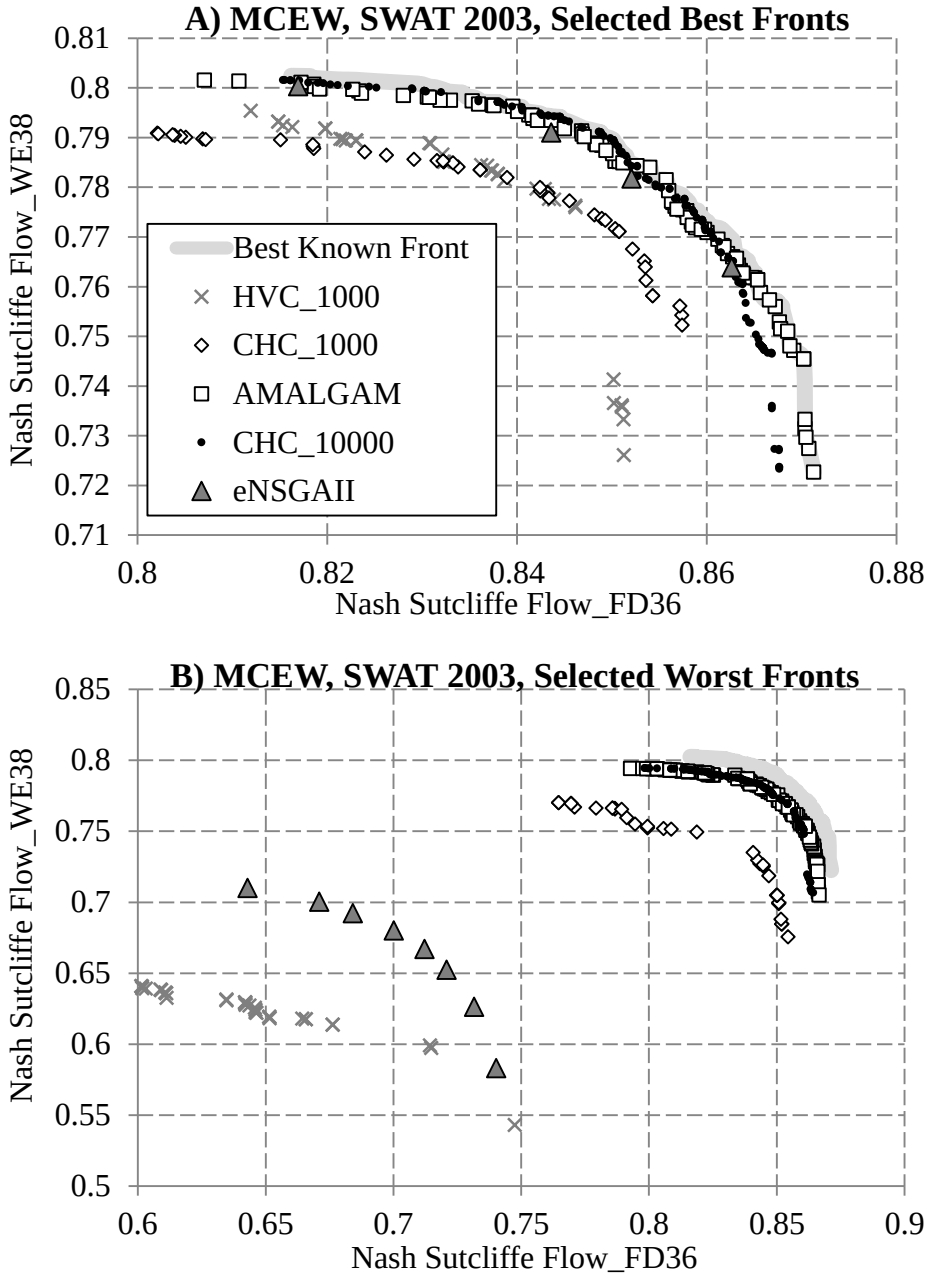

Figure 5-12 SWAT 2003 calibration problem, Pareto approximate fronts of A) selected best trial and B) selected worst trial of PA-DDS with CHC and HVC selections with 1,000 solution evaluations and PA-DDS with CHC selection, ε-NSGAII and AMALGAM with 10,000 solution evaluations.

5.4.3 Calibrated Model Simulation Results Comparison

Among all solutions on a Pareto approximate front of a MO calibration problem, the closest one to the ideal point is a reasonable selected solution for further evaluation. For example, this is a solution a modeller might select if they eventually needed to identify a single calibrated model. Such solutions are identified for all six case studies and for the best case and worst case performance of all five optimization approaches. So, ten solutions are identified for each case study. The range of objective function (simulation error metric) values for these solutions are reported in Table 5-5. Also, the objective function values of selected solutions from the best case and worst case performance of PA-DDS_CHC_1000 and alternative solutions from the best case and worst case performance of optimization approaches with 10,000 solution evaluations are reported in Table 5-5.

Table 5-5 Objective function (simulation error metric) values of selected solutions (calibrated models) from PA-DDS_CHC_1000 and the alternative solution by optimization runs with 10,000 solution evaluations. Solutions from the best and worst performed optimization are separated.

			Best Performed Trial		Worst Performed Trial	
Case Study	Objective Function	Range	PA-DDS CHC1000	Alternative	PA-DDS CHC1000	Alternative
EPANET2	SSRE FFT	(0.032, 0.0018)	0.00310	0.001753	0.03098	0.002284
C-Town	SSRE h1	$(0.0032, 1.3 \times 10^{-5})$	0.00108	0.000013	0.006264	0.000305
SWAT 2000	NS Q	(0.619, 0.697)	0.697	0.697	0.633	0.663
Town Brook 2D	NS P	(0.653, 0.749)	0.720	0.749	0.664	0.738
SWAT 2000 Town Brook 3D	NS Q	(0.541, 0.683)	0.657	0.683	0.560	0.649
	NS S	(0.637, 0.761)	0.729	0.761	0.637	0.730
	NS P	(0.649, 0.766)	0.740	0.766	0.649	0.728
HYMOD	NS	(0.870, 0.887)	0.870	0.870	0.880	0.877
Leaf River	DRMSE	(0.958, 0.921)	0.922	0.921	0.938	0.936
SAC-SMA	NS	(0.898, 0.913)	0.906	0.913	0.907	0.898
Leaf River	DRMSE	(1.032, 0.774)	0.843	0.774	0.999	0.946
SWAT 2003	NS FD36	(0.676, 0.852)	0.846	0.850	0.841	0.844
MCEW	NS WE38	(0.614, 0.789)	0.777	0.789	0.735	0.783

Beside the objective function comparison between the two alternative solutions in Table 5-5, the simulated and observed data of corresponding calibrated models are compared in Table 5-6 based on the maximum absolute simulation error and the corresponding percent relative error to the measured data and the average absolute error. This comparison represents the benefits one would expect by using PA-DDS_CHC_1000 for solving these calibration problems or using other MO algorithms in the comparison and a relatively larger computational budget of 10,000 solution evaluations.

In Table 5-6, the most significant difference between the two calibrated models occurs in the best case performance in SAC-SMA and the worst case performance in C-Town and Town Brook 3D. The simulated versus observed data for these three case studies are shown in Figure 5-13 to Figure 5-15.

Table 5-6 Some statistics of simulation error in calibrated model by PA-DDS_CHC_1000 versus the alternative calibrated model with 10,000 solution evaluations. The selected solutions from the best and worst performed optimization trials are separately studied.

			Max abs. error (%relative error)		Average abs. error	
			CHC_1000	Alternative	CHC_1000	Alternative
Best case Performance	EPANET2 C-Town	Flow (lps)	3 (1.2%)	0 (0.0%)	0.47	0.0
		Pressure (m)	1 (4.17%)	1 (4.17%)	0.18	0.03
	SWAT 2000 Town Brook 2D	Flow (cms)	7.25 (75.1%)	7.17 (74.3%)	0.365	0.354
		P (kg/day)	996.6 (75.1%)	943.0 (71.0%)	0.380	0.383
	SWAT 2000 Town Brook 3D	Flow (cms)	7.75 (80.3%)	7.19 (74.4%)	0.359	0.375
		Sed. (kg/day)	642 (65.9%)	621 63.7%)	5.75	5.31
		P (kg/day)	920 (69.3%)	871 (65.6%)	8.30	7.87
	HYMOD	Flow (cms)	139.2 (68.2%)	140.5 (68.9%)	7.41	7.40
	SAC-SMA	Flow (cms)	135.1 (195.5%)	83.0 (120.1%)	6.28	5.88
	SWAT 2003 MCEW	FD36 (cms)	0.0357 (35.1%)	0.0364 (35.8%)	0.0019	0.0019
		WE38 (cms)	0.685 (271.7%)	0.653 (259.0%)	0.041	0.040
Worst case Performance	EPANET2 C-Town	Flow (lps)	23 (9.3%)	1 (3.03%)	2.588	0.294
		Pressure (m)	3 (5.17%)	1 (4.17)	0.616	0.275
	SWAT 2000 Town Brook 2D	Flow (cms)	7.65 (82.1%)	8.20 (84.9%)	7.69	7.21
		P (kg/day)	1053 (79.3%)	1013 (76.3%)	8.41	7.07
	SWAT 2000 Town Brook 3D	Flow (cms)	8.78 (74.7%)	7.46 (77.3%)	0.422	0.406
		Sed. (kg/day)	797 (81.8%)	669 (68.7%)	5.803	5.132
		P (kg/day)	1126 (84.8%)	964 (72.6%)	8.702	7.833
	HYMOD	Flow (cms)	137.8 (67.5%)	138.7 (68.0%)	7.40	7.44
	SAC-SMA	Flow (cms)	122.4 (298%)	124.15 (302%)	6.79	6.77
	SWAT 2003 MCEW	FD36 (cms)	0.0378 (37.2%)	0.0374 (36.8%)	0.0020	0.0019
		WE38 (cms)	0.748 (296.7%)	0.632 (245.9%)	0.044	0.040

In Figure 5-13, the most significant difference between the two simulated time series occurs at day 335 where the selected solution from PA-DDS_CHC_1000 overestimates the observed flow more than the alternative solution does. Figure 5-14 shows that, the calibrated model by the selected solution from the worst performed trial of PA-DDS_CHC_1000 performs very well in estimating the pump flow and the static pressure measurements. However, it is noticeably worse than the alternative calibrated model in tank levels and fire flow test measurements. Interestingly, both of these calibrated models could identify the position of partially closed pipes. The time series of streamflow, total sediment transport, and total phosphorous delivery are presented in Figure 5-15 A), B), and C) respectively for these two calibrated problems. In general, the simulated time series are very similar.

The most significant difference between the flow time series occurs at day 100 and 881 where the calibrated model by PA-DDS_CHC_1000 underestimates the peak flow more than the alternative calibrated model does. However, the calibrated model by PA-DDS_CHC_1000 estimate the peak flows at days 180 and 481 better. The difference in the performance of these two calibrated models in simulating the total sediment transport and total phosphorus delivery are very similar. The calibrated model by PA-DDS_CHC_1000 underestimates the peak values at days 351 and 552 and overestimates the peak values at days 116 and 173 more than the alternative calibrated model does. However, it better estimates the peak values at days 545 and 557.

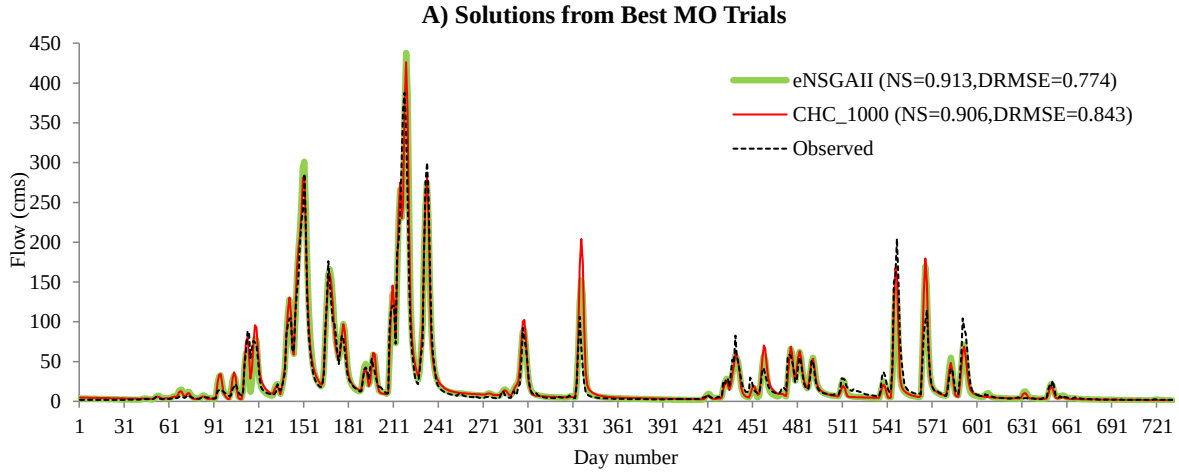

Figure 5-13 SAC-SMA, time series of observed versus simulated flow in the Leaf River watershed for selected solutions from the worst trials of PA-DDS_CHC_1000 and alternative solution from the worst MO trials with 10,000 solution evaluations.

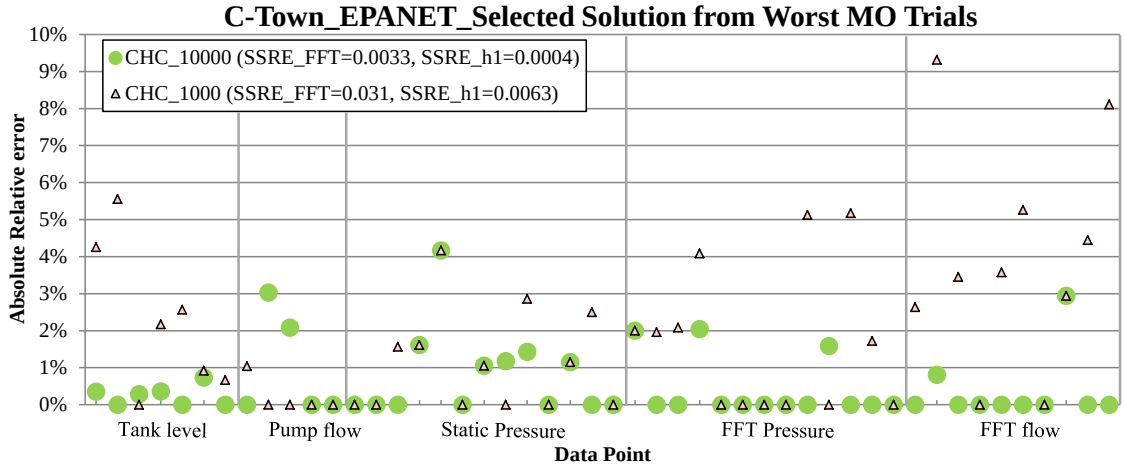

Figure 5-14 EPANET2, time series of observed versus simulated data in C-Town for selected solutions from the worst trials of PA-DDS_CHC_1000 and alternative solution from the worst MO trials with 10,000 solution evaluations.

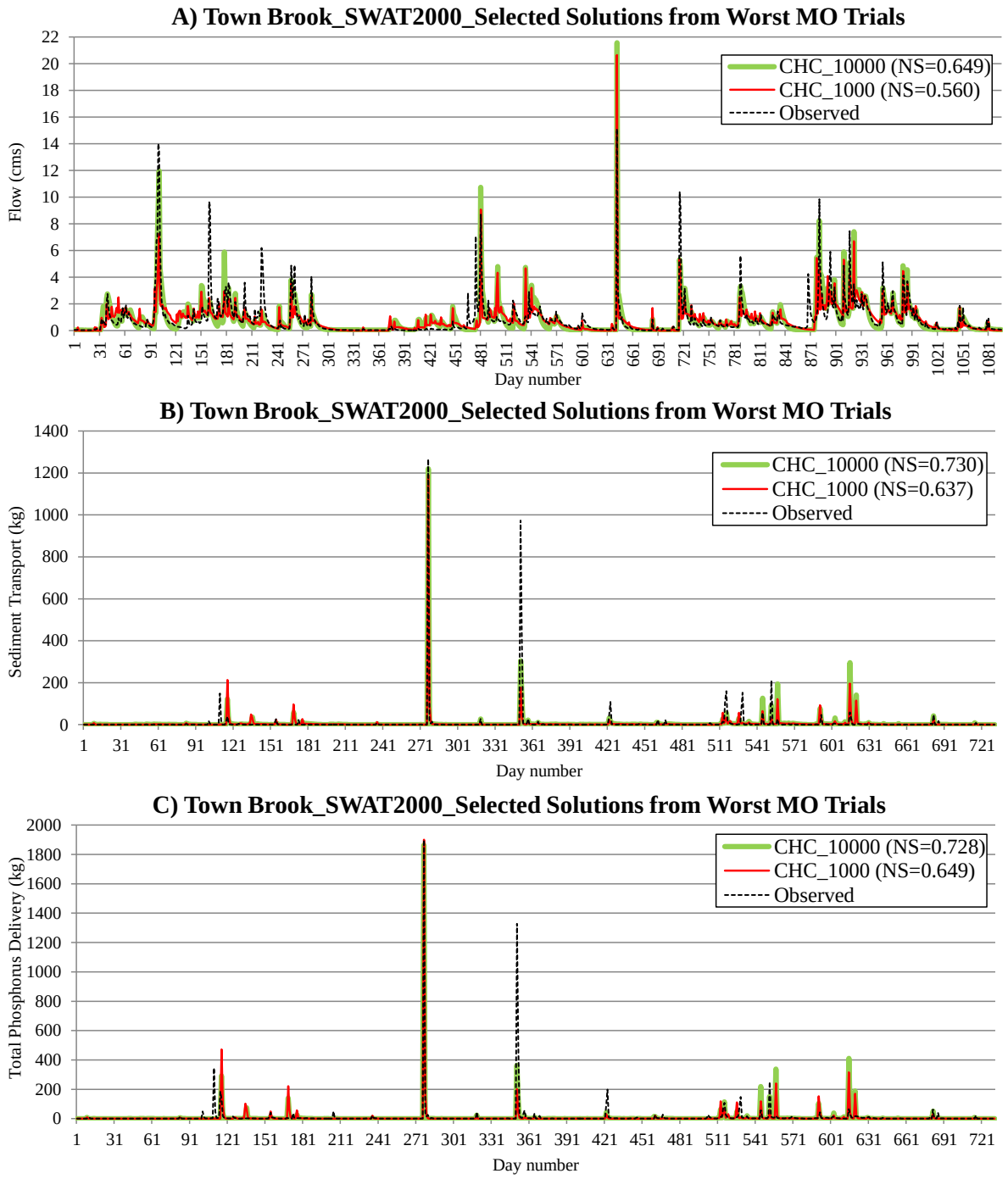

Figure 5-15 SWAT 2000, time series of observed versus simulated Flow (A), Sediment (B) and Phosphorous (C) in the Town Brook watershed for selected solutions from the worst trials of PA-DDS_CHC_1000 and alternative solution from the worst MO trials with 10,000 solution evaluations.

5.5 Conclusions

Based on the results in this section, when the computational budget is limited to 1,000 solution evaluations, PA-DDS_CHC can reach practically acceptable calibration solutions compared to the results of the best algorithm at a high budget of 10,000 solution evaluations. When solving MO problems with known/expected convex Pareto front results in this chapter demonstrate that CHC is the preferred selection metric for PA-DDS. Also, when the computational budget is relatively large, PA-DDS_CHC performs comparable to AMALGAM and ε-NSGAII.

Chapter 6

Summary, Conclusions and Recommendations for Future Work

6.1 Summary

The main goal of this thesis was to develop parsimonious and efficient heuristic optimization algorithms for solving water resources simulation-optimization problems and specifically for those that have computationally intensive simulation models. In such problems a high quality solution obtained within the analyst timeframe is preferred to the global optimal solution which can only be obtained by very large numbers of solution evaluations. Therefore, the analyst may require an acceptable solution after a limited number of solution evaluations and may not be able to spend any of those evaluations for fine-tuning the optimization algorithm.

In the first phase of this thesis, a heuristic single objective optimization algorithm, HD-DDS was developed for solving problems with discrete decision variables (demonstrated specifically for Water Distribution System, WDS, design problems). HD-DDS uses two straightforward local search techniques called L_1 and L_2. HD-DDS is an easy-to-use optimization algorithm, as it does not require users to experiment and identify good optimization algorithm parameters. Instead, it gives the users a robust optimization tool to experiment and solve multiple optimization problems with different design characteristics. In the case of WDS design, these characteristics would include different design constraint sets, different objectives and different future scenarios leading to different nodal demand scenarios.

The high efficiency and parsimony of HD-DDS motivated the development of a multi-objective optimization algorithm based on DDS. This algorithm is called PA-DDS and uses the original or discrete versions of DDS as search engine, so it can handle both discrete and continuous decision variables. PA-DDS has an unbounded archive size and archives all non-dominated solutions during the search. The unbounded archive size avoids adding any new parameter, so PA-DDS has a single parameter which is the perturbation size of the search engine. However, the unbounded archive size can make PA-DDS an inefficient algorithm for solving quick to evaluate problems, especially those with more than five objective functions solved with budgets in the order of 10^6 solution evaluations or more. The initial version of PA-DDS with a selection based on crowding distance was hybridized by a straightforward local search technique called L, and the algorithm was successfully applied to several bi-objective water distribution system design problems.

Research about the impact of selection metric on the performance of PA-DDS showed that, Hypervolume contribution (Knowles et al. 2003), referred to as HVC1 in Chapter 4, is the superior selection metric compared to other selection metrics (including crowding distance) for solving multi-objective optimization problems with a general (unknown) Pareto front shape. This improvement is not dependent to the computational budget; therefore, HVC1 is designated as the most effective and the default selection metric of PA-DDS. Whenever the exact calculation of HVC1 is applicable, i.e. problems with two or three objective functions, HVC1 does not introduce any new parameter to PA-DDS. However, in problems with more than three objective functions where the exact calculation of HVC1 is not applicable due to its computation burden, the Monte Carlo approximation of HVC1 can be used where number of Monte Carlo samples should be specified. Results indicate using 10,000 samples worked well.

Further research on the selection metric of PA-DDS focused on the development of a novel selection metric called Convex Hull Contribution (CHC) specifically designed for solving problems with a known or expected convex Pareto front. PA-DDS with selection based on CHC was successfully applied to water resources multi-objective optimization problems with up to three objective functions. Results in Chapter 5 show that CHC significantly improves the performance of PA-DDS compared to HVC1 when solving problems with a convex Pareto front and by a relatively limited computational budget. The reason is that CHC makes PA-DDS sample only from a subset of archived non-dominated solutions that form the convex approximation of the Pareto front during the search while HVC1 gives a chance to all archived non-dominated solutions to be selected and perturbed in the search for new solutions. This performance improvement by CHC is less evident when the computational budget is large because with such budgets the multi-objective optimization algorithm has enough time to sample from all archived solutions and find high quality result.

6.2 Conclusions and Guidelines for Users of the Developed Algorithms

The single-objective and multi-objective optimization algorithms developed in this research, HD-DDS and PA-DDS respectively, were successfully applied to several water resources simulation-optimization problems.

For the range of Water Distribution System design benchmark case studies considered in Chapter 2, numerical results demonstrate that the HD-DDS algorithm exhibits superior performance in specific comparisons to the MMAS ACO (Zecchin et al. 2007), GENOME GA (Reca and Martinez, 2006), and PSO variant (Montalvo et al. 2008) algorithms. For the same number of objective

function evaluations HD-DDS stochastically dominates MMAS ACO results in all three case studies for which their performance is compared. This is achieved despite the fact that no parameter tuning was conducted in HD-DDS while MMAS ACO parameters were specifically tuned to each case study (involving millions of EPANET2 simulations). The worst HD-DDS result was better than the best GENOME GA result for the 454 decision variable Balerma network even though the GENOME GA utilized 100 times more objective function evaluations. In addition, HD-DDS found a new best solution to the Balerma problem. HD-DDS found the best known solutions more frequently and easily avoided the worst solutions returned by the PSO variant despite the fact that PSO algorithm parameters were determined with preliminary tuning experiments. The parameter-free constraint handling approach based on Deb (2000) was successful. Furthermore, because the evaluation of many of the candidate solutions identified by HD-DDS does not require simulating network hydraulics (e.g., 50%-70% of objective function evaluations in this chapter), the actual HD-DDS computation time would be much less (by nearly 50%-70%) than that of the comparative algorithms in this chapter. This computational advantage of HD-DDS extends over any optimization algorithm requiring that network hydraulics be simulated for all candidate solutions.

In the most general interpretation of L_1 and L_2 local searches, HD-DDS is a general methodology that is not specific to water distribution system design problems. L_1 enumerates all solutions that differ from the current best solution by a single decision variable. L_2 enumerates all solutions that differ from the current best solution by only two decision variables. Potential applications of HD-DDS to other types of constrained discrete optimization problems in the field of water resources and environmental management include watershed best management practice optimization, e.g. Arabi et al. (2006) and groundwater management and monitoring problems, e.g. Reed et al. (2000). Also as shown in Matott et al. (2012), HD-DDS is applicable to sorptive barrier design problems to minimize the design cost by selecting from the available sorptive barrier layers that achieve desired level of containment treatment for hazardous waste. HD-DDS application to new problem types like these requires users consider whether small changes are necessary in the two local search types defined here.

In Chapter 3, the initial version of PA-DDS was successfully applied to bi-objective water distribution system design problems with discrete decision variables and results show the comparable performance of PA-DDS and two popular multi-objective optimization algorithms NSGAII and SPEA2. Also, the very straightforward neighborhood search technique L developed for hybridizing PA-DDS in Chapter 3 was applied to the other two algorithms and results show that when the

computational budget is relatively low, this local search improves the results from about 18% to more than 250% in the five case studies. This improvement decreases to 0% to slightly less than 115% as the computational budget of the global search algorithm (NSGAII or SPEA2 in Chapter 3) increases by an order of magnitude.

In Chapter 4, a more advanced version of PA-DDS with selection based on HVC1 was successfully applied to a hydrologic model calibration problem with three objective functions. Results suggest that despite its parsimony, this version of PA-DDS is promising compared to ε-NSGAII and AMALGAM for solving water resources calibration problems with a somewhat limited computational budget, e.g. 10,000 solution evaluations in this problem. However, results in Chapter 5 show that when the computational budget is even more limited, e.g. 1,000 solution evaluations in that problem, selection based on CHC is preferred to HVC1. PA-DDS with selection based on CHC shows superior performance compared to PA-DDS with selection based on HVC1 for solving five other water resources calibration problems in Chapter 5. Solving these problems by PA-DDS with selection based on CHC results in acceptable calibrated models with a computational budget that is deemed limited (one order of magnitude less than the budget that appears to be the minimum budget used in the literature to solve similar problems). This was only possible in PA-DDS by using CHC because with the same computational budget, PA-DDS with selection based on HVC1 performs noticeably worse than PA-DDS with selection based on CHC. Therefore, it is recommended that if PA-DDS users know a priori or expect that their multi-objective optimization problem has a convex Pareto front, they should choose CHC selection - especially if the computational budget is deemed limited.

The application of PA-DDS with selection based on HVC1 to general mathematical multi-objective optimization problems (with up to five objectives) in Chapter 4 shows comparable results to other recently developed multi-objective optimization algorithms. Although PA-DDS was developed to work well for limited computational budgets (computationally intensive problems), when using the rather large budget of 300,000 solution evaluations, PA-DDS ranked 7th overall out of 14 algorithms based on the CEC09 MO competition. For the 5-objective test problems in CEC09, PA-DDS ranked 2nd overall out of ten algorithms. So, it is recommended that PA-DDS users choose HVC1 if they do not have prior knowledge about the shape of the Pareto front.

The performance of PA-DDS relative to other multi-objective optimization algorithms and across a range of problems from water distribution system design problems (Chapter 3) and mathematical test problems and water resources model calibration problems (Chapter 4 and Chapter 5) suggests that the

default value of the single parameter of PA-DDS ($r = 0.2$) is robust. Researchers who seek to understand how PA-DDS performance is impacted by alternative algorithm parameter r values are referred to Hadka and Reed (2012) who demonstrate a new systematic way to assess the overall performance of algorithms under a statistically sampled parameterization and computational budget. For regular algorithm users interested in the best PA-DDS performance on their problem, instead of attempting to fine tune the parameter value, it is recommended to keep r fixed at 0.2 and spend the available computational budget solving multiple independent optimization trials and then aggregate all non-dominated solutions.

Based on relative PA-DDS performance and an assessment of computational time associated with PA-DDS algorithm procedures, there is no strong reason to bound the algorithm's archive size when solving computationally intensive problems with five or fewer objectives. Although Laumanns et al. (2002) report that archiving all non-dominated solutions does not let the multi-objective optimization algorithms focus only on interesting solutions (and thus potentially degrade results), results in this thesis show that PA-DDS performs comparable to popular multi-objective optimization algorithms with a bounded archive size for MO problems with up to five objective functions. It is systematically demonstrated in Chapter 4 that the computational cost of PA-DDS operators (solution archiving, selection, and perturbation) with an unbounded archive is negligible compared to the objective functions evaluation runtime when dealing with computationally intensive problems. In eight example problems with varying computational budgets up to 10^6 solution evaluations and varying number of objective functions up to five, it is shown that as long as solution evaluation runtime requires more than 10 seconds, only 10% or less of total runtime is attributable to PA-DDS operators including HVC1 calculation even though PA-DDS archives all non-dominated solutions. However, PA-DDS with the current implementation of the hypervolume contribution calculation is not recommended for solving quick to evaluate problems such as test problems – especially with budgets of 10^6 solution evaluations or more. Multi-objective optimization algorithms with a bounded archive size would perform better than PA-DDS for solving such problems given the same computation time.

In light of the fact that non-dominated space grows exponentially with increasing objective counts and for problems with continuous decision variables this can be infinite, there may exist computationally intensive problems (e.g., those that have more than five objective functions and/or solved with a budget over 10^6 solution evaluations) where the unbounded archive approach unacceptably slows the algorithm progress. A minor change in the implementation of PA-DDS and a minor change in the problem setting can help mitigate such an issue. First, the current linear

dominance test of PA-DDS can be modified to a data structure-based dominance test which is more efficient than the linear dominance test when the archive size is large (Mostaghim et al. 2002). Second, the user can provide PA-DDS with objective function values that are rounded so that meaningless differences in objective function values are not utilized by the algorithm. This would reduce PA-DDS archive size because candidate solutions with the same rounded objective function values as a solution currently in the archive are discarded and never archived. The same rounding approach could also be used to reduce the PA-DDS algorithm runtime when applied to problems with solution evaluation runtime that is only a fraction of a second. An alternative to rounding would be to simply utilize the epsilon dominance archiving as in ε-NSGAII. Although the runtime analysis in Chapter 4 is not conducted for PA-DDS with CHC, optimization trials in Chapter 5 did not show any runtime issue for PA-DDS with CHC compared to PA-DDS with HVC1. Therefore, a similar conclusion is applied to PA-DDS with CHC.

PA-DDS results for calibrating a multi-objective optimization problem with five decision variables (HYMOD) show that it is possible that PA-DDS converges to a portion of the Pareto front. This behavior is observed in all the worst performed trial of PA-DDS with CHC or HVC1 selections and with 1,000 or 10,000 solution evaluations. It is concluded that the DDS search scheme likely causes this behavior because in such problems DDS becomes a local search very early in the search by perturbing only one decision variable per iteration on average. This behavior motivated the hybridization of optimization algorithms in Chapter 2 and Chapter 3. This conclusion confirms the statement in Tolson and Shoemaker (2007) that recommended the original single objective DDS for problems with more than six decision variables. Therefore, PA-DDS with the settings introduced in this thesis is not recommended for solving problems with as few decision variables as six.

6.3 Recommendations for Future Work

In this thesis, PA-DDS with CHC is only applied to problems with two or three objective functions. Future work can include the application of PA-DDS with CHC for solving problems with more than three objective functions. Also, the computational budget is limited to 10,000 solution evaluations; therefore, the expected number of archived solutions in the unbounded archive of PA-DDS is not large. Further investigations are required to study the runtime of PA-DDS with CHC as a function of the number of archived solutions and the number of objective functions. Such research can follow the runtime analysis conducted in Chapter 4 of this thesis. Also, such research would require multi-objective optimization problems with more than three objective functions and convex Pareto fronts.

Beside the multi-objective calibration problems that are proposed and solved in the literature, e.g. Kollat et al. (2012); similar approach to make DTLZ2 a three objective optimization problem with convex Pareto front in Chapter 5 can be followed in higher dimensions.

The unbounded archive size of PA-DDS can limit the application hypervolume contribution as its selection for solving quick to evaluate problems, especially those with more than five objective functions solve with budgets of 10^6 solution evaluations or more. As more efficient implementations of the hypervolume calculation become available in the future, this issue will be mitigated.

Future work should also include research to identify for PA-DDS whether the archived set of non-dominated solutions represent a convex shape. The result of such research would be a criterion for PA-DDS to automatically decide during the search whether to use CHC or HVC1 selection metric at each iteration. A proposed criterion can be based on the portion of convex hull area (hypervolume in problems with more than two objective functions) corresponding to all vertices of top and bottom facets (see section 5.2.1).

Although CHC is utilized to guide the search operator of PA-DDS, it is a general selection metric. Future research can include the application of CHC to guide the search operator of other optimization algorithms such as a Genetic Algorithm based multi-objective optimization algorithm. The main difference between such algorithm and CHEA introduced by Cococcioni et al. (2007) is that CHC considers the contribution of solutions to the convex hull and therefore it measures the diversity of solutions while CHEA only considers the proximity of solutions. Also, CHEA is not applied to problems with more than two objective functions while results of Chapter 5 in this thesis suggest that the optimization algorithm that uses CHC as its selection metric can be applied to problems with at least three objective functions.

The successful application of neighborhood search techniques L_1 and L_2 in HD-DDS (Chapter 2) and L in PA-DDS (Chapter 3) suggests the importance of these algorithms in fine-tuning the final results. Future work can include the development of local search techniques for DDS and PA-DDS when dealing with continuous decision variables. It is recommended that algorithm parsimony be considered in such hybridizations. Proposed local search techniques can mimic L_1 and L_2 for DDS and L for PA-DDS by discretizing the range of continuous decision variables based on the available computational budget and enumerate all solutions that differ from the current best solution by a single decision variable and/or enumerate all solutions that differ from the current best solution by only two decision variables.

Future work can also include careful modifications of DDS algorithm for solving problems with few decision variables. A proposed research can investigate the application of DDS to various types of problems with discrete and continuous decision variables to identify when DDS search loses its effectiveness and propose an alternative search scheme. An alternative search scheme can mimic the L_2 local search in HD-DDS that enumerates all the possible two-decision-variable changes (by one discrete option) to search for better solutions. The alternative search scheme should respect the parsimony of the algorithm and therefore not introduce any new algorithm parameters.

References

Alperovits, E., and Shamir, U. (1977). "Design of optimal water distribution systems." *Water Resources Research*, 13(6), 885-900.

Arabi, M., Govindaraju, R. S., and Hantush, M. M. (2006). "Cost-effective allocation of watershed management practices using a genetic algorithm." *Water Resources Research*, 42(10).

Asadzadeh M. and Tolson B. A. (in press). "Pareto Archived Dynamically Dimensioned Search with Hypervolume Based Selection for Multi Objective Optimization." *Engineering Optimization*.

Asadzadeh, M., and Tolson, B. A. (2009). "A new multi-objective algorithm, pareto archived DDS." *11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers*, G. Raidl, E. Alba, J. Bacardit, C. B. Congdon, H.-G. Beyer, M. Birattari, C. Blum, P. A. N. Bosman, D. Corne, C. Cotta, M. D. Penta, B. Doerr, R. Drechsler, M. Ebner, J. Grahl, T. Jansen, J. Knowles, T. Lenaerts, M. Middendorf, J. F. Miller, M. O'Neill, R. Poli, G. Squillero, K. Stanley, T. Stützle, and J. v. Hemert, eds., ACM, Montreal, Quebec, Canada, 1963-1966.

Asadzadeh M., Tolson B. A., and McKillop R. (2010). "A Two Stage Optimization Approach for Calibrating Water Distribution Systems." *In 12th water distribution systems analysis symposium, (WDSA 2010 - Battle of Water Calibration Networks)*, Lansey K., Choi C., Ostfeld A., and Pepper I. , eds., ASCE, Tucson, Arizona.

Asadzadeh, M., and Tolson, B. (2012). "Hybrid Pareto archived dynamically dimensioned search for multi-objective combinatorial optimization: application to water distribution network design." *Journal of Hydroinformatics*, 14(1), 192-205.

Atiquzzaman, M., Liong, S., and Yu, X. (2006). "Alternative Decision Making in Water Distribution Network with NSGA-II." *Journal of Water Resources Planning and Management*, 132(2), 122-126.

Bader, J., Deb, K., and Zitzler, E. (2008). "Faster Hypervolume-Based Search Using Monte Carlo Sampling." *19th International Conference on Multiple Criteria Decision Making: Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems*, M. Ehrgott, B. Naujoks, T. J. Stewart, and J. Wallenius, eds., Springer, Berlin, Germany, 313-326.

Bader, J., and Zitzler, E. (2011). "HypE: An Algorithm for Fast Hypervolume-Based Many-Objective Optimization." *Evolutionary Computation*, 19(1), 45-76.

Barber, C. B., Dobkin, D. P., and Huhdanpaa, H. (1996). "The Quickhull algorithm for convex hulls." *Acm Transactions on Mathematical Software*, 22(4), 469-483.

Bekele, E. G., and Nicklow, J. W. (2007). "Multi-objective automatic calibration of SWAT using NSGA-II." *Journal of Hydrology*, 341(3-4), 165-176.

Beume, N., Naujoks, B., and Emmerich, M. (2007). "SMS-EMOA: Multiobjective selection based on dominated hypervolume." *European Journal of Operational Research*, 181(3), 1653-1669.

Beyer, H.G. (1996) "Toward a Theory of Evolution Strategies: Self-Adaptation." *Evolutionary Computation*, 3(3), 311-347.

Bhave, P. R., and Sonak, V. V. (1992). "A critical study of the linear programming gradient method for optimal design of water supply networks." *Water Resources Research*, 28(6), 1577-1584.

Boccelli, D. L., Tryby, M., Uber, J., Rossman, L., Zierolf, M., and Polycarpou, M. M. (1998). "Optimal Scheduling of Booster Disinfection in Water Distribution Systems." *Journal of Water Resources Planning and Management*, 124(2), 99-111.

Bosman, P. A. N., and de Jong, E. D. (2006). "Combining gradient techniques for numerical multi-objective evolutionary optimization." *8th Annual Conference on Genetic and evolutionary computation*, ACM, Seattle, Washington, USA, 627-634.

Boyle, D. P., Gupta, H. V., and Sorooshian, S. (2000). "Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods." *Water Resources Research*, 36(12), 3663-3674.

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 Congress on Evolutionary Computation*, G. G. Yen, S. M. Lucas, G. Fogel, G. Kendall, R. Salomon, B.-T. Zhang, C. A. C. Coello, and T. P. Runarsson, eds., IEEE Press, Vancouver, BC, Canada, 710-717.

Brown, M., and Smith, R. E. (2003). "Effective use of directional information in multi-objective evolutionary computation." *Genetic and Evolutionary Computation - GECCO 2003*, E. Cant-Paz, J. A. Foster, K. Deb, L. Davis, R. Roy, U.-M. O'Reilly, H.-G. Beyer, R. K. Standish, G. Kendall, S. W. Wilson, M. Harman, J. Wegener, D. Dasgupta, M. A. Potter, A. C. Schultz, K. A. Dowsland, N. Jonoska, and J. F. Miller, eds., Springer, 778-789.

Carrano, E. G., Wanner, E. F., and Takahashi, R. H. C. (2011). "A Multicriteria Statistical Based Comparison Methodology for Evaluating Evolutionary Algorithms." *IEEE Transactions on Evolutionary Computation*, 15(6), 848-870.

Chang, L.-C., Ho, C.-C., and Chen, Y.-W. (2010). "Applying Multiobjective Genetic Algorithm to Analyze the Conflict among Different Water Use Sectors during Drought Period." *Journal of Water Resources Planning and Management*, 136(5), 539-546.

Chen, W.-M., and Lee, W.-T. (2007). "An efficient evolutionary algorithm for multiobjective optimization problems." *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, IEEE press, 30-33.

Cococcioni, M., Ducange, P., Lazzerini, B., and Marcelloni, F. (2007). "A new multi-objective evolutionary algorithm based on convex hull for binary classifier optimization." *Congress on Evolutionary Computation (CEC '07)*, IEEE, New York, 3150-3156.

Coello, C. A. C. (2001). "A short tutorial on evolutionary multiobjective optimization." *Evolutionary Multi-Criterion Optimization*, E. Zitzler, K. Deb, L. Thiele, C. A. Coello Coello, and D. Corne, eds., Springer, Berlin, 21-40.

Coello, C. A. C., Lamont, G. B., and Van Veldhuisen, D. A. (2007). *Evolutionary Algorithms for Solving Multi-Objective Problems*, Springer.

Confesor, R. B., Jr., and Whittaker, G. W. (2007). "Automatic calibration of hydrologic models with multi-objective evolutionary algorithm and Pareto optimization." *Journal of the American Water Resources Association*, 43(4), 981-989.

Cunha, M., and Sousa, J. (1999). "Water Distribution Network Design Optimization: Simulated Annealing Approach." *Journal of Water Resources Planning and Management*, 125(4), 215-221.

Cunha, M., and Ribeiro, L. (2004). "Tabu search algorithms for water network optimization." *European Journal of Operational Research*, 157(3), 746-758.

De Jong, K. A. (1975). "An analysis of the behavior of a class of genetic adaptive systems." PhD, University of Michigan, Ann Arbor, MI, USA.

Deb, K. (1999). "Multi-objective Genetic Algorithms: Problem Difficulties and Construction of Test Problems." *Evolutionary Computation*, 7(3), 205-230.

Deb, K. (2000). "An efficient constraint handling method for genetic algorithms." *Computer Methods in Applied Mechanics and Engineering*, 186(2-4), 311-338.

Deb, K. (2001). *Multi-Objective Optimization Using Evolutionary Algorithms*, Wiley.

Deb, K., and Goel, T. (2001). "A hybrid multi-objective evolutionary approach to engineering shape design." *Evolutionary Multi-Criterion Optimization*, E. Zitzler, K. Deb, L. Thiele, C. A. Coello Coello, and D. Corne, eds., Springer, 385-399.

Deb, K., Thiele, L., Laumanns, M., and Zitzler, E. (2001). "Scalable Test Problems for Evolutionary Multi-Objective Optimization." *KANGAL REPORT*, Indian Institute of Technology Kanpur, Kanpur, India.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). "A fast and elitist multiobjective genetic algorithm: NSGA-II." *IEEE Transactions on Evolutionary Computation* 6(2), 182-197.

Deb, K., Mohan, M., and Mishra, S. (2003). "A Fast Multi-objective Evolutionary Algorithm for Finding Well-Spread Pareto-Optimal Solutions." *KANGAL REPORT*, Indian Institute of Technology Kanpur, Kanpur, India.

di Pierro, F., Khu, S.-T., Savic, D., and Berardi, L. (2009). "Efficient multi-objective optimal design of water distribution networks on a budget of simulations using hybrid algorithms." *Environmental Modelling & Software*, 24(2), 202-213.

Eiben, A. E., Hinterding, R., and Michalewicz, Z. (1999). "Parameter control in evolutionary algorithms." *IEEE Transactions on Evolutionary Computation*, 3(2), 124-141.

Eiger, G., Shamir, U., and Ben-Tal, A. (1994). "Optimal design of water distribution networks." *Water Resources Research*, 30(9), 2637-2646.

Emmerich, M., Beume, N., and Naujoks, B. (2005). "An EMO algorithm using the hypervolume measure as selection criterion." *Evolutionary Multi-Criterion Optimization*, C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, eds., Springer, 62-76.

Eusuff, M., and Lansey, K. (2003). "Optimization of Water Distribution Network Design Using the Shuffled Frog Leaping Algorithm." *Journal of Water Resources Planning and Management*, 129(3), 210-225.

Farmani, R., Savic, D. A., and Walters, G. A. (2005). "Evolutionary multi-objective optimization in water distribution network design." *Engineering Optimization*, 37(2), 167-183.

Farmani, R., Wright, J. A., Savic, D. A., and Walters, G. A. (2005). "Self-adaptive fitness formulation for evolutionary constrained optimization of water systems." *Journal of Computing in Civil Engineering*, 19(2), 212-216.

Feng, C. W., Liu, L. A., and Burns, S. A. (1997). "Using genetic algorithms to solve construction time-cost trade-off problems." *Journal of Computing in Civil Engineering*, 11(3), 184-189.

Fenicia, F., Solomatine, D. P., Savenije, H. H. G., and Matgen, P. (2007). "Soft combination of local models in a multi-objective framework." *Hydrology and Earth System Sciences*, 11(6), 1797-1809.

Fieldsend, J. E., Everson, R. M., and Singh, S. (2003). "Using unconstrained elite archives for multiobjective optimization." *IEEE Transactions on Evolutionary Computation*, 7(3), 305-323.

Fonseca, C. M., and Fleming, P. J. (1998). "Multiobjective optimization and multiple constraint handling with evolutionary algorithms - Part I: A unified formulation." *IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans*, 28(1), 26-37.

Fonseca, C. M., Paquete, L., and Lopez-Ibanez, M. (2006). "An improved dimension-sweep algorithm for the hypervolume indicator." *IEEE Congress on Evolutionary Computation*, IEEE press, Vancouver, Canada, 1142-1148.

Fujiwara, O., and Khang, D. B. (1990). "A two-phase decomposition method for optimal design of looped water distribution networks." *Water Resources Research*, 26(4), 539-549.

Geem, Z. W. (2006). "Optimal cost design of water distribution networks using harmony search." *Engineering Optimization*, 38(3), 259-277.

Gessler, J. (1985). "Pipe network optimization by enumeration." *Computer Applications for Water Resources*, H. C. Torno, ed., American Society of Civil Engineers, New York, 572-581.

Gibbons, J. D., and Chakraborti, S. (2003). *Nonparametric Statistical Inference*, Marcel Dekker Incorporated.

Gibbs, M. S., Dandy, G. C., and Maier, H. R. (2008). "A genetic algorithm calibration method based on convergence due to genetic drift." *Information Sciences*, 178(14), 2857-2869.

Gibbs, M. S., Maier, H. R., and Dandy, G. C. (2010). "Comparison of Genetic Algorithm Parameter Setting Methods for Chlorine Injection Optimization." *J. Water Resour. Plan. Manage.-ASCE*, 136(2), 288-291.

Goel, T., and Deb, K. (2001). "Hybrid Methods for Multi-Objective Evolutionary Algorithms." *KANGAL REPORT*, Indian Institute of Technology Kanpur, Kanpur, India.

Gupta, H. V., Sorooshian, S., and Yapo, P. O. (1998). "Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information." *Water Resources Research*, 34(4), 751-763.

Hadka, D., and Reed, P. (2012). "Diagnostic Assessment of Search Controls and Failure Modes in Many-Objective Evolutionary Optimization." *Evolutionary Computation*, 20(3), 423-452.

Hanne, T. (1999). "On the convergence of multiobjective evolutionary algorithms." *European Journal of Operational Research*, 117(3), 553-564.

Huband, S., Hingston, P., Barone, L., and While, L. (2006). "A review of multiobjective test problems and a scalable test problem toolkit." *IEEE Transactions on Evolutionary Computation*, 10(5), 477-506.

Igel, C., Hansen, N., and Roth, S. (2007). "Covariance matrix adaptation for multi-objective optimization." *Evolutionary Computation*, 15(1), 1-28.

Ishibuchi, H., and Murata, T. (1996). "Multi-objective genetic local search algorithm." *IEEE International Conference on Evolutionary Computation*, IEEE, Nayoya University, Japan, 119-124.

Ishibuchi, H., Yoshida, T., and Murata, T. (2003). "Balance between genetic search and local search in memetic algorithms for multiobjective permutation flowshop scheduling." *IEEE Transactions on Evolutionary Computation*, 7(2), 204-223.

Jaszkiewicz, A. (2002). "Genetic local search for multi-objective combinatorial optimization." *European Journal of Operational Research*, 137(1), 50-71.

Jourdan, L., Corne, D., Savic, D., and Walters, G. (2005). "Preliminary Investigation of the 'Learnable Evolution Model' for Faster/Better Multiobjective Water Systems Design." *Evolutionary Multi-Criterion Optimization*, C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, eds., Springer, 841-855.

Kaylani, A., Georgiopoulos, M., Mollaghasemi, M., Anagnostopoulos, G. C., Sentelle, C., and Zhong, M. (2010). "An Adaptive Multiobjective Approach to Evolving ART Architectures." *IEEE Transactions on Neural Networks*, 21(4), 529-550.

Keedwell, E., and Khu, S. T. (2006). "Novel cellular automata approach to optimal water distribution network design." *Journal of Computing in Civil Engineering*, 20(1), 49-56.

Khu, S. T., and Madsen, H. (2005). "Multiobjective calibration with Pareto preference ordering: An application to rainfall-runoff model calibration." *Water Resources Research*, 41(3).

Kim, J. H., Kim, T. G., Kim, J. H., and Yoon, Y. N. (1994). "A study on the pipe network system design using non-linear programming." *Korean Water Resources Association*, 27(4), 59-67.

Kleeman, M. P., Lamont, G. B., Cooney, A., and Nelson, T. R. (2007). "A multi-tiered memetic multiobjective evolutionary algorithm for the design of quantum cascade lasers." *Evolutionary Multi-Criterion Optimization*, S. Obayashi, K. Deb, C. Poloni, T. Hiroyasu, and T. Murata, eds., Springer, 186-200.

Knowles, J. D., and Corne, D. W. (2000). "Approximating the Nondominated Front Using the Pareto Archived Evolution Strategy." *Evolutionary Computation*, 8(2), 149-172.

Knowles, J. D., Corne, D. W., and Fleischer, M. (2003). "Bounded archiving using the lebesgue measure." *Congress on Evolutionary Computation, CEC '03*, R. Sarker, R. Reynolds, H.

Abbass, K. C. Tan, B. McKay, Daryl Essam, and T. Gedeon, eds., IEEE press, Canberra, Australia, 2490-2497 Vol.2494.

Knowles, J. D. (2006). "ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems." *IEEE Transactions on Evolutionary Computation*. 10(1), 50-66.

Kollat, J., and Reed, P. (2005). "The Value of Online Adaptive Search: A Performance Comparison of NSGAII, ε-NSGAII and εMOEA." *Evolutionary Multi-Criterion Optimization*, C. Coello Coello, A. Hernández Aguirre, and E. Zitzler, eds., Springer 386-398.

Kollat, J. B., and Reed, P. M. (2006). "Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design." *Advances in Water Resources*, 29(6), 792-807.

Kollat, J. B., Reed, P. M., and Wagener, T. (2012). "When are multiobjective calibration trade-offs in hydrologic models meaningful?" *Water Resources Research*, 48.

Lamont, P. A. (1981). "Common Pipe-Flow Formulas Compared with the Theory of Roughness." *J. Am. Water Work Assoc.*, 73(5), 274-280.

Laumanns, M., Thiele, L., Zitzler, E., and Deb, K. (2002). "Archiving With Guaranteed Convergence And Diversity In Multi-objective Optimization." *Genetic and Evolutionary Computation Conference (GECCO '02)*, W. B. Langdon, E. Cantú-Paz, K. E. Mathias, R. Roy, D. Davis, R. Poli, K. Balakrishnan, V. G. Honavar, G. Rudolph, J. Wegener, L. Bull, M. A. Potter, A. C. Schultz, J. F. Miller, E. Burke, and N. Jonoska, eds., Morgan Kaufmann Publishers Inc., New York, NY, USA, 439-447.

Lee, G., Tachikawa, Y., and Takara, K. (2011). "Comparison of model structural uncertainty using a multi-objective optimisation method." *Hydrological Processes*, 25(17), 2642-2653.

Levy, H. (1992). "Stochastic-Dominance and Expected Utility - Survey and Analysis." *Management Science*, 38(4), 555-593.

Li, H., and Zhang, Q. (2009). "Multiobjective Optimization Problems With Complicated Pareto Sets, MOEA/D and NSGA-II." *IEEE Transactions on Evolutionary Computation*, 13(2), 284-302.

Lippai, I., Heaney, J. P., and Laguna, M. (1999). "Robust water system design with commercial intelligent search optimizers." *Journal of Computing in Civil Engineering*, 13(3), 135-143.

Lobo, F. G., and Goldberg, D. E. (2004). "The parameter-less genetic algorithm in practice." *Information Sciences*, 167(1-4), 217-232.

Lobo, F. G., Lima, C. F., and Michalewicz, Z. (2007). *Parameter Setting in Evolutionary Algorithms*, Springer Publishing Company, Incorporated.

Madsen, H. (2000). "Automatic calibration of a conceptual rainfall-runoff model using multiple objectives." *Journal of Hydrology*, 235(3-4), 276-288.

Madsen, H. (2003). "Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives." *Advances in Water Resources*, 26(2), 205-216.

Maier, H. R., Simpson, A. R., Zecchin, A. C., Foong, W. K., Phang, K. Y., Seah, H. Y., and Tan, C. L. (2003). "Ant Colony Optimization for Design of Water Distribution Systems." *Journal of Water Resources Planning and Management*, 129(3), 200-209.

Matott, L. S., Tolson, B. A., and Asadzadeh, M. (2012). "A benchmarking framework for simulation-based optimization of environmental models." *Environmental Modelling & Software*, 35, 19-30.

Minsker, B. (2005). *Genetic Algorithms*, CRC press.

Montalvo, I., Izquierdo, J., Pérez, R., and Iglesias, P. L. (2008). "A diversity-enriched variant of discrete PSO applied to the design of water distribution networks." *Engineering Optimization*, 40(7), 655-668.

Mostaghim, S., Teich, J., and Tyagi, A. (2002). "Comparison of data structures for storing Pareto-sets in MOEAs." *Congress on Evolutionary Computation (Cec'02)*, IEEE press, Honolulu, Hawaii, 843-848.

Moussa, R., and Chahinian, N. (2009). "Comparison of different multi-objective calibration criteria using a conceptual rainfall-runoff model of flood events." *Hydrology and Earth System Sciences*, 13(4), 519-535.

Mugunthan, P., Shoemaker, C. A., and Regis, R. G. (2005). "Comparison of function approximation, heuristic, and derivative-based methods for automatic calibration of computationally expensive groundwater bioremediation models." *Water Resources Research*, 41(11), 1-17.

Nash, J. E., and Sutcliffe, J. V. (1970). "River flow forecasting through conceptual models, part I. A discussion of principles." *Journal of Hydrology*, 10(3), 282– 290.

Nicklow, J., Reed, P., Savic, D., Dessalegne, T., Harrell, L., Chan-Hilton, A., Karamouz, M., Minsker, B., Ostfeld, A., Singh, A., and Zechman, E. (2010). "State of the Art for Genetic Algorithms and Beyond in Water Resources Planning and Management." *ASCE, Journal of Water Resources Planning and Management*., 136(4), 412–432.

Okabe, T., Jin, Y. C., Olhofer, M., and Sendhoff, B. (2004). "On test functions for evolutionary multi-objective optimization." *Parallel Problem Solving from Nature - PPSN VIII*, X. Yao, E. Burke, J. A. Lozano, J. Smith, J. J. Merelo-Guervos, J. A. Bullinaria, J. Rowe, P. Tino, A. Kaban, and H. P. Schwefel eds., Springer, 792-802.

Ostfeld, A., Salomons, E., Ormsbee, L., Uber, J., Bros, C., Kalungi, P., Burd, R., Zazula-Coetzee, B., Belrain, T., Kang, D., Lansey, K., Shen, H., McBean, E., Wu, Z., Walski, T., Alvisi, S., Franchini, M., Johnson, J., Ghimire, S., Barkdoll, B., Koppel, T., Vassiljev, A., Kim, J., Chung, G., Yoo, D., Diao, K., Zhou, Y., Li, J., Liu, Z., Chang, K., Gao, J., Qu, S., Yuan, Y., Prasad, T., Laucelli, D., Vamvakeridou Lyroudia, L., Kapelan, Z., Savic, D., Berardi, L., Barbaro, G., Giustolisi, O., Asadzadeh, M., Tolson, B., and McKillop, R. (In press). "The Battle of the Water Calibration Networks (BWCN)." *Journal of Water Resources Planning and Management*, 0(ja), 137.

Parajka, J., Merz, R., and Bloeschl, G. (2007). "Uncertainty and multiple objective calibration in regional water balance modelling: case study in 320 Austrian catchments." *Hydrological Processes*, 21(4), 435-446.

Parmee, I. C., and Purchase, G. (1994). "The development of a directed genetic search technique for heavily constrained design spaces." *Adaptive Computing in Engineering Design and Control*, I. C. Parmee, ed.University of Plymouth, Plymouth, UK.

Parsopoulos, K. E., and Vrahatis, M. N. (2002). "Particle swarm optimization method in multiobjective problems." *The 2002 ACM symposium on Applied computing*, ACM, Madrid, Spain, 603-607.

Perelman, L., and Ostfeld, A. (2007). "An adaptive heuristic cross-entropy algorithm for optimal design of water distribution systems." *Engineering Optimization*, 39(4), 413-428.

Perelman, L., Ostfeld, A., and Salomons, E. (2008). "Cross Entropy multiobjective optimization for water distribution systems design." *Water Resources Research*, 44(9).

Ponweiser, W., Wagner, T., Biermann, D., Vincze, M. (2008). "Multiobjective optimization on a limited budget of evaluations using model-assisted S-metric selection." *In Parallel Problem Solving from Nature – PPSN X*, Rudolph, G., Jansen, T., Lucas, S., Poloni, C., and N. Beume eds., vol. 5199, pp. 784–794. Springer, Heidelberg

Reca, J., and Martinez, J. (2006). "Genetic algorithms for the design of looped irrigation water distribution networks." *Water Resources Research*, 42(5).

Reca, J., Martínez, J., Gil, C., and Baños, R. (2008). "Application of Several Meta-Heuristic Techniques to the Optimization of Real Looped Water Distribution Networks." *Water Resources Management*, 22(10), 1367-1379.

Reed, P., Minsker, B., and Valocchi, A. J. (2000). "Cost-effective long-term groundwater monitoring design using a genetic algorithm and global mass interpolation." *Water Resources Research*, 36(12), 3731-3741.

Reed, P. M., Hadka, D., Herman, J. D., Kasprzyk, J. R., and Kollat, J. B. (In press). "Evolutionary multiobjective optimization in water resources: The past, present, and future." *Advances in Water Resources*.

Rogers, A., and Prugel-Bennett, A. (1999). "Generic drift in genetic algorithm selection schemes." *IEEE Transactions on Evolutionary Computation*, 3(4), 298-303.

Sato, H., Aguirre, H. E., and Tanaka, K. (2004). "Local dominance using polar coordinates to enhance multiobjective evolutionary algorithms." *Congress on Evolutionary Computation (CEC '04)*, IEEE press, 188-195.

Savic, D. A., and Walters, G. A. (1997). "Genetic Algorithms for Least-Cost Design of Water Distribution Networks." *Journal of Water Resources Planning and Management*, 123(2), 67-77.

Schaake, J. C., and Lai, F. H. (1969). *Linear Programming and Dynamic Programming Application to Water Distribution Network Design*, M.I.T. Hydrodynamics Laboratory.

Schaffer, J. D. (1984). "Some experiments in machine learning using vector evaluated genetic algorithms (artificial intelligence, optimization, adaptation, pattern recognition)." PhD, Vanderbilt University, Vanderbilt University, Nashville, TN, USA.

Schaffer, J. D., and Morishima, A. (1987). "An adaptive crossover distribution mechanism for genetic algorithms." *Second International Conference on Genetic Algorithms and their application*, J. J. Grefenstette, ed., L. Erlbaum Associates Inc., Cambridge, Massachusetts, United States, 36-40.

Schoups, G., Hopmans, J. W., Young, C. A., Vrugt, J. A., and Wallender, W. W. (2005). "Multi-criteria optimization of a regional spatially-distributed subsurface water flow model." *Journal of Hydrology*, 311(1-4), 20-48.

Simpson, A., Dandy, G., and Murphy, L. (1994). "Genetic Algorithms Compared to Other Techniques for Pipe Optimization." *Journal of Water Resources Planning and Management*, 120(4), 423-443.

Smith, K. I., Everson, R. M., Fieldsend, J. E., Murphy, C., and Misra, R. (2008). "Dominance-Based Multiobjective Simulated Annealing." *IEEE Transactions on Evolutionary Computation*, 12(3), 323-342.

Sorooshian, S., Duan, Q. Y., and Gupta, V. K. (1993). "Calibration of Rainfall-Runoff Models - Application of Global Optimization to the Sacramento Soil-Moisture Accounting Model." *Water Resources Research*, 29(4), 1185-1194.

Srinivas, N., and Deb, K. (1994). "Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms." *Evolutionary Computation*, 2(3), 221-248.

Srinivasa, K. G., Venugopal, K. R., and Patnaik, L. M. (2007). "A self-adaptive migration model genetic algorithm for data mining applications." *Information Sciences*, 177(20), 4295-4313.

Suribabu, C. R., and Neelakantan, T. R. (2006). "Design of water distribution networks using particle swarm optimization." *Urban Water Journal*, 3(2), 111-120.

Talbi, E. G., Rahoual, M., Mabed, M. H., and Dhaenens, C. (2001). "A hybrid evolutionary approach for multicriteria optimization problems: Application to the Flow Shop." *Evolutionary Multi-Criterion Optimization*, E. Zitzler, K. Deb, L. Thiele, C. A. Coello Coello, and D. Corne, eds., Springer, 416-428.

Tang, Y., Reed, P., and Wagener, T. (2006). "How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration?" *Hydrology and Earth System Sciences*, 10(2), 289-307.

Tang, Y., Reed, P. M., and Kollat, J. B. (2007). "Parallelization strategies for rapid and robust evolutionary multiobjective optimization in water resources applications." *Advances in Water Resources*, 30(3), 335-353.

Tolson, B. A., Maier, H. R., Simpson, A. R., and Lence, B. J. (2004). "Genetic algorithms for reliability-based optimization of water distribution systems." *J. Water Resour. Plan. Manage.-ASCE*, 130(1), 63-72.

Tolson, B. A., and Shoemaker, C. A. (2007). "Dynamically dimensioned search algorithm for computationally efficient watershed model calibration." *Water Resources Research*, 43(1), W01413.

Tolson, B. A., Asadzadeh, M., Maier, H. R., and Zecchin, A. C. (2008). "A New Algorithm for Water Distribution System Optimization: Discrete Dynamically Dimensioned Search." *World Environmental and Water Resources Congress 2008*, R. W. Babcock, and R. Walton, eds., American Society of Civil Engineers, Honolulu, Hawaii, 1-10.

Tolson, B. A., Asadzadeh, M., Maier, H. R., and Zecchin, A. (2009). "Hybrid discrete dynamically dimensioned search (HD-DDS) algorithm for water distribution system design optimization." *Water Resources Research*, 45(12), W12416.

Van Veldhuizen, D. A., and Lamont, G. B. (1998). "Evolutionary Computation and Convergence to a Pareto Front." *Late Breaking Papers at the Genetic Programming 1998 Conference*, J. R. Koza, ed., Stanford University Bookstore, University of Wisconsin, Madison, Wisconsin, USA.

Van Veldhuizen, D. A. (1999). "Multiobjective evolutionary algorithms: classifications, analyses, and new innovations." Air Force Institute of Technology, Wright Patterson AFB, OH, USA.

Varma, K. V. K., Narasimhan, S., and Bhallamudi, S. M. (1997). "Optimal Design of Water Distribution Systems Using an NLP Method." *Journal of Environmental Engineering*, 123(4), 381-388.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S. (2003). "Effective and efficient algorithm for multiobjective optimization of hydrologic models." *Water Resources Research*, 39(8).

Vrugt, J. A., and Robinson, B. A. (2007). "Improved evolutionary optimization from genetically adaptive multimethod search." *Proceedings of the National Academy of Sciences of the United States of America*, 104(3), 708-711.

Walski, T. M. (2001). "The wrong paradigm - Why water distribution optimization doesn't work." *Journal of Water Resources Planning and Management*, 127(4), 203-205.

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

While, L. (2005). "A New Analysis of the LebMeasure Algorithm for Calculating Hypervolume." *Evolutionary Multi-Criterion Optimization*, C. A. Coello Coello, A. Hernández Aguirre, and E. Zitzler, eds., Springer, 326-340.

Wolpert, D. H., and Macready, W. G. (1997). "No free lunch theorems for optimization." *IEEE Transactions on Evolutionary Computation*, 1(1), 67-82.

Wu, Z. Y., Boulos, P. F., Orr, C. H., and Ro, J. J. (2001). "Using genetic algorithms to rehabilitate distribution systems." *J. Am. Water Work Assoc.*, 93(11), 74-85.

Wu, Z. Y., and Simpson, A. R. (2002). "A self-adaptive boundary search genetic algorithm and its application to water distribution systems." *Journal of Hydraulic Research*, 40(2), 191-203.

Xia, Y., Pittman, A. J., Gupta, H. V., Leplastrier, M., Henderson-Sellers, A., and Bastidas, L. A. (2002). "Calibrating a land surface model of varying complexity using multicriteria methods and the Cabauw dataset." *Journal of Hydrometeorology*, 3(2), 181-194.

Yang, I. T. (2007). "Pareto archived PSO optimization for time-cost tradeoff analysis." *IEEE Congress on Evolutionary Computation*, IEEE press, Singapore, 3329-3334.

Yapo, P. O., Gupta, H. V., and Sorooshian, S. (1998). "Multi-objective global optimization for hydrologic models." *Journal of Hydrology*, 204(1-4), 83-97.

Zecchin, A., Maier, H., Simpson, A., Leonard, M., and Nixon, J. (2007). "Ant Colony Optimization Applied to Water Distribution System Design: Comparative Study of Five Algorithms." *Journal of Water Resources Planning and Management*, 133(1), 87-92.

Zecchin, A. C., Simpson, A. R., Maier, H. R., and Nixon, J. B. (2005). "Parametric study for an ant algorithm applied to water distribution system optimization." *IEEE Transactions on Evolutionary Computation*, 9(2), 175-191.

Zecchin, A. C., Simpson, A. R., Maier, H. R., Leonard, M., Roberts, A. J., and Berrisford, M. J. (2006). "Application of two ant colony optimisation algorithms to water distribution system optimisation." *Math. Comput. Model.*, 44(5-6), 451-468.

Zhang, Q., and Suganthan, P. N. (2008). "Final Report on CEC'09 MOEA Competition." School of Computer Science and Electrical Engineering, University of Essex, CES-887.

Zhang, Q., Zhou, A., Zhao, S., Suganthan, P. N., W., L., and Tiwari, S. (2008). "Multiobjective optimization test instances for the CEC 2009 special session and competition." *IEEE Congress on Evolutionary Computation (IEEE CEC 2009)*, University of Essex and Nanyang Technological University.

Zhang, X., Srinivasan, R., and Liew, M. V. (2010). "On the use of multi-algorithm, genetically adaptive multi-objective method for multi-site calibration of the SWAT model." *Hydrological Processes*, 24(8), 955-969.

Zitzler, E., and Thiele, L. (1998). "Multiobjective optimization using evolutionary algorithms - A comparative case study." *Parallel Problem Solving from Nature - Ppsn V*, A. E. Eiben, T. Bäck, M. Schoenauer, and H.-P. Schwefel, eds., Springer, 292-301.

Zitzler, E., Deb, K., and Thiele, L. (2000). "Comparison of Multiobjective Evolutionary Algorithms: Empirical Results." *Evolutionary Computation*, 8(2), 173-195.

Zitzler, E., Laumanns, M., and Thiele, L. (2001). "{SPEA2}: Improving the strength pareto evolutionary algorithm for multiobjective optimization." *Evolutionary Methods for Design Optimization and Control with Applications to Industrial Problems*, K. C. Giannakoglou, D. T. Tsahalis, J. Periaux, K. D. Papailiou, and T. Fogarty, eds., International Center for Numerical Methods in Engineering, Athens, Greece, 95--100.

Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C. M., and da Fonseca, V. G. (2003). "Performance assessment of multiobjective optimizers: An analysis and review." *IEEE Transactions on Evolutionary Computation*, 7(2), 117-132.

Appendix A

Pseudo Codes

REQUIRES: computational budget NFE, perturbation size r (default: 0.2), vectors of discrete options for all D decision variables $\boldsymbol{x}^{max}$, note that $\boldsymbol{x}^{min} = [1, 1, \dots, 1]$, initial solution $\boldsymbol{x}^0 = [x^1, \dots, x^D]$

STEP 1. Set solution evaluation counter $i = 1$, and evaluate objective f at initial solution, $f(\boldsymbol{x}^0)$:

- $f_{best} = f(\boldsymbol{x}^0)$, and $\boldsymbol{x}^{best} = \boldsymbol{x}^0$

STEP 2. Randomly select J of the D decision variables for inclusion in neighborhood, $\{N\}$:

- calculate probability each decision variable is included in $\{N\}$: $p = 1 - ln(i)/ln(\mathrm{NFE}_{\mathrm{total}})$
- FOR $d = 1, \dots D$ decision variables, add d to $\{N\}$ with probability p
- IF $\{N\}$ is empty, select one random d for $\{N\}$

STEP 3. FOR $j = 1, \dots, J$ decision variables in $\{N\}$, perturb x_j^{best} by sampling from a discrete probability distribution. This discrete distribution approximates a normal probability distribution:

- Sample a standard normal random variable, $\mathbf{N}(0,1)$
- $x_j^{new} = x_j^{best} + \sigma_j \mathbf{N}(0,1)$, where $\sigma_j = r\left(x_j^{max} - x_j^{min}\right)$
- IF $x_j^{new} < \left(x_j^{min} - 0.5\right)$, reflect perturbation at $x_j^{min} - 0.5$:
  - $x_j^{new} = \left(x_j^{min} - 0.5\right) + \left(\left(x_j^{min} - 0.5\right) - x_j^{new}\right) = 2x_j^{min} - x_j^{new} - 1$
  - IF $x_j^{new} > \left(x_j^{max} + 0.5\right)$, set $x_j^{new} = \mathrm{x_j^{min}}$
- ELSEIF $x_j^{new} > \left(x_j^{max} + 0.5\right)$, reflect perturbation at $x_j^{max} + 0.5$:
  - $x_j^{new} = \left(x_j^{max} + 0.5\right) - \left(x_j^{new} - \left[x_j^{max} + 0.5\right]\right) = 2x_j^{max} - x_j^{new} + 1$
  - IF $x_j^{new} < \left(x_j^{min} - 0.5\right)$, set $x_j^{new} = x_j^{max}$
- Round x_j^{new} to the nearest integer representing the discrete option number
- IF $x_j^{new} = x_j^{best}$, sample a discrete uniform random variable, $\mathbf{U}\left(x_j^{max}, x_j^{min}\right)$, until $x_j^{new} \neq x_j^{best}$

STEP 4. Evaluate $f(\boldsymbol{x}^{new})$ and update current best solution if necessary:

- IF $f(\boldsymbol{x}^{new}) \leq f_{best}$, update new best solution:
  - $f_{best} = f(\boldsymbol{x}^{new})$ and $\boldsymbol{x}^{best} = \boldsymbol{x}^{new}$

STEP 5. Update solution evaluation counter, $i = i + 1$, and check stopping criterion:

- IF $(p < 1/D)$ OR IF $(i =$ NFE), STOP, save f_{best} and $\boldsymbol{x}^{best}$
- ELSE, set $\boldsymbol{x}^{new} = \boldsymbol{x}^{best}$, and go to STEP 3

Figure A- 1 Discrete Dynamically Dimensioned Search (Discrete DDS) algorithm. $\mathbf{N}(0,1)$ generates random numbers from the standard Normal Distribution. $\mathbf{U}(lb, ub)$ generates random numbers from the Uniform distribution in the (lb, ub) interval.

REQUIRES: new solution $\boldsymbol{x}$, current best solution $\boldsymbol{x}^{best}$ and its sum of pressure violations $g(\boldsymbol{x}^{best})$, minimum required head $\boldsymbol{h}^{min}$ at all nodes, and solution with all pipes at their maximum diameter $\boldsymbol{x}^{max}$

- Evaluate $\text{cost}(\boldsymbol{x})$
- IF $\text{cost}(x) > \text{cost}(x^{best})$ AND $g(\boldsymbol{x}^{best}) = 0$
  - $f(\boldsymbol{x}) = \text{cost}(\boldsymbol{x})$ (no EPANET2 run)
- ELSE
  - run EPANET2 and calculate $g(\boldsymbol{x}) = \sum_{j=1}^{\#nodes} \max\left[0, h_j^{min} - h_j(\boldsymbol{x})\right]$
  - IF $g(\boldsymbol{x}) = 0$ then $f(\boldsymbol{x}) = \text{cost}(\boldsymbol{x})$
  - IF $g(\boldsymbol{x}) > 0$ then $f(\boldsymbol{x}) = \text{cost}(\boldsymbol{x}^{max}) + g(\boldsymbol{x})$

Figure A- 2 Evaluating the objective function in HD-DDS. Note that $\text{cost}(\boldsymbol{x})$ calculates the cost of the network based on the diameter and length of pipes, $f(\boldsymbol{x})$ is the objective function, $g(\boldsymbol{x})$ is the summation of pressure violations at all nodes in the network.

REQUIRES: initial solution $\boldsymbol{x}^{ini}$, computational budget NFE

- $\boldsymbol{x}^{best} = \boldsymbol{x}^{ini}$
- $\boldsymbol{x}^{pbest} = [\ \ ]$
- $i = 0$
- WHILE L_1 is not converged $\left(\text{x}^{\text{best}} \neq \text{x}^{\text{pbest}}\right)$:
  - $\boldsymbol{x}^{pbest} = \boldsymbol{x}^{best}$
  - FOR $j = 1$ to D decision variables
    - $\boldsymbol{x}^{test} = \boldsymbol{x}^{best}$
    - WHILE pipe j can have diameter decreased from x_j^{best} (and thus reduced cost):
      - decrease the diameter of pipe j by 1 discrete option, ${x_j}^{test} = {x_j}^{best} - 1$
      - evaluate objective function $f(\boldsymbol{x}^{test})$
      - IF $\boldsymbol{x}^{test}$ is feasible, update $\boldsymbol{x}^{best}$ by $\boldsymbol{x}^{test}$ and $i = i + 1$
      - IF $i =$ NFE, STOP L_1
      - BREAK inner loop when $\boldsymbol{x}^{test}$ is infeasible because smaller pipe j would also be infeasible

STOP L_1: Return $\boldsymbol{x}^{best}$, $f\left(\boldsymbol{x}^{best}\right)$, i, and whether L_1 is converged to a local solution such that no further improvement is possible by changing one pipe at a time

Figure A- 3 Outline of local search L_1 for constrained WDS optimization problem that can identify a local minimum such that no further improvement is possible by changing one decision variable (pipe) at a time

REQUIRES: initial solution $\boldsymbol{x}^{ini}$, computational budget NFE

- $\boldsymbol{x}^{best} = \boldsymbol{x}^{ini}$
- $\boldsymbol{x}^{pbest} = [\quad]$
- $i = 0$
- WHILE L_2 is not converged $\left(\mathrm{x^{best}} \neq \mathrm{x^{pbest}}\right)$
  - $\boldsymbol{x}^{pbest} = \boldsymbol{x}^{best}$
  - $r = \max\left(x_1^{best}, x_2^{best}, \ldots, x_D^{best}\right)$ the maximum number of pipe diameter reductions to consider
  - FOR $l = 1\ to\ r$
    - FOR $j = 1\ to\ D$ decision variables
      - $\boldsymbol{x}^{test} = \boldsymbol{x}^{pbest}$
      - IF $\left(x_j^{pbest} - l\right) >$ minimum diameter option number of pipe j (and thus a reduced cost)
        - FOR $k = 1\ to\ D$
          - $\boldsymbol{x}^{test} = \boldsymbol{x}^{pbest}$
          - ${x_j}^{test} = {x_j}^{pbest} - l$
          - IF $k \neq j$ evaluate increased diameters for pipe k starting at largest diameter
            - Set ${x_k}^{test} = {x_k}^{max} + 1$
            - WHILE ${x_k}^{test}$ can have diameter decreased by one option and ${x_k}^{test} > {x_k}^{pbest} + 1$:
              - decrease the diameter of pipe k by 1 discrete option, ${\boldsymbol{x}_k}^{test} = {\boldsymbol{x}_k}^{best} - 1$
              - calculate cost($\boldsymbol{x}^{test}$). no EPANET run. ($\boldsymbol{x}^{test}$ differs from $\boldsymbol{x}^{pbest}$ only in pipe j^{th}, k^{th})
              - IF cost($\boldsymbol{x}^{test}$) < cost$\left(\boldsymbol{x}^{best}\right)$
                - evaluate objective function $f(\boldsymbol{x}^{test})$
                - IF $\boldsymbol{x}^{test}$ is feasible, update $\boldsymbol{x}^{best}$ by $\boldsymbol{x}^{test}$ and $i = i + 1$
                - IF $i =$ NFE, STOP L_2
                - BREAK inner loop when $\boldsymbol{x}^{test}$ infeasible because smaller pipe k would be infeasible

STOP L_2: Return $\boldsymbol{x}^{best}$, F$\left(\boldsymbol{x}^{best}\right)$, i, and whether L_2 is converged to a local solution such that no further improvement is possible by changing two pipes at a time

Figure A- 4 Outline of local search L_2 for constrained WDS optimization problem that can identify a local minimum such that no further improvement is possible by changing two decision variables (pipe) at a time

REQUIRES: case study and network inputs (layout, available pipe diameters, pipe costs, etc.), maximum number of solution evaluations, $\mathrm{NFE}_{\mathrm{total}}$, $\mathrm{NFE} = \mathrm{NFE}_{\mathrm{total}}$

STEP 1. Perform global search with discrete DDS:

- $\boldsymbol{x}_A^{best} = [\,]$
- $\boldsymbol{x}_B^{best} = [\,]$
- Run the Discrete DDS with NFE as number of solution evaluation limit
- Return the best solution of this step, $\boldsymbol{x}^{best}$ and the corresponding objective function value, $f(\boldsymbol{x}^{best})$
- update available computational budget: $\mathrm{NFE} = (\mathrm{NFE} - \text{computational effort in this step})$

STEP 2. Perform local search L_1 that changes current best solution by only one pipe at a time:

- Run L_1 initialized at $\boldsymbol{x}^{best}$
- Check if update needed for best solution, $\boldsymbol{x}^{best}$, and $F(\boldsymbol{x}^{best})$
- update available computational budget: $\mathrm{NFE} = (\mathrm{NFE} - \text{computational effort in } \mathrm{L}_1)$
- IF $\boldsymbol{x}_A^{best}$ is empty
  - set $\boldsymbol{x}_A^{best} = \boldsymbol{x}^{best}$
- ELSE IF $\boldsymbol{x}_B^{best}$ is empty
  - set $\boldsymbol{x}_B^{best} = \boldsymbol{x}^{best}$
- IF NFE ≤ 0
  - STOP HD-DDS
- ELSE IF $\boldsymbol{x}_B^{best}$ is not empty
  - Go to STEP 4
- ELSE, Go to STEP 3

STEP 3. Perform the second independent global search with discrete DDS followed by L_1:

- Go to STEP 1.

STEP 4. Perform local search L_2 that changes current best solution by only two-pipes at a time:

- Run L_2 initialized at the best of $\boldsymbol{x}_A^{best}$ and $\boldsymbol{x}_B^{best}$ (this is referred to as $\mathrm{L'}_2$)
- Check if update needed for $\boldsymbol{x}_A^{best}$, $\boldsymbol{x}_B^{best}$, $\boldsymbol{x}^{best}$, and $f(\boldsymbol{x}^{best})$
- update available computational budget: $\mathrm{NFE} = (\mathrm{NFE} - \text{computational effort in } \mathrm{L'}_2)$
- IF NFE≤ 0 or $\boldsymbol{x}_B^{best} = \boldsymbol{x}_A^{best}$
  - STOP HD-DDS
- ELSE, Go to STEP 5

STEP 5. Perform another L_2 local search:

- Run L_2 initialized at the worst of $\boldsymbol{x}_A^{best}$ and $\boldsymbol{x}_B^{best}$ (this is referred to as $\mathrm{L''}_2$)
- Check if update needed for $\boldsymbol{x}_A^{best}$, $\boldsymbol{x}_B^{best}$, $\boldsymbol{x}^{best}$, and $f(\boldsymbol{x}^{best})$
- update available computational budget: $\mathrm{NFE} = (\mathrm{NFE} - \text{computational effort in } \mathrm{L''}_2)$

STOP HD-DDS

- Return $\boldsymbol{x}^{best}$ as the best of $\boldsymbol{x}_A^{best}$ and $\boldsymbol{x}_B^{best}$, $f(\boldsymbol{x}^{best})$ and total solutions evaluated $(\mathrm{NFE}_{\mathrm{total}} - \mathrm{NFE})$
- Report if $\boldsymbol{x}^{best}$ is local optimum with respect to L_1 or L_2.

Figure A- 5 The Hybrid Discrete Dynamically Dimensioned Search (HD-DDS) algorithm.

REQUIRES: computational budget NFE, vectors of decision variable boundaries $\boldsymbol{x}^{min}$, $\boldsymbol{x}^{max}$, vector of decision variable type $\boldsymbol{x}^{type}$: 0s for real−, 1s for discrete − valued decision variable, number of decision variables n, number of objective functions m, q number of initial solutions, selection metric type (either RND, CD, HVC1, HVC2, or CHC)

- $\text{PS}^{\text{a}} = \varnothing$, $\text{PF}^{\text{a}} = \varnothing$
- FOR $i = 1, \dots, q$
  - $\boldsymbol{x} = \mathbf{U}(\boldsymbol{x}^{min}, \boldsymbol{x}^{max})$
  - $\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \dots, f_m(\boldsymbol{x}))$
  - $[\text{PS}^{\text{a}}, \text{PF}^{\text{a}}] = \textbf{Archive}\,(\text{PS}^{\text{a}}, \text{PF}^{\text{a}}, \boldsymbol{x}, \boldsymbol{f}(\boldsymbol{x}))$
- Using selection metric type, calculate selection metric values $(\boldsymbol{z})$ for solutions in PS^{a} and PF^{a}
- $\boldsymbol{x}^{cur} = \textbf{Select}\,(\text{PS}^{\text{a}}, \text{PF}^{\text{a}}, \boldsymbol{z})$
- FOR $i = 1, \dots, \text{NFE} - q$
  - $p = 1 - ln(i) \,/\, ln(NFE - q)$
  - $\boldsymbol{x}^{new} = \textbf{Perturb}\,(\boldsymbol{x}^{cur}, p, n, \boldsymbol{x}^{min}, \boldsymbol{x}^{max}, \boldsymbol{x}^{type}, r)$
  - $\boldsymbol{f}(\boldsymbol{x}^{new}) = (f_1(\boldsymbol{x}^{new}), \dots, f_m(\boldsymbol{x}^{new}))$
  - IF $\boldsymbol{x}^{cur} \preccurlyeq \boldsymbol{x}^{new}$ (if $\boldsymbol{x}^{new}$ is not dominated by $\boldsymbol{x}^{cur}$)
    - $[\text{PS}^{\text{a}}, \text{PF}^{\text{a}}] = \textbf{Archive}\,(\text{PS}^{\text{a}}, \text{PF}^{\text{a}}, \boldsymbol{x}^{new}, \boldsymbol{f}(\boldsymbol{x}^{new}))$
    - IF $\boldsymbol{x}^{new}$, is archived
      - $\boldsymbol{x}^{cur} = \boldsymbol{x}^{new}$
      - Calculate metric values $(\boldsymbol{z})$ for solutions in PS^{a} and PF^{a}
    - ELSE
      - $\boldsymbol{x}^{cur} = \textbf{Select}\,(\text{PS}^{\text{a}}, \text{PF}^{\text{a}}, \boldsymbol{z})$
  - ELSE
    - $\boldsymbol{x}^{cur} = \textbf{Select}\,(\text{PS}^{\text{a}}, \text{PF}^{\text{a}}, \boldsymbol{z})$

END procedure

Figure A- 6 Pareto Archived Dynamically Dimensioned Search (PA-DDS) Algorithm. $\mathbf{U}(lb, ub)$ generates random numbers from the Uniform distribution in the (lb, ub) interval.

REQUIRES: archived solutions PS^{a} and objective function values PF^{a}, new solution $\boldsymbol{x}^{new}$ and its objective function values $\boldsymbol{f}(\boldsymbol{x}^{new})$

- IF $\text{PS}^{\text{a}} = \varnothing$
  - $\text{PS}^{\text{a}} = \boldsymbol{x}^{new}$, $\text{PF}^{\text{a}} = \boldsymbol{f}(\boldsymbol{x}^{new})$
- ELSE
  - FOR $i = 1, \dots, \|\text{PS}^{\text{a}}\|$
    - IF $\boldsymbol{x}^{i} \preccurlyeq \boldsymbol{x}^{new}$ (Solution i in PS^{a} is dominated by the new solution)
      - RETURN PS^{a} and PF^{a} unchanged
    - ELSEIF $\boldsymbol{x}^{new} \preccurlyeq \boldsymbol{x}^{i}$
      - $\text{PS}^{\text{a}} = \text{PS}^{\text{a}} \backslash \boldsymbol{x}^{i}$, $\text{PF}^{\text{a}} = \text{PF}^{\text{a}} \backslash \boldsymbol{f}(\boldsymbol{x}^{i})$
  - $\text{PS}^{\text{a}} = \text{PS}^{\text{a}} \cup \boldsymbol{x}^{new}$, $\text{PF}^{\text{a}} = \text{PF}^{\text{a}} \cup \boldsymbol{f}(\boldsymbol{x}^{new})$
- RETURN PS^{a} and PF^{a}

END procedure

Figure A- 7 Archive strategy of PA-DDS.

REQUIRES: archived solutions PS^{a} along with corresponding objective function values PF^{a} and selection metric values $\boldsymbol{z}$

- $n = \|\text{PS}^{\text{a}}\|$
- SORT PS^{a}, PF^{a} in ascending order of metric values $\boldsymbol{z}$
- $z_0^{cum} = 0$
- FOR $i = 1, \dots, n$
  - $z_i^{cum} = z_{i-1}^{cum} + z_i$
- $t = \mathbf{U}(0,1) \times z_n^{cum}$
- FOR $i = 1, \dots, n$
  - IF $z_i^{cum} \geq t$
    - RETURN $\boldsymbol{x}^{cur} = \text{PS}_i^{\text{a}}, \boldsymbol{f}_{cur} = \text{PF}_i^{\text{a}}$

END procedure

Figure A- 8 Select function based on roulette wheel selection. $\mathbf{U}(0,1)$ generates random numbers from Uniform distribution in the $(0,1)$ interval.

REQUIRES: solution $\boldsymbol{x}^{cur}$, perturbation probability p, number of decision variables n, $\boldsymbol{x}^{min}$, $\boldsymbol{x}^{max}$, vector of decision variable type $\boldsymbol{x}^{type}$: 0s for real, 1s for discrete − valued decision variable, , vector of decision variable type $\boldsymbol{x}^{type}$: 0s for real, 1s for discrete − valued decision variable, perturbation size r (default: 0.2)

- $\boldsymbol{d} = \varnothing$
- FOR $i = 1, \dots, n$
  - IF $\mathbf{U}(0,1) \leq p$
    - $\boldsymbol{d} = \boldsymbol{d} \cup \{i\}$
- FOR $i = 1, \dots, \|\boldsymbol{d}\|$
  - $j = \boldsymbol{d}(i)$
  - $\sigma_j = r \times (x_j^{max} - x_j^{min})$
  - $x_j^{new} = x_j^{cur} + \sigma_j \mathbf{N}(0,1)$
  - IF $x^{type}{}_j = 0$
    - IF $x_j^{new} < x_j^{min}$
      - IF $\mathrm{U}(0,1) \geq 0.5 \rightarrow x_j^{new} = x_j^{min}$
      - ELSE $\rightarrow x_j^{new} = x_j^{min} + \left(x_j^{min} - x_j^{new}\right)$
      - IF $x_j^{new} > x_j^{max} \rightarrow x_j^{new} = x_j^{min}$
    - ELSEIF $x_j^{new} > x_j^{max}$
      - IF $\mathrm{U}(0,1) \geq 0.5 \rightarrow x_j^{new} = x_j^{max}$
      - ELSE $\rightarrow x_j^{new} = x_j^{max} - \left(x_j^{new} - x_j^{max}\right)$
      - IF $x_j^{new} < x_j^{min} \rightarrow x_j^{new} = x_j^{max}$
  - ELSEIF $x^{type}{}_j = 1$
    - IF $x_j^{new} < \left(x_j^{min} - 0.5\right)$
      - IF $\mathrm{U}(0,1) \geq 0.5 \rightarrow x_j^{new} = x_j^{min}$
      - ELSE $\rightarrow x_j^{new} = x_j^{min} + \left(x_j^{min} - x_j^{new}\right) - 1$
      - IF $x_j^{new} > \left(x_j^{max} + 0.5\right)$, set $x_j^{new} = x_j^{min}$
    - ELSEIF $x_j^{new} > \left(x_j^{max} + 0.5\right)$
      - IF $\mathrm{U}(0,1) \geq 0.5 \rightarrow x_j^{new} = x_j^{max}$
      - ELSE $\rightarrow x_j^{new} = x_j^{max} - \left(x_j^{new} - x_j^{max}\right) + 1$
      - IF $x_j^{new} < \left(x_j^{min} - 0.5\right)$, set $x_j^{new} = x_j^{max}$
  - Round x_j^{new} to the nearest integer representing the discrete option number
  - IF $x_j^{new} = x_j^{cur}$, sample a discrete uniform random variable, $U\left(x_j^{max}, x_j^{min}\right)$, until $x_j^{new} \neq x_j^{cur}$

RETURN $\boldsymbol{x}^{new}$

END procedure

Figure A- 9 Perturb function (search strategy) of PA-DDS. $\mathbf{U}(0,1)$ Uniform distribution in the $(0,1)$ interval. $\mathbf{N}(0,1)$ generates random numbers from Standard Normal distribution.

REQUIRES: computational budget NFE, number of decision variables n, number of objectives to be minimized m, vectors of discrete options for all n decision variables $\boldsymbol{x}^{max}$, $\boldsymbol{x}^{min} = [1, 1, \ldots, 1]$, $\boldsymbol{x}^{type} = [1, 1, \ldots, 1]$, selection metric type (CD), network inputs (layout, available pipe diameters, costs, ...)

STEP 1. Global Search

- $i = \left(\text{NFE}^{(n-1)/n}\right)$: solution evaluation count at which DDS perturbs 1 decision variable per iteration
- $q = min(5, 0.005 \times i)$:number of initial solutions
- $\boldsymbol{ND\_set} = \textbf{PA} - \textbf{DDS}\left(\boldsymbol{x}^{min}, \boldsymbol{x}^{max}, \boldsymbol{x}^{type}, n, m, i, q\right)$
- NFE = NFE $- i$

STEP 2. Local Search for Extreme Points of ND_set, (Best Solutions for Individual Objectives)

- FOR $k = 1, \ldots, \#obj$
  - find current extreme point corresponding to objective k
  - WHILE NFE > 0 and L has not yet converged for selected extreme point
    - perform local search L for selected extreme point
    - update NFE = NFE $-$ i, i is the number of solutions evaluated in this step
    - IF i $= 0$, Go to STEP4

STEP 3. Local Search for Other Solutions in ND_set

- divide first objective range to $j = \text{abs}\left(\text{NFE}/(2n)\right)$ equal intervals
- IF j < (# solutions in $\boldsymbol{ND\_set}$ that are not polished by L)
  - FOR $k = 1, \ldots j$
    - IF there is any $\boldsymbol{ND\_set}$ solution in interval j that is not polished by L, choose a solution at random, $\boldsymbol{x}^{cur}$ and perform local search L starting at $\boldsymbol{x}^{cur}$
- ELSE
  - FOR k = 1 , …, number of solutions in $\boldsymbol{ND\_set}$
    - IF solution k in $\boldsymbol{ND\_set}$ has not been polished by L
      - choose solution k as $\boldsymbol{x}^{cur}$
      - perform local search L starting at $\boldsymbol{x}^{cur}$
    - update NFE = NFE $-$ i, i is the number of solutions evaluated in this step

STEP 4. Check Termination Criteria

- IF NFE > 0 and there is any solution in $\boldsymbol{ND\_set}$ for which L is not converged
  - Go to STEP 2
- ELSE
  - Return $\boldsymbol{ND\_set}$ as the Pareto approximate front, STOP

Figure A- 10 Pseudo Code of the Hybrid PA-DDS. abs is the absolute function.

REQUIRES: initial solution $\boldsymbol{x}^{curr}$, and $F(\boldsymbol{x}^{curr})$, computational budget M, vectors with number of discrete options for all D decision variables $\boldsymbol{x}^{max}$, note that $\boldsymbol{x}^{min} = [1, 1, \dots, 1]$, set of non-dominated solutions, ND_set, solution evaluation counter, $i = 0$

STEP 1. Reduce decision variables by one option

- $\boldsymbol{x}^{ini} = \boldsymbol{x}^{cur}$
- FOR $d = 1, \dots, D$
  - $\boldsymbol{x}^{new} = \boldsymbol{x}^{cur}$
  - IF $\boldsymbol{x}^{new}(d) > \boldsymbol{x}^{min}(d)$
    - set $i = i + 1$
    - IF $i > M$, STOP L and return $\boldsymbol{ND\_set}$
    - $\boldsymbol{x}^{new}(d) = \boldsymbol{x}^{new}(d) - 1$
    - Evaluate $\boldsymbol{F}^{new} = \boldsymbol{F}(\boldsymbol{x}^{new})$
    - IF $\boldsymbol{x}^{new}$ is not dominated by $\boldsymbol{x}^{cur}$
      - Check dominance of $\boldsymbol{x}^{new}$ against $\boldsymbol{ND\_set}$ solutions
      - IF $\boldsymbol{x}^{new}$ dominates $\boldsymbol{x}^{cur}$ then $\boldsymbol{x}^{cur} = \boldsymbol{x}^{new}$, $\boldsymbol{F}^{cur} = \boldsymbol{F}^{new}$

STEP 2. Increase decision variables by one option

- $\boldsymbol{x}^{cur} = \boldsymbol{x}^{ini}$
- FOR $d = 1, \dots, D$
  - $\boldsymbol{x}^{new} = \boldsymbol{x}^{cur}$
  - IF $\boldsymbol{x}^{new}(d) < \boldsymbol{x}^{max}(d)$
    - set $i = i + 1$
    - IF $i > M$, STOP L and return $\boldsymbol{ND\_set}$
    - $\boldsymbol{x}^{new}(d) = \boldsymbol{x}^{new}(d) + 1$
    - Evaluate $\boldsymbol{F}^{new} = \boldsymbol{F}(\boldsymbol{x}^{new})$
    - IF $\boldsymbol{x}^{new}$ is not dominated by $\boldsymbol{x}^{cur}$
      - Check dominance of $\boldsymbol{x}^{new}$ against $\boldsymbol{ND\_set}$ solutions
      - IF $\boldsymbol{x}^{new}$ dominates $\boldsymbol{x}^{cur}$ then $\boldsymbol{x}^{cur} = \boldsymbol{x}^{new}$, $\boldsymbol{F}^{cur} = \boldsymbol{F}^{new}$

Figure A- 11 Local Search L to polish one solution on the approximate front