H.R. Maier, Z. Kapelan, J. Kasprzyk, L.S. Matott
Thematic issue on evolutionary algorithms in water resources
Environmental Modelling and Software, 2015; 69:222-225

Citation: Maier H.R., Kapelan Z., Kasprzyk J. and Matott L.S. (2015) Thematic issue on evolutionary algorithms in water resources, *Environmental Modelling and Software*, **69**, 222-225, DOI:10.1016/j.envsoft.2015.05.003

Preface

Thematic Issue on Evolutionary Algorithms in Water Resources

H.R. Maier, Z. Kapelan, J. Kasprzyk, L.S. Matott

Evolutionary algorithms (EAs) and other similar optimisation approaches have become very popular in the water resources research literature over the last two decades. One reason for the emergence of EAs in the literature is that they use evolutionary principles found in nature, "evolving" to find better solutions to complex water resources problems. Another reason is that evolutionary optimisation provides a natural extension to the use of simulation models, as EAs simply "bolt onto" existing models. Consequently, the resulting optimisation process is very intuitive, as the way EAs try different solutions and then learn from the outcomes of these trials is similar to the process humans adopt when manually "optimising" or adjusting solutions to problems via a simulation based approach. The only differences when EAs are used are that the decisions as to which options to try are made with the aid of evolutionary operators, rather than human judgement, intuition and experience, and that the number of options considered is much larger. Moreover, outputs of the EA process are equivalent to outputs of trusted simulation models. Therefore, the optimisation results from EAs tend to have more credibility than those obtained using alternative approaches, such as mathematical programming, since the latter generally require gross simplifications of problem representation.

Another attractive feature of EAs is that they are not necessarily prescriptive in the sense of suggesting "the" optimal solution. This is because they work with populations of solutions and therefore produce a number of near-optimal solutions, which might be similar in objective function space, but quite different in solution space. This enables consideration of factors other than those captured in the mathematical formulation of the optimisation problem when selecting the solution to be implemented. As a result of the loose coupling between the optimisation engine, which decides which parts of the solution space to explore, and the simulation model, which evaluates how well the selected solutions perform in relation to the objectives and/or whether constraints have been violated, EAs can deal with discontinuities and non-linearities with ease, as long as these have been captured appropriately in the simulation model. Another advantage of EAs is that they are well suited to multi-objective problems, as they can evolve optimal trade-offs between objectives (i.e. Pareto fronts) in a single optimisation trial.

Given the fascination and intrigue associated with the ability to use evolutionary processes to optimise water resources problems, the practicality and intuitiveness associated with being able to make use of existing simulation models and the advantage of being able to solve complex problems, it is not surprising that research involving EAs has received significant attention. This research has demonstrated the undoubted potential of EAs in the sense that they can be applied to and perform well in a wide range of application areas. In addition, significant research effort on the development and testing of different types of EAs, evolutionary operators and algorithm parameterisation has resulted in the ability to find better solutions with reduced computational effort. However, while there are pockets of research that continue to significantly push the boundaries of knowledge in this field, there is also a large amount of research that continues to re-visit the same themes. For example:

- There continue to be a large number of papers on using an ever increasing number of EA variants for solving an ever increasing number of water resources problems, with little focus on *understanding why* certain algorithm variants perform better for certain case studies than others. In addition, there is no consistency in algorithms, algorithm implementations, performance criteria and case studies in the papers. The above factors make it extremely difficult to draw conclusions that are applicable to the wider research field and enable meaningful guidelines for the application of different algorithms to be developed.
- There continue to be a large number of studies that use theoretical or very simplistic case studies. However, there are significant challenges associated with the application of EAs to real-world problems that need to be addressed in order to increase their uptake in industry.

In order to counteract potential repetition and stagnation in this field, Maier et al. (2014) identified a number of research questions that should be addressed. They suggest that the main areas in which research efforts should be directed include improving our understanding of algorithm performance and how to apply EAs to real-world problems, as summarised in Table 1. The 18 papers in this thematic issue begin to address some of these research questions, as summarised in Table 2 and discussed below.

Table 1: Summary of key research questions identified in Maier et al. (2014)

1. Research questions associated with improving our understanding of algorithm performance	
1.1	Can we develop knowledge of the fundamental ***characteristics of the problem being optimised*** at the level at which optimisation algorithms operate?
1.2	Can we develop knowledge of the ***underlying searching behaviour*** of different search methodologies?
1.3	How can we rigorously ***measure and improve the performance*** of a selected search methodology?
2. Research questions associated with applying EAs to real-world problems	
2.1	How do we best change the ***formulation*** of optimisation problems to cater to real-world problems?
2.2	What can be done to ***reduce the size of the search space*** for real-world problems?
2.3	How can ***computational efficiency be increased*** for real-world problems?
2.4	Which ***searching mechanisms*** are best for solving real-world problems?
2.5	What ***termination / convergence criteria*** are most appropriate for real-world problems?
2.6	What is the best way is to ***convey the results*** of the optimisation of real-world problems ***to decision makers*** and what is the ***role of optimisation*** in the ***decision-making process***?
2.7	What is the best way to ***take account of uncertainty*** in the optimisation of realistic systems?
2.8	What is the best way to ***implement*** optimisation algorithms for realistic systems?

Table 2: Research questions addressed in papers in this thematic issue

Paper	Research Question										
	1.1	1.2	1.3	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8
Gibbs et al. (2015)	X					X					
Zheng et al. (2015a)		X									
Piscopo et al. (2015)				X							
Yang et al. (2015)				X			X				
Fowler et al. (2015)	X			X			X				
Zheng et al. (2015b)					X	X					
Creaco and Pezzinga (2015)					X	X					
Li et al. (2015)					X	X	X				
Zimmer et al. (2015)					X	X	X				
Dumedah (2015)					X	X					
Hadka and Reed (2015)						X	X				
Bi et al. (2015)						X					
Broad et al. (2015)						X				X	
Tsoukalas and Makropoulos (2015)						X	X			X	
McClymont et al. (2015)	X						X				
Lerma et al. (2015)				X			X	X	X		
Mortazavi-Naeini et al. (2015)										X	
Stokes et al. (2015)											X

Gibbs et al. (2015) develop a relationship between metrics that quantify fitness function characteristics and the number of generations needed for a genetic algorithm to converge in a pre-determined number of generations for a large number of synthetically generated test problems with different attributes. This relationship is then validated on two water distribution system optimisation problems, including the Cherry Hill-Brushy Plains network, which is a commonly used test problem, and the optimal operation of the Woranora water distribution near Sydney, Australia, which is a real-world case study. The ability to select the population size that results in convergence for a given computational budget based on problem characteristics is likely to be very useful for solving real-world problems where computational issues are a problem, particularly in operational settings.

Zheng et al (2015a) use a number of run-time behaviour analysis measures to better understand how a differential evolution (DE) EA explores the solution space and why it produces the solutions it does at various stages of searching for three water distribution system optimisation problems of varying complexity and different parameterisations of the DE. The ability to understand how algorithms and algorithm parameterisations navigate through the solution space throughout the search for different problems is vital in terms of the ability to select the most appropriate algorithms and their parameters, to design better algorithms and to dynamically adjust searching behaviour during an optimisation run in order to maximise performance.

Piscopo et al. (2015) address the important issue of problem formulation for real-world problems in the context of the application of the Borg multi-objective evolutionary algorithm to the problem of optimising engineered injection and extraction for groundwater remediation. While in the vast majority of optimisation studies in literature problem formulation is established *a priori* and treated as fixed, this is generally not the case when dealing with real-world problems. In this paper, a novel iterative optimisation approach is introduced, as part of which problem formulation is updated based on the results of prior rounds of optimisation.

Yang et al. (2015) tackle the issue of problem formulation for the real-world case study of optimising the hydropower reservoir operation of the Oroville-Thermalito Complex in California, USA. Particular attention is given to the impact of the simplification of the reservoir's highly non-linear storage-elevation relationship. In addition, the performance of a new multi-objective search technique (Multi-Objective Complex Evolution Global Optimization Method with Principal Component Analysis and Crowding Distance Operator) is compared with that of a number of other techniques, including the Multi-Objective Complex Evolution Global Optimization method, the Multi-Objective Differential Evolution method, the Multi-Objective Genetic Algorithm, the Multi-Objective Simulated Annealing approach and the Multi-Objective Particle Swarm Optimization scheme, in order to determine which searching behaviour performs best.

Fowler et al. (2015) introduce a formulation for the real-world problem of deciding which crop planting choices farmers should make when faced with competing revenue, water use and demand objectives. The formulation is tested on a hypothetical case study where MODFLOW-FMP2 is used as the simulation package and a Multi-objective genetic algorithm is used as the optimisation engine. An extensive sensitivity analysis is used to obtain a better understanding of the relationship between algorithm parameterisation, algorithm performance and problem characteristics.

Zheng et al. (2015b) address the issue of increasing computational efficiency of multi-objective optimisation problems by means of search space size reduction. This is achieved by decomposing the optimisation problem into a subset of smaller problems via graph theoretic approaches and optimising each of these sub-problems independently. A novel approach is then used to propagate the Pareto fronts of the sub-problems towards the Pareto front of the original problem without the need to analyse the full problem. The approach is developed for the design of water distribution systems and its advantages demonstrated for two large case studies, including a real-world system from a suburb in a city in the south of China, with multi-objective differential evolution as the optimisation engine.

Creaco and Pezzinga (2015) also address the problem of reducing the size of the search space and increasing computational efficiency, but for problems that have both discrete integer and real decision variables. This is achieved by dividing the search space into two sub-regions consisting of the different types of decision variables and solving the problem using a hybrid optimisation approach, as part of which the EA is used to search through the discrete variables and linear programming is used to identify the optimal values of the real-valued decision variables for each of the solutions identified by the EA. The approach is demonstrated for the multi-objective optimisation of the location of control valves for leakage attenuation in water distribution systems, where the multi-objective optimisation algorithm NSGA-II is used as the EA.

Similar to Creaco and Pezzinga (2015), Li et al. (2015) also tackle the issue of search space size reduction by means of a hybrid EA-linear programming approach. However, their approach is applicable to the optimisation of multi-reservoir systems with heterogeneous hydropower units and involves the use of EAs for multi-reservoir optimisation as part of an outer loop and the use of linear programming for optimal unit scheduling as part of an inner loop. The approach is applied to the Three Gorges system in China. As part of the analysis, the performance of seven different EAs (or similar heuristic search methods) for use in the outer loop is compared, including a simple genetic algorithm, an improved genetic algorithm, particle swarm optimisation, simulated annealing, dynamically dimensioned search, dynamic coordinate search using response surface models and the stochastic radial basis function method.

Zimmer et al. (2015) explore an approach to reducing the size of the search space and increasing the computational efficiency of real-valued EAs in the context of model predictive control for time-varying systems with moving decision windows. This is achieved by investigating the impact of a range of modifications to standard genetic algorithms, including gene shifting, use of a reduced alphabet, application of the compact genetic algorithm and use of the micro genetic algorithm. The efficacy of these approaches is evaluated on a portion of the Chicago combined sewer and interceptor system with the aim of minimising combined sewer overflow during real-time use.

Dumedah (2015) combines the power of EAs with that of data assimilation to introduce a unified evolutionary data assimilation (EDA) approach that results in the provision of a genome-like data set, which can be used for search-space size reduction. The approach is applied to the multi-objective calibration of a Sacramento Soil Moisture Accounting model in the Fairchild creek catchment in southern Ontario, Canada, using NSGA-II.

Hadka and Reed (2015) tackle the issue of long run-times associated with the optimisation of real-world water resources problems by exploring the effectiveness of two different parallel implementations of the Borg multi-objective evolutionary algorithm, including master-slave and multi-master implementations. These implementations are applied to a case study of risk-based urban water portfolio planning for a city located in the Lower Rio Grande Valley in southern Texas and their performance is assessed for 1024, 2048, 4096, 8192 and 16384 processors. The performance of Borg is also compared with that of the large-cluster master-slave ε-NSGA-II as a benchmark.

In order to increase the computational efficiency of EA runs for real-world problems with long run-times, Bi et al. (2015) introduce a novel sampling approach for initialising EA search from good starting position in decision variable space for the water distribution system design problem. The sampling approach is based on domain knowledge of the problem under consideration and is applied to seven water distribution system design problems of varying size and complexity. The performance of the proposed sampling approach is compared with that of an existing sampling approach that considers domain knowledge, as well as random and Latin hypercube sampling. EPANet is used as the hydraulic solver and a simple genetic algorithm is used as the optimisation engine.

Broad et al. (2015) also address the issue of increasing computational efficiency, but by focussing on a reduction of the computational effort associated with the use of simulation models for objective function and/or constraint evaluation with the aid of surrogate/metamodels. They introduce a novel framework for identifying which component of the objective function and constraint evaluation process is most suitable for replacement by a surrogate/metamodel and apply it to the risk-based optimal design of water distribution systems. The framework is tested on two case studies, including a benchmark problem and a real-world case study system from the USA, called Pacific City.

Both case studies include hydraulic and water quality considerations, as well as reliability-based performance criteria. Artificial neural networks are used as the metamodels to replace the computationally expensive EPANet hydraulic simulation models and a single-objective genetic algorithm is used as the optimisation engine.

In order to enable the robust, multi-objective optimisation of long-term operating rules for multi-reservoir systems considering stochastic system inputs, Tsoukalas and Makropoulos (2015) develop a surrogate modelling based optimisation framework for increasing computational efficiency. The framework is demonstrated on a real-world hydrosystem with three hydro-electric power stations on the River Nestos in Greece. WEAP21 is used as the hydrosystem simulation model and kriging is used as the surrogate modelling technique. The performance of difference optimisation approaches is compared, including ParEGO, the Surrogate MOdeling (SUMO) Toolbox and the SMS-EGO algorithm. In addition, the performance of these surrogate-based optimisation approaches is compared with that of two standard multi-objective algorithms, including NSGA-II and SMS-EMOA.

McClymont et al. (2015) address the problem of identifying which algorithm parameterisation should be used for particular problems. This is achieved by introducing an approach that enables the performance of different search methods to be compared for problems with different characteristics/features. The approach is applied to three benchmark water distribution system design problems. Problem features are characterised in terms of topology and assets and search methods are characterised by different genetic operators, including mutation, crossover, pipe smoothing and pipe expansion. The water distribution systems are simulated using EPANet.

Lerma et al. (2015) identify which searching mechanisms and termination/convergence criteria are best for the determination of optimal operating rules for water resources systems. This is achieved by means of an extensive sensitivity analysis considering two EAs, including SCE-UA and Scatter Search, a number of optimisation algorithm parameters and a number of different stopping/convergence criteria for a theoretical case study. The two EAs with optimised parameters and stopping criteria are used to solve a real-world, complex case study, the Tirso-Flumendosa-Campidano system located on the island of Sardinia, Italy, which is simulated using the SIMGES water allocation model. Consultation with stakeholders provides important insight in terms of problem formulation and the communication of the results of the optimisation process.

Mortazavi-Naeini et al. (2015) address the problem of incorporating uncertainty into the optimisation of real-world systems, and develop an approach to finding robust optimal solutions that secure real urban bulk water systems against extreme drought in the presence of deep uncertainty about future climate change. The approach is applied to the Lower Hunter urban bulk water system in New South Wales, Australia. The ε-multi-objective optimization evolutionary algorithm is used as the optimisation engine and the system is simulated using WATHNET5.

Stokes et al. (2015) tackle the issue of providing a uniform computational platform that enables the results of different studies to be compared with confidence. They outline some general principles for the development of computational software frameworks and present one such framework for the minimisation of costs and greenhouse gas emissions from water distribution systems. The software is easily accessible and freely available for others to use. The utility of various aspects of the software tool is demonstrated for a theoretical case study.

The above articles cover a wide range of issues related to Evolutionary Algorithms in Water Resources. However, the list of issues covered is by no means exhaustive. In addition, the papers do not answer all of the key research questions posed by Maier et al (2014). However, they provide an excellent starting point and will hopefully encourage and inspire you to make your own contribution towards meeting the challenges outlined in Maier et al. (2014). In the meantime, we hope you enjoy reading the papers in this thematic issue.

References

W. Bi, G.C. Dandy, H.R. Maier (2015) Improved genetic algorithm optimization of water distribution system design by incorporating domain knowledge, Environmental Modelling & Software, Available online 29 September 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.09.010.

Darren R. Broad, Graeme C. Dandy, Holger R. Maier (2015) A systematic approach to determining metamodel scope for risk-based optimization and its application to water distribution system design, Environmental Modelling & Software, Available online 3 December 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.11.015.

E. Creaco, G. Pezzinga (2015) Embedding linear programming in multi objective genetic algorithms for reducing the size of the search space with application to leakage minimization in water distribution networks, Environmental Modelling & Software, Available online 4 November 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.10.013.

Gift Dumedah (2015) Toward essential union between evolutionary strategy and data assimilation for model diagnostics: An application for reducing the search space of optimization problems using hydrologic genome map, Environmental Modelling & Software, Available online 30 October 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.09.025.

K.R. Fowler, E.W. Jenkins, C. Ostrove, J.C. Chrispell, M.W. Farthing, M. Parno (2015) A decision making framework with MODFLOW-FMP2 via optimization: Determining trade-offs in crop selection, Environmental Modelling & Software, Available online 8 January 2015, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.11.031.

Matthew S. Gibbs, Holger R. Maier, Graeme C. Dandy (2015) Using characteristics of the optimisation problem to determine the Genetic Algorithm population size when the number of evaluations is limited, Environmental Modelling & Software, Available online 7 September 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.08.023.

David Hadka, Patrick Reed (2015) Large-scale parallelization of the Borg multiobjective evolutionary algorithm to enhance the management of complex environmental systems, Environmental Modelling & Software, Available online 6 November 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.10.014.

Néstor Lerma, Javier Paredes-Arquiola, Joaquín Andreu, Abel Solera, Giovanni M. Sechi (2015) Assessment of evolutionary algorithms for optimal operating rules design in real Water Resource

Systems, Environmental Modelling & Software, Available online 25 October 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.09.024.

Fang-Fang Li, Christine A. Shoemaker, Jun Qiu, Jia-Hua Wei (2015) Hierarchical multi-reservoir optimization modeling for real-world complexity with application to the Three Gorges system, Environmental Modelling & Software, Available online 24 December 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.11.030.

Maier H.R., Kapelan Z., Kasprzyk J., Kollat J., Matott L.S., Cunha M.C., Dandy G.C., Gibbs M.S., Keedwell E., Marchi A., Ostfeld A., Savic D., Solomatine D.P., Vrugt J.A., Zecchin A.C., Minsker B.S., Barbour E.J., Kuczera G., Pasha F., Castelletti A., Giuliani M. and Reed P.M. (2014) Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions, Environmental Modelling and Software, 62, 271-299, DOI: 10.1016/j.envsoft.2014.09.013.

K. McClymont, E. Keedwell, D. Savic (2015) An analysis of the interface between evolutionary algorithm operators and problem features for water resources problems. A case study in water distribution network design, Environmental Modelling & Software, Available online 7 February 2015, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.12.023.

Mortazavi-Naeini M., Kuczera G., Kiem A.S., Cui L., Henley B., Berghout B. and Turner E. (2015) Robust optimization to secure urban bulk water supply against extreme drought and uncertain climate change, Environmental Modelling & Software, accepted (this issue).

A.N. Piscopo, J.R. Kasprzyk, R.M. Neupauer (2015) An iterative approach to multi-objective engineering design: Optimization of engineered injection and extraction for enhanced groundwater remediation, Environmental Modelling & Software, Available online 30 September 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.08.030.

Christopher S. Stokes, Angus R. Simpson, Holger R. Maier (2015)A computational software tool for the minimization of costs and greenhouse gas emissions associated with water distribution systems, Environmental Modelling & Software, Available online 3 December 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.11.004.

Ioannis Tsoukalas, Christos Makropoulos (2015) Multiobjective optimisation on a budget: Exploring surrogate modelling for robust multi-reservoir rules generation under hydrological uncertainty, Environmental Modelling & Software, Available online 16 October 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.09.023.

Tiantian Yang, Xiaogang Gao, Scott Lee Sellars, Soroosh Sorooshian (2015) Improving the multi-objective evolutionary optimization algorithm for hydropower reservoir operations in the California Oroville–Thermalito complex, Environmental Modelling & Software, Available online 18 December 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.11.016.

Feifei Zheng, Aaron C. Zecchin, Angus R. Simpson (2015a) Investigating the run-time searching behavior of the differential evolution algorithm applied to water distribution system optimization, Environmental Modelling & Software, Available online 16 October 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.09.022.

Feifei Zheng, Angus Simpson, Aaron Zecchin (2015b) Improving the efficiency of multi-objective evolutionary algorithms through decomposition: An application to water distribution network design, Environmental Modelling & Software, Available online 16 September 2014, ISSN 1364-8152, http://dx.doi.org/10.1016/j.envsoft.2014.08.022.

Zimmer A., Schmidt A., Ostfeld A. and Minsker B. (2015) Evolutionary algorithm enhancement for model predictive control and real-time decision support, Environmental Modelling & Software, accepted (this issue).