Multiobjective optimisation of urban wastewater systems using ParEGO: a comparison with NSGA II

G. Fu*, S.-T. Khu and D. Butler

Centre for Water Systems, School of Engineering, Computing and Mathematics, University of Exeter, North Park Road, Harrison Building, Exeter EX4 4QF, UK

**Corresponding author, email: G.Fu@exeter.ac.uk*

ABSTRACT
Commercial and research-based simulation models are now available to represent the performance and control of the sewer network, wastewater treatment plant and receiving water as a whole. To improve overall system performance, these models can be combined with optimisation methods to derive optimal control strategies. The popular evolutionary algorithms (EAs) have been proven to be a powerful method in developing optimal control strategies; however, the high computational requirements of these methods impose a limit on their application due to the complexity of the system. This paper explores the potential of a surrogate-based multi-objective optimisation method, ParEGO, for real time control of urban wastewater systems. An existing integrated model is used to evaluate the multiple objectives. This method is compared with NSGA II by using two performance indicators: the hypervolume indicator and the additive binary ε-indicator. Comparative results show that ParEGO is an efficient and effective method in deriving optimal control strategies for the multiple objective control problems. It is suggested that ParEGO can greatly improve the computational efficiency, particularly for complex systems.

KEY WORDS
Integrated modelling; multi-objective optimisation; NSGA II; ParEGO; real time control; urban wastewater system

INTRODUCTION
There is growing recognition of the need to take an integrated approach to water management, which has led to the development of integrated models of the various hydraulic and water quality processes in the drainage/sewer system, treatment plant and receiving water body as a whole (Rauch *et al.*, 2002; Butler and Schütze, 2005; Vanrolleghem *et al.*, 2005). The development of integrated models provides us with the opportunity to control the urban wastewater system as a whole, and it enables two kinds of integration according to Schütze *et al.* (2002): objective integration by which control of one subsystem may be based on the objective measured in other subsystems, and information integration by which control of one subsystem is based on the state information from other subsystems. Integration helps achieve improved system performance through development of optimal control strategies; however, it also makes the system more complex, and thus more challenging and expensive for optimisation methods.

Control of urban wastewater systems is usually regarded as a non-linear mathematical optimisation problem, and in many situations, a multiobjective optimisation problem (Fu *et al.*, 2008). Evolutionary algorithms (EAs) have been proven as promising to derive the optimal control strategies, compared with the conventional optimisation techniques (Rauch and Harremoës, 1999; Muschalla *et al.*, 2006). However, this technique generally needs tens of thousands of model simulations in order to reach the optimal control strategies. The computational burden makes EAs very inefficient and impractical for real time control, which requires a rapid decision making on selection of control strategies.

To improve computational efficiency, some forms of surrogate modelling have been used for fitness approximation in evolutionary computation, and a good summary was given by Jin (2005). The most used methods include polynomials, the kriging model, neural networks, and support vector machines. This paper explores the potential and the benefit of a fast surrogate method, ParEGO (Knowles, 2006), for the multi-objective control problem in urban wastewater systems. This method is based on the popular kriging approach, the Design and Analysis of Computer Experiments (DACE) and can usually achieve a satisfying set of Pareto solutions within a few hundreds of objective evaluations. This method is compared with one of the state-of-the-art EAs, NSGA II (Deb *et al.*, 2002), and is demonstrated by a semi-hypothetical case study.

THE PAREGO ALGORITHM

ParEGO is an extension of the single objective efficient global optimisation (EGO) for multiobjective optimisation problems (Knowles, 2006). The approximate method used is one of the kriging approaches, the Design and Analysis of Computer Experiments (DACE).

DACE

In the Kriging approach, the model y with n variables, $\mathbf{x} = (x_1, \ \cdots, \ x_n)$, is described as

$$y(\mathbf{x}) = g(\mathbf{x}) + z(\mathbf{x})$$

where $g(\mathbf{x})$ is the regression term, usually a polynomial function, and $z(\mathbf{x})$ is the error term, represented by a Gaussian random function with zero mean and non-zero covariance. In the stochastic process, the errors for N samples are related or ‘corelated’ and the correlation is related to the distance between the corresponding samples, and usually expressed as:

$$R\left(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}\right) = \exp\left[-\sum_{h=1}^{n} \mathrm{q}_h \left|x_h^{(i)} - x_h^{(j)}\right|^{p_h}\right], \ (\mathrm{q}_h \geq 0, \ p_h \in [1,2], \text{ and } i, j = 1, \ \cdots, \ N)$$

where q_h is a parameter measuring the importance of the variable x_h, $x_h^{(i)}$ and $x_h^{(j)}$ are the value of the variable x_h in sample points $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(j)}$, respectively. The covariance of $z(\mathbf{x})$ is denoted as

$$\mathrm{Cov}\left[z\left(\mathbf{x}^{(i)}\right), z\left(\mathbf{x}^{(j)}\right)\right] = \mathrm{s}^2 \mathbf{R}\left[R\left(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}\right)\right]$$

where $z\left(\mathbf{x}^{(i)}\right)$ is Normal (0, s^2), $i = 1, \ \cdots, \ N$, and $\mathbf{R}$ is the symmetric correlation matrix for N samples. It proves that modelling the correlation in this way is so powerful that the regression term $g(\mathbf{x})$ can be reduced to a simple constant term b, which is regarded as the mean of the stochastic process (Jones *et al.*, 1998).

This model has a total of $2n+2$ parameters: b, s^2, $q_1, \cdots, q_n$ and $p_1, \cdots, p_n$. The maximum likelihood method can be used to estimate these parameters. The prediction for a point $\mathbf{x}^*$ can be calculated as

$$\hat{y} = \hat{b} + \mathbf{r}^{\mathrm{T}}\left(\mathbf{x}^*\right)\mathbf{R}^{-1}\left(\mathbf{y} - \hat{b}\mathbf{I}\right)$$

where $\hat{b}$ is the estimated value of b, $\mathbf{y}$ is a vector of model outputs for the N samples, $\mathbf{I}$ is a unit vector of length N, and $\mathbf{r}$ is the correlation vector between the error term at the predicted point and the error terms at the previously sampled points. The ith element of $\mathbf{r}$ is $R\left(\mathbf{x}^*, \mathbf{x}^{(i)}\right)$.

One of the advantages of DACE is that a confidence interval of the prediction can be obtained, which is explicitly used by EGO and ParEGO to guide the search. A MATLAB toolbox developed by Lophaven *et al.* (2002) is used in this research for implementing the DACE model.

Implementation of ParEGO

The implementing process is shown in Figure 1. An internal genetic algorithm is used to search for the solution that maximizes the expected improvement, and to update the solution set, which consists of solutions evaluated by the real objective functions.

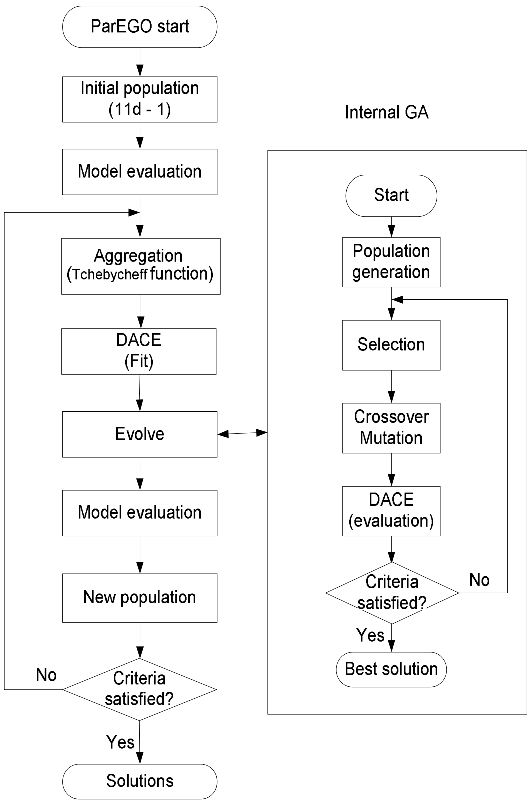

Figure 1. Flow chart of ParEGO

ParEGO is basically an aggregation-based algorithm, and the non-linear Tchebycheff function was suggested by Knowles (2006) to combine the n objectives into one single objective

$$f = \max_{j=1}^{n} \lambda_j f_j + \rho \sum_{j=1}^{n} \lambda_j f_j$$

Where f_j and λ_j ($j = 1, 2, \cdots, n$) are the *j*th normalized objective value and its weight, and ρ is a small positive parameter and was set to 0.05 according to Knowles (2006). The objectives are assumed to be simultaneously minimized in the aggregation function, and thus maximization objectives should be converted to minimization. In order to explore the whole region of the Pareto front, a varying weight vector is used in ParEGO, and is drawn randomly from the evenly distributed vector set,

$$\Lambda = \left\{ \left(\lambda_1, \quad \lambda_2, \quad \cdots, \quad \lambda_n \right) \middle| \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \in \left\{ 0, \quad \frac{1}{s}, \quad \cdots, \quad 1 \right\} \right\}$$

The total number of vectors in the above set is determined by $|\Lambda| = \begin{pmatrix} s+k-1 \\ k-1 \end{pmatrix}$, s is set to 10 for the two objective case in this paper.

PERFORMANCE INDICATORS

Performance indicators are used to assess and compare the properties of an approximation set derived from multiobjective evolutionary algorithms: convergence and diversity. Convergence is measured by the distance of the approximation set from the true Pareto front or a reference set, and diversity measures the extent of the approximation set in the objective space. Two good indicators were selected to compare the performance of ParEGO with NSGA II, and attainment surface plots were used to visualize the approximate sets.

Hypervolume indicator

This indicator, also known as the *S* metric or the Lebesgue measure, measures the size of the region of objective space dominated by a set of solutions. The hypervolume not only indicates the closeness of the solutions to the optimal set, but also captures the spread of the solutions over the objective space. This measure has been incorporated into multiobjective genetic algorithms as a selection criterion to improve the diversity of the solutions.

Several algorithms exist for calculating hypervolume, such as, the inclusion-exclusion (Wu and Azarm, 2001), LebMeasure (Fleischer, 2003), and HSO (While *et al.*, 2006). The HSO algorithm is used due to its computational efficiency. In calculating the region, a reference solution must be chosen, which should be dominated by every solution in the Pareto set. In this research, the reference solution is chosen from all the Pareto solutions obtained from each run.

Additive binary *ε*-indicator

This indicator was first defined by Zitzler *et al.* (2003) as follows: for a minimization problem with n positive objectives, a solution $z^1 = \left(z_1^1, \quad z_2^1, \quad \cdots, \quad z_n^1\right)$ is said to ε-dominate another solution $z^2 = \left(z_1^2, \quad z_2^2, \quad \cdots, \quad z_n^2\right)$, denoted as $z_i^1 \leq_{\varepsilon+} z_i^2$, if and only if

$\forall 1 \le i \le n : z_i^1 \le \varepsilon + z_i^2$

Then a pair of numbers (I_A , I_B) is defined as the binary ε -indicator

$$I_A = I_{\varepsilon+}(A,B) = \inf_{\varepsilon \in \Re}\{\forall z^2 \in B \ \exists z^1 \in A : z^1 \le_{\varepsilon+} z^2\}$$

$$I_B = I_{\varepsilon+}(B,A) = \inf_{\varepsilon \in \Re}\{\forall z^2 \in A \ \exists z^1 \in B : z^1 \le_{\varepsilon+} z^2\}$$

for a pair of Pareto set *A* and *B*. the Pareto set *A* is strictly better than *B* if ($I_A \le 0$, $I_B > 0$), and the two sets are incomparable if ($I_A > 0$, $I_B > 0$). However, *A* could be interpreted to be better than *B* in a weaker sense if $I_A < I_B$ (Knowles, 2006).

Since 10 runs were used to compare the performance of ParEGO and NSGA II, so the hypervolume and binary ε -indicators were calculated for the Pareto set from each run, and the mean and standard deviation are also computed for comparison.

RESULTS AND DISCUSSION

A case study is used to demonstrate the potential of ParEGO for multi-objective control of urban wastewater systems, and its performance is compared with NSGA II in two experiments with 160 (150 for ParEGO) and 260 objective evaluations respectively. The parameters for these two algorithms are set according to Knowles (2006).

The case study

The approach is demonstrated by an integrated case study, consisting of a combined sewer system, a treatment plant and receiving river. The catchment was first defined by Schütze (1998) and is semi-hypothetical in origin. It has been studied in detail for real time control optimisation (Schütze *et al.*, 2002; Butler and Schütze, 2005; Fu *et al.*, 2008).

The sewer system has seven sub-catchments with a total area of 725.8 ha, and four on-line pass-through storage tanks linked to sub-catchments 2, 4, 6 and 7 respectively, which are controlled by a pump. The wastewater treatment plant includes an off-line pass-through storm tank, a primary clarifier, aerator, and secondary clarifier. The treatment plant effluent and storm tank overflow are discharged to the river at Reach 10, and CSO discharges at Reach 3. In this research, the selected control variables include the maximum outflow rate of the storage tank linked to sub-catchment 7, the maximum inflow rate to the treatment plant, the threshold starting to empty the storm tank and its emptying flow rate, and return sludge rate. For the detailed set-up of the case study, the reader is referred to Schütze (1998) and Schütze *et al.* (2002).

This urban system was simulated by an existing integrated model developed using the SIMBA tool in the MATLAB/SIMULINK environment (IFAK, 2005). This model allows a holistic simulation of system dynamics and interactions between subsystems, and enable for assessment of system performance using receiving water quality indicators directly, rather than their surrogates, such as discharged CSO volume or pollutant loads. A storm event of a total depth of 27 mm is used for simulation of real time control. Two water quality indicators for the receiving river are considered in this paper, i.e., minimum DO concentration (DO-M) and maximum ammonium concentration (AMM-M) along all the river reaches.

Hypervolume and additive binary ε-indicators

Tables 1 and 2 show the 10-run results of the hypervolume and additive binary ε-indicators after 160 (150 for ParEGO) and 260 objective evaluations, respectively. For the hypervolume indicator, ParEGO can achieve a higher value for each run, which means a larger region in the objective space. For the additive binary ε-indicator, ParEGO weakly dominates NSGA II for all the runs with 260 evaluations, however, they achieve a relatively equivalent performance for the runs with 160/150 evaluations, considering that NSGA II achieves a smaller value in four cases. For both of the indicators, the standard deviations for ParEGO are always smaller than those of NSGA II, so it shows that ParEGO is more reliable.

Table 1. Hypervolume and additive binary ε-indicators for NSGA II and ParEGO with 160 and 150 objective evaluations, respectively.

Run number	Hypervolume		Additive binary ε-indicator	
	NSGA II	ParEGO	NSGA II	ParEGO
1	1.8405	2.5383	1.1653	0.0933
2	1.8314	2.3412	0.2499	0.2510
3	2.1233	2.5545	0.5191	0.1395
4	1.8314	2.4335	0.2486	0.2502
5	1.8454	2.5114	1.1535	0.0077
6	1.6781	2.5381	0.7404	0.1888
7	1.8314	2.3734	0.2292	0.2310
8	2.1233	2.4146	0.5191	0.1395
9	1.8949	2.5165	0.6035	0.0184
10	1.8314	2.4341	0.2399	0.2467
Mean	1.8831	2.4656	0.5668	0.1566
Std	0.1380	0.0758	0.3592	0.0937

Table 2. Hypervolume and additive binary ε-indicators for NSGA II and ParEGO with 260 objective evaluations.

Run number	Hypervolume		Additive binary ε-indicator	
	NSGA II	ParEGO	NSGA II	ParEGO
1	1.9154	2.4229	0.2518	0.1772
2	1.8748	2.4307	0.2879	0.0709
3	2.185	2.6128	0.6982	0.0953
4	1.8748	2.6128	0.2870	0.0619
5	1.6313	2.591	1.2910	0.0298
6	1.8078	2.6128	0.9195	0.2189
7	1.8748	2.5141	0.2779	0.0705
8	2.185	2.5141	0.6982	0.0953
9	1.9088	2.591	0.7713	0.0387
10	1.8748	2.5316	0.2790	0.0621
Mean	1.9132	2.5434	0.5762	0.0921
Std	0.16474	0.0731	0.3568	0.0604

The attainment surface

An attainment surface can show the division of the objective space by a set of the Pareto solutions. The best attainment surface from multiple runs visualises the biggest objective space that is achieved in all the runs, so it gives a good indication about the performance of an algorithm. Figure 2 shows the best surfaces for NSGA II and ParEGO after 10 runs. The plots show that ParEGO can dominate a larger space, particularly in the left hand side of the space. This probably was affected by the weights in the aggregation function, which gives a greater exploration on the space of the AMM-M objective.

The best surfaces are also compared with a set of Pareto solutions from a NSGA II run with 10,000 objective evaluations as shown in Figure 2. The best surfaces from ParEGO give a good approximate to the Pareto front from 10,000 evaluations.

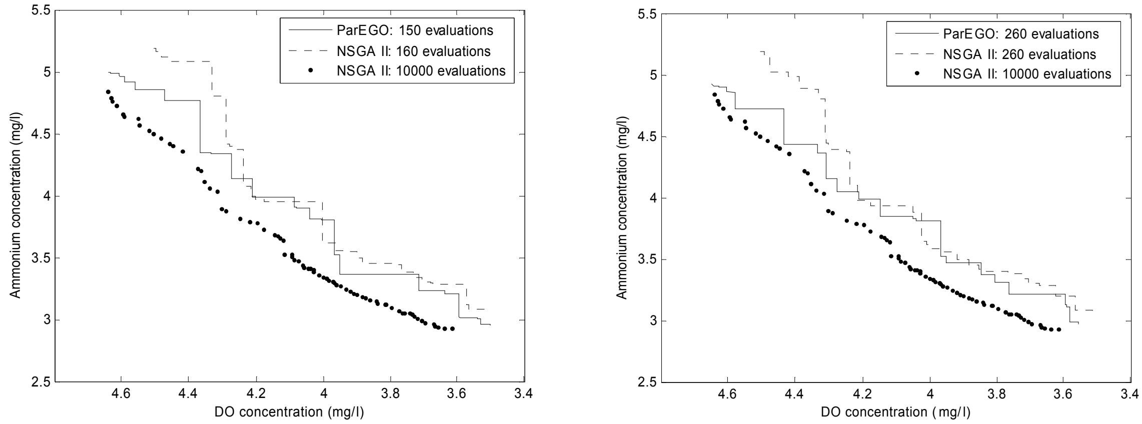

Figure 2. The best attainment surfaces for ParEGO and NSGA II

CONCLUSIONS

This paper explores the potential and the benefit of a fast surrogate method, ParEGO, for multiobjective real time control of urban wastewater systems. This method is demonstrated by an integrated case study in which the receiving water quality parameters (the minimum DO and maximum ammonium concentrations) are directly used as control objectives. The comparative results with NSGA II show that ParEGO shows that ParEGO can achieve a relatively better performance, particularly for the runs with 260 objective evaluations. ParEGO has less variation in performance, and it also gives a good approximate to the set of Pareto solutions derived by NSGA II with 10,000 objective evaluations. For the optimal control problem in the urban wastewater system, ParEGO is an efficient and effective method in deriving optimal control strategies.

ACKNOWLEDGEMENT

The study is partly funded by UK EPSRC grant no GR/S86846/01 and is also part of the Integrative Systems and the Boundary Problem (ISBP) project (www.tigress.ac/isbp), supported by the European Union's Framework 6 Programme New and Emerging Science and Technology Pathfinder initiative.

REFERENCES

Butler D. and Schütze M. (2005). Integrating simulation models with a view to optimal control of u rban wastewater systems. Environmental Modelling and Software, **20**, 415-426.

Deb K., Pratap A, Agarwal S., Meyarivan T. (2002). A fast and elitist multi -objective genetic algorithm: NSGA-II. IEEE Transaction on Evolutionary Computation, **6**(2) 181-197.

Fleischer M. (2003). The measure of Pareto optima: Applications to multi -objective metaheuristics. In: C. M. Fonseca et al., (eds.), Evolutionary Multiobjective Optimisation. ser. Lecture Notes in Computer Science, Berlin, Springer-Verlag, vol. 2632, pp. 519-533.

Fu G., Butler D. and Khu S.T. (2008). Multiple objective optimal control of integrated urban wastewater s ystems. Environmental Modelling and Software, **23**(2), 225-234.

IFAK (2005). SIMBA 5.0: simulation of wastewater systems, user's guide. Institut für Au tomation und Kommunikation e.V. Magdeburg, Germany.

Jin Y. (2005). A comprehensive survey of fitness approximation in evolutionary computation. Soft Computing, **9**, 3-12.

Jones D., Schonlau M., and Welch W. (1998). Efficient global optimisation of expensive black-box functions. Journal of Global Optimisation, **13**, 455-492.

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

Lophaven S.N., Nielsen H.B., Søndergaard J. (2002). DACE - A Matlab Kriging Toolbox, Version 2.0. Technical Report IMM-TR-2002-12, Technical University of Denmark. Available at: http://www2.imm.dtu.dk/~hbn/dace.

Muschalla D., Schrφter K., and Schütze M. (2006). Multi-objective evolutionary algorithms in the field of urban drainage. Seventh International Conference on Hydroinformatics, Nice.

Rauch W., and Harremoës P., (1999). On the potential of genetic algorithms in urban drainage modelling. Urban Water, **1**, 79-89.

Rauch W., Bertrand-Krajewski J.-L., Krebs P., Mark O., Schilling W., Schütze M., and Vanrolleghem P.A., (2002). Mathematical modelling of integrated urban drainage systems. W ater Science and Technology, **45**(3), 81-94.

Schütze M. (1998). Integrated Simulation and Optimum Control of the Urban Wastewater System. PhD thesis. Department of Civil Engineering; Imp erial College of Science, Technology and Medicine, University of London, London.

Schütze M., Butler D. and Beck M.B. (2002). Modelling Simulation and Control of Urban Wastewater Systems, Springer, Berlin.

Vanrolleghem P.A., Benedetti L., and Meirlaen J., (2005). Modelling and real-time control of the integrated urban wastewater system. Environmental Modelling & Software, **20** (4), 427-442.

While L., Hingston P., Barone L., Huband S. (2006). A faster algorithm for calculating hypervolume. IEEE Transactions on Evolutionary computation, **10** (1), 29-38.

Wu J. and Azarm S. (2001). Metrics for quality assessment of a multiobjective design optimisation solution set. J. Mechanical Design, **123**, 18-25.

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