The current issue and full text archive of this journal is available at
www.emeraldinsight.com/0332-1649.htm

Scalarizing cost-effective multi-objective optimization algorithms made possible with kriging

Glenn Hawe
Vector Fields Ltd, Oxford, UK and
School of Electronics and Computer Science, University of Southampton, Southampton, UK, and
Jan Sykulski
School of Electronics and Computer Science, University of Southampton, Southampton, UK

Abstract

Purpose – The purpose of this paper is threefold: to make explicitly clear the range of efficient multi-objective optimization algorithms which are available with kriging; to demonstrate a previously uninvestigated algorithm on an electromagnetic design problem; and to identify algorithms particularly worthy of investigation in this field.

Design/methodology/approach – The paper concentrates exclusively on scalarizing multi-objective optimization algorithms. By reviewing the range of selection criteria based on kriging models for single-objective optimization along with the range of methods available for transforming a multi-objective optimization problem to a single-objective problem, the family of scalarizing multi-objective optimization algorithms is made explicitly clear.

Findings – One of the proposed algorithms is demonstrated on the multi-objective design of an electron gun. It is able to identify efficiently an approximation to the Pareto-optimal front.

Research limitations/implications – The algorithms proposed are applicable to unconstrained problems only. One future development is to incorporate constraint-handling techniques from single-objective optimization into the scalarizing algorithms.

Originality/value – A family of algorithms, most of which have not been explored before in the literature, is proposed. Algorithms of particular potential (utilizing the most promising developments in single-objective optimization) are identified.

Keywords Optimization techniques, Programming and algorithm theory

Paper type Research paper

1. Introduction

Multi-objective optimization algorithms which seek to locate a set of Pareto-optimal solutions may be categorized broadly into two families: non-scalarizing and scalarizing. Non-scalarizing algorithms consider each objective separately, and select new designs using a measure of how non-dominated they are; such methods are known as multi-objective evolutionary algorithms (Deb, 2001). Scalarizing algorithms on the other hand, transform the multi-objective optimization problem (MOOP) to a single-objective optimization problem (SOOP) using some particular method (Miettinen, 1999); then an algorithm from single-objective optimization is used to

COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering
Vol. 27 No. 4, 2008
pp. 836-844
© Emerald Group Publishing Limited
0332-1649
DOI 10.1108/03321640810878243

select design vectors for evaluation. Varying the parameters in the scalarization method used between iterations then enables an approximation of the Pareto set to be built up.

Owing to the high-computational cost of evaluating objective functions and constraints in electromagnetic optimal design problems, cost-saving techniques, such as surrogate-modeling, are typically employed. One popular type of surrogate model is kriging (Lebensztain *et al.*, 2004). Functions which utilize statistical information regarding the uncertainty in the prediction of the kriging model are then commonly used to select the next design vector to evaluate (Jones, 2001). The use of surrogate models in multi-objective optimization is relatively new; in particular, only recently have (non-trivial) non-scalarizing algorithms been proposed (Hawe and Sykulski, 2007a).

In single-objective optimization, a wide range of utility functions now exist, each of which may be combined with a scalarizing function to give rise to a (scalarizing) multi-objective optimization algorithm. Despite this simplicity, relatively few have been explored in the literature. The purpose of this paper is threefold:

(1) to provide a review of existing scalarizing cost-effective multi-objective optimization algorithms;

(2) to demonstrate a previously uninvestigated algorithm on an electromagnetic design problem; and

(3) to suggest algorithms worthy of investigation in this field.

The paper begins with a brief review of the state-of-the-art in selection criteria for single-objective optimization with kriging. This is then followed by a description of scalarizing methods commonly used to transform MOOPs to SOOPs. The full range of scalarizing algorithms made possible with kriging is then made explicitly clear, and attention is drawn to such algorithms already investigated in the literature. A previously uninvestigated algorithm is then demonstrated on the optimal design of an electron gun; finally, suggestions are made for further investigation.

2. Selection criteria with kriging models for single-objective optimization

Owing to lack of space (and the fact that many descriptions may be found elsewhere, e.g. Lebensztain *et al.*, 2004), the details of constructing a kriging model are omitted. Here, we concern ourselves only with the details of using a kriging model to determine where to evaluate next (for the purpose of minimizing a single-objective function f).

It is well known that, the statistical foundation of kriging models enables the uncertainty in their predictions to be quantified (Jones, 2001). Having a measure of the uncertainty in the prediction y of an unknown objective function f allows the concept of improvement to be defined for unevaluated design vectors. Specifically, suppose that objective function f is to be minimized, and that $f_{\min}$ is the lowest objective function value so far observed. Then, by modeling the kriging prediction of an unknown design vector $\mathbf{x}$ as the realization of a Gaussian distribution Y, with mean $y(\mathbf{x})$ and standard error $s(\mathbf{x})$ as given by the kriging model (Jones *et al.*, 1998), the improvement associated with design vector $\mathbf{x}$ is defined to be (Jones, 2001):

$$I[\mathbf{x}] = \max[(f_{\min} - Y), 0] \tag{1}$$

This concept has proven to be extremely useful in defining criteria for selecting design vectors to evaluate. For example, by specifying a certain target level which should be achieved T ($<f_{\min}$), using this model of uncertainty the probability of an unevaluated design vector attaining a value less than T is given by (Jones, 2001):

$$P(f(\mathbf{x}^*) < T) = \Phi\left(\frac{T - y(\mathbf{x}^*)}{s(\mathbf{x}^*)}\right) \tag{2}$$

where Φ is the Normal cumulative distribution function. This is shown in Figure 1: the probability of the unevaluated design vector $\mathbf{x}^*$ having an objective function value less than T is represented by the shaded area under the Gaussian distribution modeling the uncertainty in the value of $f(\mathbf{x}^*)$. By maximizing the value of $P(f(\mathbf{x}) < T)$ over design variable space, the design vector which has the greatest probability of achieving a value less than T may be selected for evaluation.

Rather than evaluate the design vector which maximizes the probability of achieving a certain level of improvement, the EGO algorithm (Jones *et al.*, 1998) chooses to evaluate the design vector which maximizes the expectation value of the improvement (the expected improvement), given by:

$$E[I[\mathbf{x}]] = (f_{\min} - y(\mathbf{x}))\Phi\left(\frac{f_{\min} - y(\mathbf{x})}{s(\mathbf{x})}\right) + s(\mathbf{x})\varphi\left(\frac{f_{\min} - y(\mathbf{x})}{s(\mathbf{x})}\right) \tag{3}$$

where Φ and φ are the normal cumulative and density functions, respectively. The first term in equation (3) favours design vectors close to the current minimum (exploitation), whilst the second term favours design vectors with high uncertainty in their value (exploration); by introducing weighting factors w_1 and w_2 before these two terms, a weighted expected improvement utility function (Sobester *et al.*, 2005) may be formed which allows control over the balance between exploration and exploitation. Furthermore, by defining the generalized improvement as:

$$I^g[\mathbf{x}] = \max[(f_{\min} - Y)^g, 0], \tag{4}$$

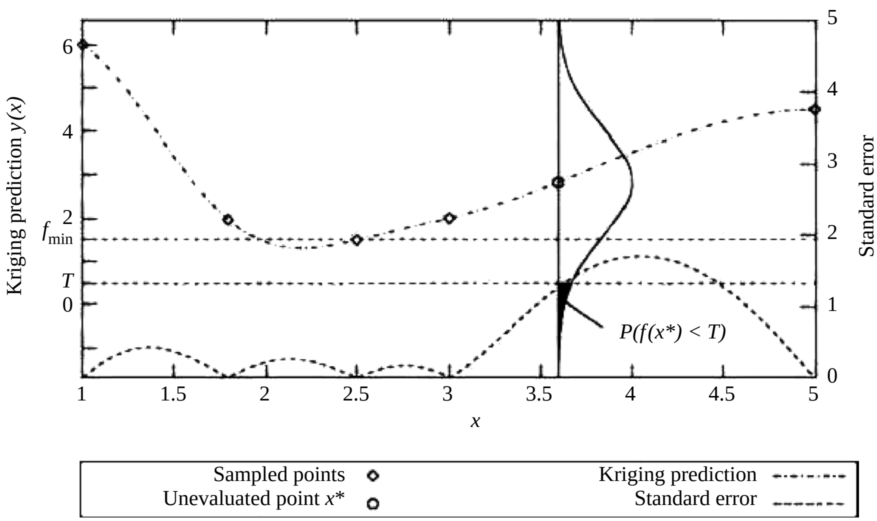

Figure 1.
Modeling the uncertainty in the prediction made by a kriging model using a Gaussian distribution

the generalized expected improvement may be shown to be (Sasena *et al.*, 2002):

$$E\left[I^g[\mathbf{x}]\right] = s^g(\mathbf{x})\sum_{k=0}^{g}(-1)^k\left(\frac{g!}{k!(g-k)!}\right)\left(\frac{f_{\min}-y(\mathbf{x})}{s(\mathbf{x})}\right)^{g-k}T_k \tag{5}$$

where:

$$T[\mathbf{x}] = -\phi\left(\frac{f_{\min}-y(\mathbf{x})}{s(\mathbf{x})}\right)\left(\frac{f_{\min}-y(\mathbf{x})}{s(\mathbf{x})}\right)^{k-1}+(k-1)T_{k-2} \tag{6}$$

with:

$$T_0[\mathbf{x}] = \Phi\left(\frac{f_{\min}-y(\mathbf{x})}{s(\mathbf{x})}\right) \tag{7}$$

and:

$$T_1[\mathbf{x}] = -\phi\left(\frac{f_{\min}-y(\mathbf{x})}{s(\mathbf{x})}\right). \tag{8}$$

Higher values of g correspond to higher levels of improvement, and so tend to favour exploration of empty regions of design variable space. Variation of the parameter g between iterations allow the balance between exploration and exploitation to be balanced (Figure 2).

Not all utility functions depend on the concept of improvement. For example, the conditional likelihood method (Jones, 2001) is quite different: by hypothesizing the value of the objective function f^*, the design vector $\mathbf{x}^*$ which yields the most credible response surface (given by the likelihood of the sampled points conditional on the surface passing through the hypothesized point) is chosen for evaluation.

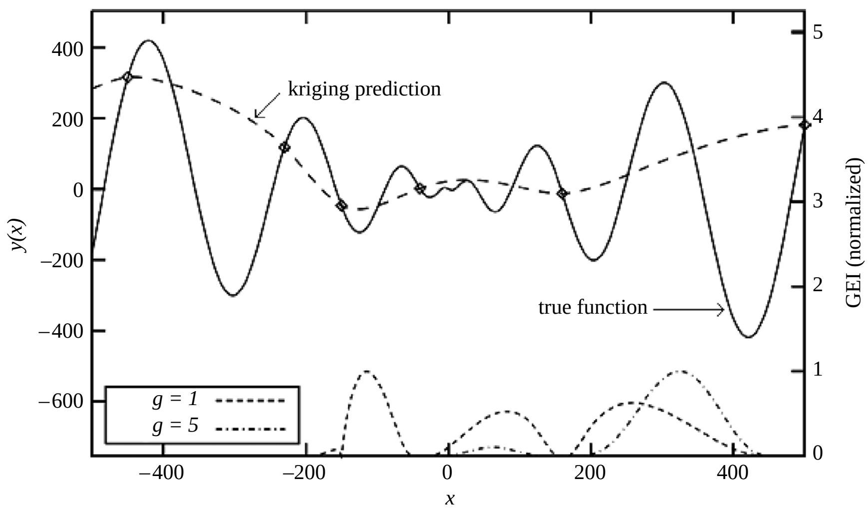

Figure 2. Illustration of the generalized expected improvement functions on a test function

This concept is shown in Figure 3, and used in the algorithm proposed in Hawe and Sykulski (2007b).

Recently, the minimizer entropy utility function has been proposed (Villemonteix *et al.*, 2008) and implemented in an algorithm known as informational approach to global optimization (IAGO). Rather than evaluate the design vector which is most likely to be the global minimum of the objective function, IAGO chooses to evaluate the design vector which maximizes the information gain on the position of the global minimum. Initial results show IAGO to be very effective; it promises to be very suitable for optimal electromagnetic design.

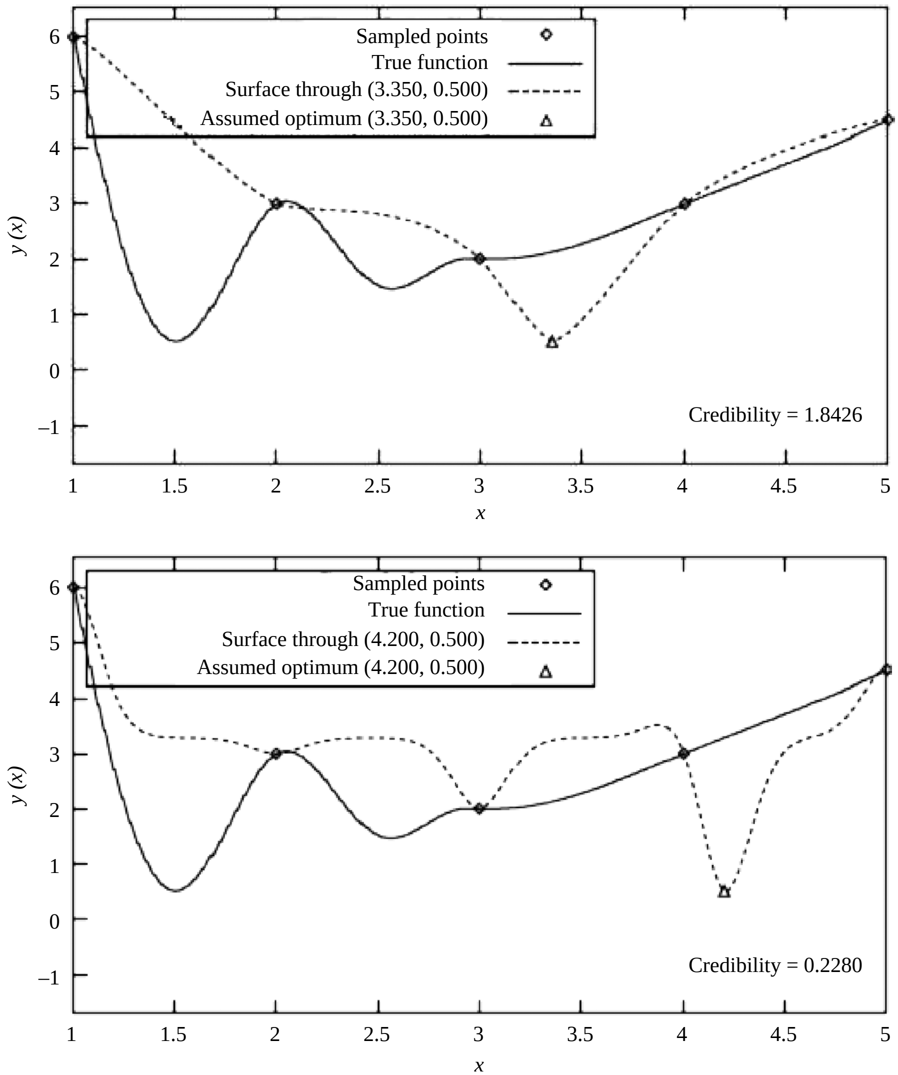

Figure 3.
Illustration of the credibility of hypothesis method on a test function

3. Scalarizing cost-effective multi-objective algorithms

3.1 Scalarizing methods

Many scalarizing methods exist for transforming a MOOP to a SOOP. This section discusses three of the most popular: the ε-constraint method, the weighting method, and the weighted metric method.

The ε-constraint method is unlike most other scalarizing methods as it does not involve weighting the multiple objectives into one function; instead one objective (f_l say) is chosen for minimization, whilst the others (f_i, $i \neq l$) are constrained below suitable limits ε_i:

$$\text{Minimize } f_l(\mathbf{x}) \quad \text{subject to } f_i(\mathbf{x}) < \varepsilon_i \quad i = 1, 2, \ldots, M, \quad i \neq l \tag{9}$$

By varying the limits ε_i, an approximation to the Pareto-front can be built up. This is illustrated for a two-objective problem in Figure 4(a), where f_2 is chosen as the objective to be minimized, with different upper bounds ε_i, $i = 0, 1, 2, 3$ shown for objective f_1. In this case, ε_0 is too low; no feasible solution has a value of $f_1 < \varepsilon_0$, and so no solution is found; however Pareto solutions z_1, z_2, and z_3 may be located using ε_1, ε_2, and ε_3, respectively. This method has the attractive feature of being capable of locating all Pareto-optimal solutions (Miettinen, 1999); however, it does require solving SOOPs with (potentially) high numbers of constraints, which is undesirable.

Another popular scalarizing method is to combine the multiple objectives using a weighted sum, i.e:

$$\text{Minimize } \hat{f}(\mathbf{x}) = \sum_{i=1}^{M} w_i \bar{f}_i(\mathbf{x}) \tag{10}$$

where $\bar{f}_i$ is the normalized value of objective f_i. One major drawback of the weighting method is its inability to capture solutions on concave parts of Pareto-optimal fronts. This is shown schematically in Figure 4(b): for every contour line which could potentially capture a solution on the concave part of the Pareto-front (such as that shown), a better (i.e. lower) contour line exists which yields a solution on the non-concave part of the front instead. Despite its drawbacks, the simplicity of this method means it is one of the most popular methods used to scalarize multi-objective problems.

The problem of locating concave solutions may be overcome by using different metrics to define the distance of a solution from the Utopian point z^*, such as the weighted L_p metric:

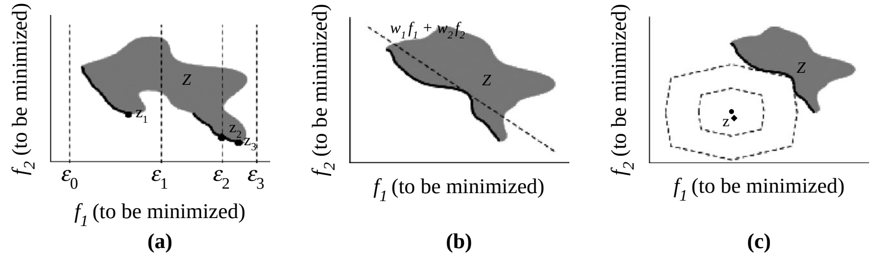

Figure 4. Illustration of three different scalarizing methods: (a) ε-constraint method; (b) contour lines of the weighting method; and (c) contour lines of the augmented Tchebycheff method

$$\left\| f_i(\mathbf{x}) - z^* \right\|_p = \left(\sum_{i=1}^{M} w_i \left| f_i(\mathbf{x}) - z_i^* \right|^p \right)^{1/p}. \qquad (11)$$

For the case $p = \infty$, the metric becomes the weighted Tchebycheff metric, variants of which exist such as the augmented Tchebycheff metric. Minimizing this metric:

$$\text{Minimize } \hat{f}(\mathbf{x}) = \max_{i=1,2,\ldots,M} \left[w_i \left| f_i(\mathbf{x}) - z_i^* \right| + \rho \sum_{i=1}^{M} \left| f_i(\mathbf{x}) - z_i^* \right| \right] \qquad (12)$$

(where ρ is a small positive constant) has the advantage that all Pareto solutions may be located (Miettinen, 1999), even those on the concave part of the Pareto-optimal front, as shown in Figure 4(c).

3.2 Scalarizing cost-effective algorithms made possible with kriging

Given the wide selection of kriging-assisted utility functions for solving SOOPs, and methods for transforming a MOOP to a SOOP, the range of potential cost-effective multi-objective algorithms is huge. Despite this, relatively few have been explored in the literature. Table I shows the algorithms made possible using just the selection of utility functions and scalarizing functions discussed in this paper alone: it should be noted however that, just as the discussion of selection criteria and scalarizing methods within this paper is by no means exhaustive, neither is the list of potential scalarizing algorithms in Table I.

Table I.
The family of scalarizing multi-objective optimization algorithms made possible with kriging

	Scalarizing method		
Kriging-based selection criteria	ε-constraint	Weighting method	Weighted metric
Probability of improvement			
Expected improvement	Jones (2001)		Knowles (2006)
Weighted EI			
Generalized EI			X
Credibility of hypothesis			Hawe and Sykulski (2007c)
Minimizer entropy			

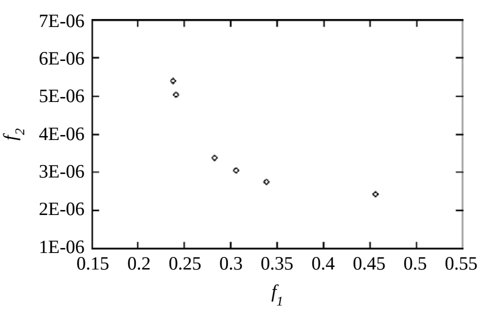

Figure 5.
Pareto solutions as found by the algorithm combining the generalized expected improvement selection criteria with the augmented Tchebycheff metric

Three of the potential algorithms in Table I have already been investigated in the literature: combining the expected improvement criteria with the ε-constraint (Jones, 2001) and weighted metric (Knowles, 2006) methods, and combining the credibility of hypothesis criteria with the weighted metric method (Hawe and Sykulski, 2007c). The following section demonstrates the algorithm combining the generalized expected improvement criteria with the weighted metric method (marked X in Table I); a suitable cooling scheme was used to vary the value of g between iterations.

3.3 Demonstration of algorithm on electromagnetic design problem

The voltage on, and position of, the focus electrode of an electron gun was varied so as to achieve two objectives: to focus the beam of electrons on the center of the anode as much as possible, and to make the electrons hit the anode face as perpendicular as possible. Formally, denoting the voltage on the focus electrode by V Volts, and its perpendicular distance from the emitting surface by d cm, the objective functions to be minimized are:

$$\text{Minimize } f_1(V, d) = \int_{\text{anode}} J(r) r^2 \mathrm{d}S \tag{13}$$

and:

$$f_2(V, d) = \int_{\text{anode}} \frac{\left(v_x^2 + v_y^2\right)}{\left(v_x^2 + v_y^2 + v_z^2\right)} \mathrm{d}S \tag{14}$$

subject to $V \in [0, 1000]$ and $d \in [4, 10]$, where r is the radial distance from the center of the anode surface, $J(r)$ is the current density at r, and the integrals are taken over the surface of the anode. v_x, v_y, and v_z are the components of the electron velocities as they hit the surface of the anode, which lies in the xy-plane. Each analysis was carried out using the vector fields space charge solver, SCALA. The Pareto optimal points found, along with one of the solutions, are shown in Figure 5.

4. Conclusions and further work

There are a huge number of potential cost-effective scalarizing multi-objective optimization algorithms possible with kriging. This paper has made explicitly clear how to construct such algorithms using selection criteria from kriging-assisted optimization, and demonstrated one such algorithm on the optimization of an electron gun. Clearly, many further algorithms are possible, and worthy of investigation; of particular potential is a scalarizing algorithm which utilizes the recently proposed information-based minimizer entropy selection criteria (Villemonteix *et al.*, 2008).

References

Deb, K. (2001), *Multi-objective Optimization Using Evolutionary Algorithms*, Wiley, New York, NY.

Hawe, G. and Sykulski, J.K. (2007a), "An enhanced probability of improvement utility function for locating Pareto-optimal solutions", *Proceedings of the 16th Conference on the Computation of Electromagnetic Fields, COMPUMAG 2007*, pp. 965-6.

Hawe, G. and Sykulski, J.K. (2007b), "A hybrid one-then-two stage algorithm for computationally expensive electromagnetic design optimisation", *COMPEL*, Vol. 26, No. 2, pp. 236-46.

Hawe, G. and Sykulski, J.K. (2007c), "A scalarizing one-stage algorithm for efficient multi-objective optimisation", *Proceedings of the 16th Conference on the Computation of Electromagnetic Fields, COMPUMAG 2007*, pp. 967-8.

Jones, D.R. (2001), "A taxonomy of global optimization methods based on response surfaces", *Journal of Global Optimization*, Vol. 21, pp. 345-83.

Jones, D.R., Schonlau, M. and Welch, W.J. (1998), "Efficient global optimization of expensive black-box functions", *Journal of Global Optimization*, Vol. 13, pp. 455-92.

Knowles, J. (2006), "ParEGO: a hybrid algorithm with on-line landscape approximation for expensive multi-objective optimization problems", *IEEE Transactions on Evolutionary Computation*, Vol. 10 No. 1, pp. 50-66.

Lebensztain, L., Marretto, C.A.R., Costa, M.C. and Coulomb, J-L. (2004), "Kriging: a useful tool for electromagnetic devices optimisation", *IEEE Transactions on Magnetics*, Vol. 40 No. 2, pp. 1196-9.

Miettinen, K.M. (1999), *Nonlinear Multiobjective Optimization*, Kluwer Academic Publishers, Dordrecht.

Sasena, M.J., Papalambros, P. and Goovaerts, P. (2002), "Exploration of metamodelling sampling criteria for constrained global optimisation", *Engineering Optimization*, Vol. 34, pp. 263-78.

Sobester, A., Leary, S.J. and Keane, A.J. (2005), "On the design of optimization strategies based on global response surface approximation models", *Journal of Global Optimization*, Vol. 33, pp. 31-59.

Villemonteix, J., Vazquez, E. and Walter, E. (2008), "An informational approach to the global optimization of expensive-to-evaluate functions", *Journal of Global Optimization* (submitted).

About the authors

Glenn Hawe graduated in 2004 from the University of Warwick with a Masters degree in Mathematics and Physics. He is currently a KTP Associate working with Vector Fields Ltd, Oxford and the School of Electronics and Computer Science at the University of Southampton. He is registered at the university as a part-time PhD student, where his research area is cost-effective optimization techniques. Glenn Hawe is the corresponding author and can be contacted at: glenn.hawe@manchester.ac.uk

Jan Sykulski is a Professor of Applied Electromagnetics and Head of Electrical Power Engineering Group of the School of Electronics and Computer Science at the University of Southampton. His research interests are in the areas of power applications of high-temperature superconductivity, modelling of superconducting materials, advances in simulations of coupled field systems, development of fundamental methods of computational electromagnetics and new concepts in design and optimization of electromechanical devices. He has 224 publications listed on the official database of the University. He is a founding Secretary/Treasurer/Editor of International Compumag Society, a Visiting Professor at universities in Canada, Italy, Poland, France and China, a founding member of International Steering Committees of major international conferences, including Compumag, IEEE CEFC, IEE CEM, ISEF, EPNC and others. He is fellow of IEE, Institute of Physics and British Computer Society.

To purchase reprints of this article please e-mail: **reprints@emeraldinsight.com**
Or visit our web site for further details: **www.emeraldinsight.com/reprints**