International Journal of Applied Electromagnetics and Mechanics 46 (2014) 427–436
DOI 10.3233/JAE-141954
IOS Press

Considerations of uncertainty in robust optimisation of electromagnetic devices

Song Xiao, Yinjiang Li, Mihai Rotaru and Jan K. Sykulski*
Electronics and Computer Science, University of Southampton, Southampton, UK

Abstract. Due to unavoidable uncertainties related to material properties and manufacturing processes, the robustness of the optimal solution must be considered when designing electromagnetic devices. In this paper, the worst-case optimisation (WCO) and the worst-vertex-based WCO are proposed to evaluate the robustness of both performance and constraints under uncertainty. To reduce computing times when searching for the robust solution a predicted objective function is used, obtained with the help of a kriging algorithm which explores the searching space using the concept of rewards. Finally, to avoid some of the shortcomings of WCO, the concept of average performance evaluation is developed.

Keywords: Robust optimisation, CAD in electromagnetics, kriging

1. Introduction

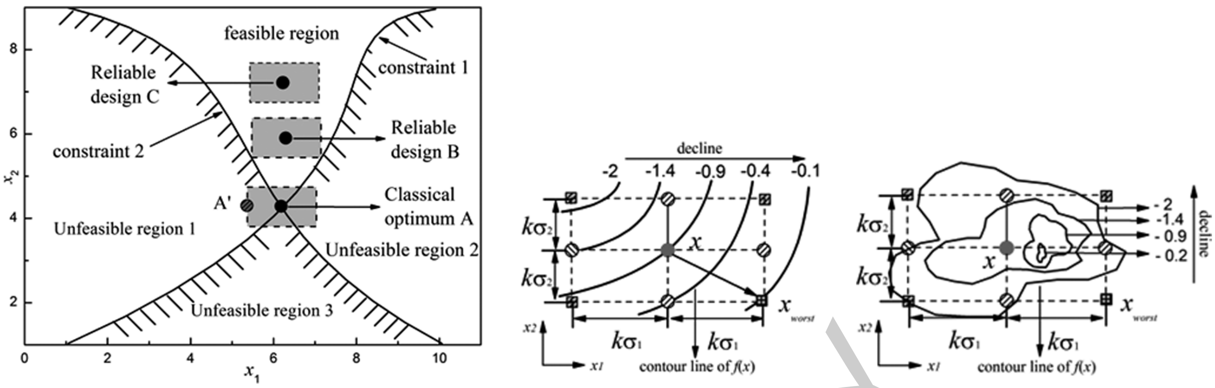
In electromagnetic design, uncertainties in design variables are inevitable, thus the ability to evaluate the robustness is critical while pursuing the theoretical optimum. This is particularly true when constrained optimisation is considered, as illustrated later by an example involving a quenching condition for a superconducting material, in order to maintain the solutions within the feasible region when perturbation occurs. The worst-case optimization (WCO) method has been selected to evaluate the accuracy of the prediction of the objective function provided by an improved kriging model, implemented for the sake of reducing computational effort associated with direct application of information produced by a time-consuming Finite Element Method (FEM). However, in this study some shortcomings of the WCO approach have been identified associated with the inability to assess the performance variation under conditions of uncertainty. Therefore the concept of average performance evaluation was suggested as an improved measure of robustness.

2. Robust optimization algorithms exploiting kriging modelling

In general, seeking the minimum (maximum) of an objective function, while the search space is restricted by certain constraints, is the aim of conventional optimization

$$\text{Minimize } f(x) \qquad \text{Subject to } g_i(x) \leqslant 0, i = 1, \ldots, m \tag{1}$$

*Corresponding author: Jan K. Sykulski, Electronics and Computer Science (ECS), University of Southampton, Southampton, SO17 1BJ, UK. Tel.: +44 2380 593 448; Fax: +44 2380 593 709; E-mail: jks@soton.ac.uk.



(a) (b) (c)

Fig. 1. (a) The constrained optimization problem; (b) The worst case optimization (WCO); (c) The worst-vertex-based WCO.

When designing real devices, however, many design variables are subject to specific uncertainties (manufacturing tolerances, variation of material properties, etc.). Hence the assessment of the influence caused by these uncertainties on the performance becomes essential in practical design problems. As a consequence, finding a theoretical optimum may not be sufficient and the robustness of the theoretical optimal solution needs to be considered as well. Algorithms applying different strategies to evaluate robustness, such as the sensitivity analysis [1], the worst-case optimization method [2–6], and the mean value and variance of performance [7–9], have been developed to assist in the design tasks (see also [19]). In our previous work [10] a multi-objective optimization method, which included sensitivity analysis using gradient index, was developed and demonstrated.

A widely used approach to evaluate the reliability of a robust solution is the worst-case method. This technique can maintain a certain level of robustness by avoiding solutions that may push the function into unfeasible region when searching for the optimum, as shown in Fig. 1(a). A set of typical examples is shown in Fig. 1(a) to illustrate the principle of the worst-case optimization technique. The theoretical optimum A, for example, may be abandoned in favour of a 'worse' solution B because the actual design A', which accounts for the uncertainties of the variables, might violate constraints and enter unfeasible area. The uncertainties can be defined mathematically as

$$U(x) = \{\xi \in R^n | x - k\sigma \leqslant \xi \leqslant x + k\sigma\} \tag{2}$$

where σ is standard deviation of uncertain variables and k is determined by confidence level [11]. The uncertainties may also be specified directly (e.g. as machining tolerances, say Δ). As an algorithm which can predict the worst scenario considering the uncertainties, as well as constraints with respect to specific designs, the worst case optimization (WCO) method [2–5] may be applied to analyse the reliability of the solution as follows

$$\text{Minimize } f_w(x) \equiv \max_{\xi \in U(x)} f(\xi) \quad \text{Subject to } g_{w,i}(x) = \max_{\xi \in U(x)} g(x) \leqslant 0, i = 1, \ldots, m \tag{3}$$

The worst values of the objective function and the i-th constraint function are chosen to substitute the original values of the nominal design x.

Numerical methods, such as finite elements, are often used when searching for the worst objective function value under imposed constraints which may be an extremely time-consuming process. To reduce the computational burden the worst-vertex-based WCO (W-WCO) [4] was proposed; this algorithm only needs to observe the vertices within the region restricted by uncertainties rather than evaluating every design value. For example, in the problem illustrated in Fig. 1(b), in addition to x there are 8 more points required, located at the corners and the middle of the specified boundary. However, in certain cases assessing only these 8 points might still not be sufficient. Figure 1(c) illustrates such a case were a large variation of the function will not be identified by the W-WCO method. Therefore finding a balance between mitigating the heavy computational burden of the original worst-case method and pursuing more detailed evaluation is the main issue to be addressed.

One of the possibilities is to employ the improved kriging method, assisted by a set of strategies capable of balancing exploration and exploitation [12,13,17] using the concept of rewards [14], which can be used to predict the objective function value instead of directly calculating it using computationally expensive FEM models. Based on such a kriging prediction, the worst-case method can be directly implemented. In other words, the WCO method uses the predicted information rather than the expensive FEM models. The accuracy of the predicted objective function using the improved kriging model has been considered in [10]. In this paper the suitability of directly using WCO with the predicted function model is discussed and demonstrated.

3. Kriging surrogate model

As a kind of regression model, kriging [15] is able to predict the shape of the objective function via spatial correlation of data using limited information. The accuracy of this prediction can also be estimated by kriging, which may be extremely helpful when making a decision where to place the next evaluation point at any stage of the optimization process. To accomplish this aim kriging needs to exploit the spatial correlation between the known points (vectors) of the objective function and all the unknown points, as well as the correlation between the known points (newly found points and initial sampling points), in order to build a correct surrogate model of the real objective function through interpolation. This relies on the linear regression model

$$\hat{y}(x) = \sum_{k=1}^{m} \beta_k f_k(x) + \varepsilon(x) \tag{4}$$

and the Gaussian correlation model

$$R\left(\varepsilon\left(x^i\right), \varepsilon\left(x^i\right)\right) = \prod_{k=1}^{n} e^{-\theta_k \left|x_k^i - x_k^j\right|^{p_k}}. \tag{5}$$

where the global function $\sum_{k=1}^{m} \beta_k f_k(x)$ and an additive Gaussian noise $\varepsilon(x)$ are integrated to the predicted value $\hat{y}(x)$ of the objective function; θ_k is the correlation amongst the data in k-direction and p_k determines the 'smoothness' of Eq. (5). The most popular correlation function is given by the Gaussian model where the value of p_k is simply taken as equal to 2. For a given set of data, the maximum likelihood estimation optimizes the value of θ and then the correlation model is brought into the regression model to evaluate the function with the best linear unbiased predictor [4].

Although kriging can potentially solve large multi-parameter optimization problems, it has some inherent limitations making the implementation difficult. In particular, for multi-parameter problems, the correlation model built by the kriging algorithm can grow very fast resulting in a 'combinatorial explosion' of correlation data filling very quickly the memory of standard computer workstations. As a result

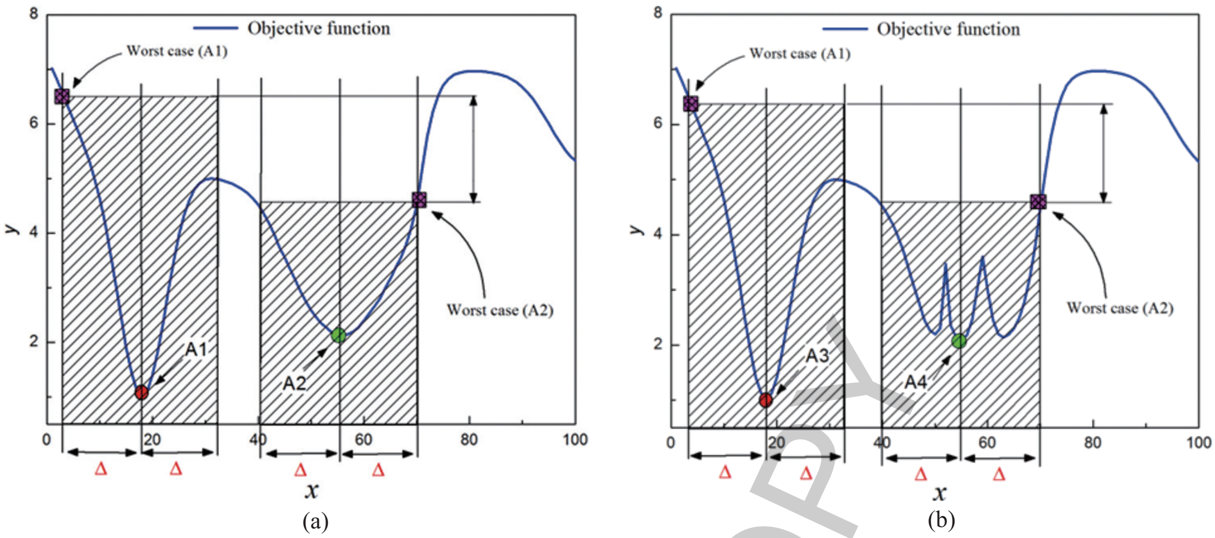


Fig. 2. (a) The objective function with a robust local minimum subject to uncertainty Δ; (b) The objective function with a non-robust local minimum subject to uncertainty Δ.

the process can become slow and inefficient. To solve this issue a scheme that adaptively partitions the correlation matrices was developed [10]. Using this approach the size of the data is managed to use efficiently the available memory throughout the iterative process of kriging. The scheme mentioned above has another advantage as the kriging predictor and Mean Squared Error are being calculated at the same time hence more computational time is saved in the process of building the kriging model. Therefore this modification makes kriging suitable for solving multi-parameter optimization problems and could be linked with WCO which needs detailed data to work effectively.

4. Average performance

As explained in the previous sections, the WCO method can be used to find robust solutions for a particular problem once the constrains and the uncertainties of the variables have been defined. In this section, however, we address the extreme case depicted in Fig. 2; it is argued that WCO on its own is not sufficient to find a reliable and robust solution. To deal with such a situation the concept of an average performance has been suggested and will now be explained. Figures 2(a) and (b) show two similar functions, otherwise identical, except the region around the points A2 and A4, respectively. If WCO is used to find a robust solution using the same uncertainties Δ for the case depicted in Fig. 2(a), point A2 will be found, whereas for the case shown in Fig. 2(b) point A4 is likely to be returned as a robust solution. It is clear from this example that A4 is by no means a robust solution and although this may be considered a somewhat extreme situation it illustrates the fact that the WCO algorithm cannot differentiate between a robust and non-robust solution in such cases. To resolve such problems the concept of average performance within the uncertainty range is introduced.

The initial idea of evaluating average performance was to simply calculate the average value of all the potential perturbed values in the uncertain region with respect to the solution. This, however, proved infeasible, as although the shape of the two objective functions is clearly different, they share some common characteristics as explained in Table 1.

Table 1
The common features of the two functions

Global minimum (A1 and A3)	Local minimum (A2 and A4)	Uncertainty Δ	Worst case (A1 and A3)	Worst case (A2 and A4)	Average value (A2 and A4)
$x = 18$, $y = 1.0216$	$x = 55$, $y = 2.12657$	15	$x = 3$, $y = 6.557$	$x = 55$, $y = 2.12657$	2.9556

By visual inspection we can see that the variation around the solution A4 is much more intense than around A2, but the average values within the shaded areas with middle-points A2 and A4 for the given uncertainty are the same and equal to 2.9556. This means that the average value criterion may not be useful in assessing the variation of the function close to the point of interest. Hence an average value of the gradient index (GI) [16] has been introduced as an alternative way of assessing the average performance. The average value of the gradient index is calculated as:

$$\text{Minimize Average } G\text{I}\,(x) = \left[\sum_{xi=xL}^{xi=xU} \max\left(\left|\frac{\partial f\,(x)}{\partial x_i}\right|\right)\right] \Big/ n \tag{6}$$

where $x_i \in R^2$ ($x_L \leqslant x_i \leqslant x_U$) is the i-dimensional design variable vector with lower and upper bounds x_L and x_U, respectively, and n is the sum of the design vectors. The sum of the maximum gradients is divided by the total number of design vectors. The average first-order gradient for the first case is then found to be 0.0594, while in the second case it is three times larger 0.1416. Therefore the average gradient index could be used as a more reliable criterion to evaluate the average performance. This criterion can therefore be combined with the WCO method to resolve difficult problems such as the one described by Fig. 2(b). By generalizing this methodology it can be argued that a robust optimization problem can be transformed into a three-objective optimization problem defined as

$$\text{Minimize } f(x)$$

$$\text{Minimize } f_w(x) \equiv \max_{\xi \in U(x)} f\,(\xi) \text{ Subject to } g_{w,i}(x) = \max_{\xi \in U(x)} g(x) \leqslant 0, i = 1, \ldots, m \tag{7}$$

$$\text{Minimize Average } G\text{I}\,(x) = \left[\sum_{xi=xL}^{xi=xU} \max\left(\left|\frac{\partial f\,(x)}{\partial x_i}\right|\right)\right] \Big/ n \tag{8}$$

5. Robust optimization algorithms exploiting kriging modelling

In our previous work [10] the robustness of an optimal solution was evaluated using the gradient index, where the task of robust optimization was transformed into a two-objective optimization. One objective was to minimize the difference between the absolute value of the largest and the smallest gradients within the uncertain range, called the 'sensitivity', while minimizing the objective function that was the second objective. Although the sensitivity calculated by the gradient index method is able to provide information on the rate of change of the objective function, the WCO method can also be employed to obtain similar information. However, as shown in the previous section, the WCO method has some limitations, especially for extreme cases (Fig. 2(b)). The average performance assessment described above can thus be added to the WCO method to improve the overall reliability of the result. To verify the concept and to analyse further the average performance criterion, two problems have been tested and the results are reported below.

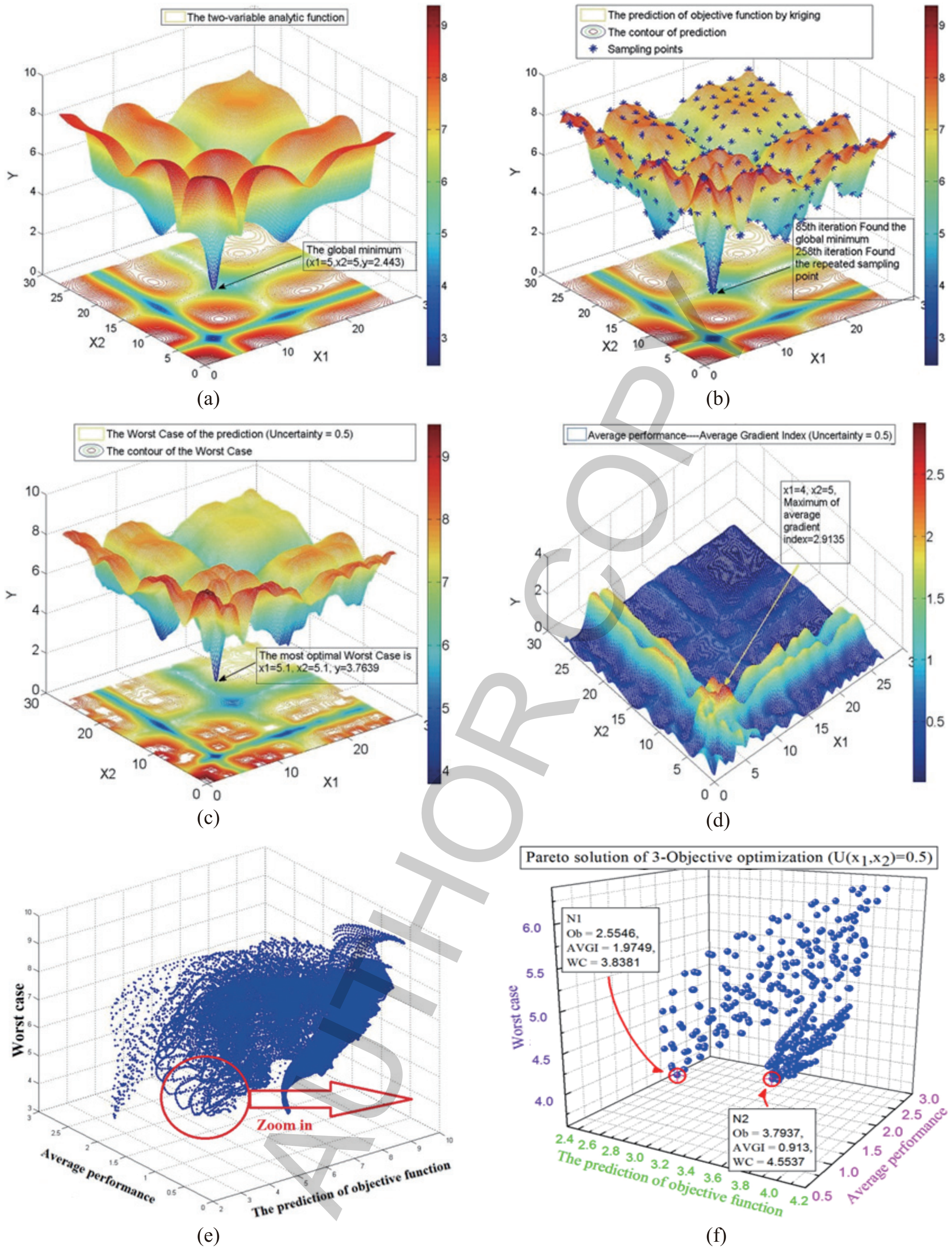


Fig. 3. (a) Analytic function; (b) The kriging prediction; (c) The performance of the WCO method; (d) Average performance; (e) The full-field solution for three objectives; (f) The zoomed-in optimal part.

First, the two-variable analytic function (8) depicted in Fig. 3(a), which was also used to assess the accuracy of the improved kriging model, has been tested.

$$f(x) = 10 - \sum_{i=1}^{n} \left[\frac{3.5}{1 + (x_i - 5)^2} + \frac{2.2}{1 + (x_i - 15)^2/10} + \frac{1.2}{1 + (x_i - 25)^2/30} \right], \quad (0 < x_i < 27) \tag{9}$$

Kriging with adaptive weighted expected improvement (AWEI) [16] has provided an approximation of this analytic function with a test step size of 0.1. Within 85 iterations, the kriging model can find the global minimum successfully; however – for a better approximation of the shape of the objective function – the model continued to run until the 285^{th} iteration (Fig. 3(b)).

The uncertainty with respect to variables is set as $\mathrm{U}(x_1, x_2) = 0.5$, and the WCO method is used to obtain the surface of the worst case for each solution as shown in Fig. 3(c). For this case the 'best' solution shifts from the theoretical optimum ($x_1 = 5$, $x_2 = 3$, $y = 2.443$) to the location ($x_1 = 5.1$, $x_2 = 5.1$, *y(worst case)* $= 3.7639$) which provides a more robust result, for the given conditions. If the values of the uncertainties were to keep increasing, up to a certain extent, ultimately the robust optimum would thoroughly shift from the sharp global minimum to one of the preferable local minima with higher robustness. Figure 3(d) depicts the average gradient index values in the search space. Finally, the full-scale ($0 < x_i < 27$) optimal solutions including all three objectives have been presented in Fig. 3(e). For clearer presentation of the pareto front, the full-scale version is zoomed in Fig. 3(f). Two typical pareto solutions are labelled in the zoomed-in graph: solution N1 delivers a more optimal value of the prediction of the objective value, while N2 offers a relatively better average performance.

6. Application to electromagnetic design

The second example tested with the proposed WCO procedure involves a multi-objective version of the TEAM 22 benchmark problem [18]. The full description of the TEAM problem 22 may be found elsewhere and will not be repeated here. The target for this problem is to achieve an arrangement of the two superconducting coils such that the stored energy within the system is $\mathrm{E_{ref}} = 180$ MJ while a minimal stray field $\mathrm{B_{stray}}$ is maintained. The objective function is defined as

$$\mathrm{OF} = \frac{\mathrm{B^2_{stray}}}{\mathrm{B^2_{norm}}} + \frac{|\mathrm{E} - \mathrm{E_{ref}}|}{\mathrm{E_{ref}}}, \tag{10}$$

where $\mathrm{B_{norm}}$ 3 μT and $\mathrm{B^2_{stray}} = \frac{\sum_{i=1}^{22} |\mathrm{B_{stray,i}}|^2}{22}$, subject to geometrical and 'quench' constraints. The approach taken here combines WCO method with kriging and commercially available FEM based software. A 2D model of the TEAM 22 problems is solved throughout this procedure. The three parameter case of TEAM 22, which includes three geometric variables R_2, H_2 and D_2, while R_1, H_1 and D_1 are fixed, has been tried under different settings of uncertainties. The uncertainties are assumed to exist in the current densities $J(J_1, J_2)$ of the two coils, because normally they are limited within certain range by a current controller for compensating perturbation. A constraint is imposed that the superconducting coils should not violate the quench condition which links together the value of the current density and the maximum value of magnetic flux density as follows

$$g_i(x) = |J_i| + 6.4 \cdot |B_{m,i}| - 54.0 \leqslant 0, \quad i = 1, 2 \tag{11}$$

Table 2
The initial setup of the prediction by the kriging model

	Three variables			Uncertainties		
	R_2 (m)	$h_2/2$ (m)	d_2 (m)	U(J_1, $J_2 = 0.1$) (MA/m^2)	U(J_1, $J_2 = 0.2$) (MA/m^2)	U(J_1, $J_2 = 0.35$) (MA/m^2)
Lower bound	3.03	0.211	0.367	J_1: 22.4 J_2: −22.6	J_1: 22.3 J_2: −22.7	J_1: 22.15 J_2: −22.85
Upper bound	3.13	0.281	0.397	J_1: 22.6 J_2: −22.4	J_1: 22.7 J_2: −22.3	J_1: 22.85 J_2: −22.15
Step size	0.01	0.007	0.003	0.02	0.04	0.07
No of steps	11	11	11	11	11	11

Table 3
Results for the case when uncertainty is 0.1

	R_2 (m)	$h_2/2$ (m)	d_2 (m)
P1	3.04	0.492	0.397
P2	3.09	0.464	0.382
P3	3.03	0.562	0.397

Table 4
Results for the case when uncertainty is 0.2

	R_2 (m)	$h_2/2$ (m)	d_2 (m)
P4	3.11	0.492	0.376
P5	3.09	0.478	0.385
P6	3.03	0.562	0.397

Table 5
Results for the case when uncertainty is 0.35

	R_2 (m)	$h_2/2$ (m)	d_2 (m)
P7	3.07	0.492	0.382
P8	3.03	0.562	0.397

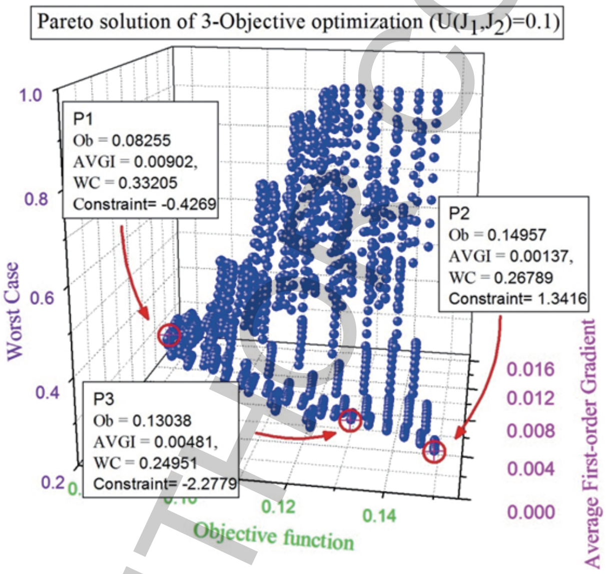


Fig. 4. The three objective optimization including (Worst case (WC), Average gradient index performance (AVGI), the prediction of objective functions (Ob)) (U(J_1, J_2) = 0.1).

Three tests were performed for this TEAM 22 problem. The initial data set-up, as well as the number of steps, for the three tests is listed in Table 2. One of the differences between the three sets is the size of the uncertainty; the first set had the uncertainty set as 0.1, the second test 0.2, while in the third test 0.35 was used. Another difference was in the number of iterations that were used to approximate the objective function by the kriging model with AWEI [10]. For the first data set (uncertainty 0.1) kriging generated 185 sampling points, for the second case (uncertainty 0.2) kriging produced 257 sampling points, while for the third set (uncertainty 0.35) only 188 points were necessary. The results returned by the WCO, coupled with kriging, for these three cases are summarised in Tables 3 to 5. Figures 4–6

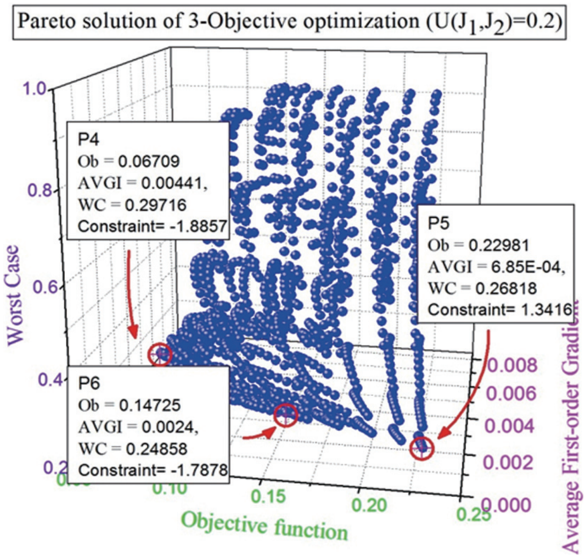


Fig. 5. The three objective optimization ($\mathrm{U}(J_1, J_2) = 0.2$).

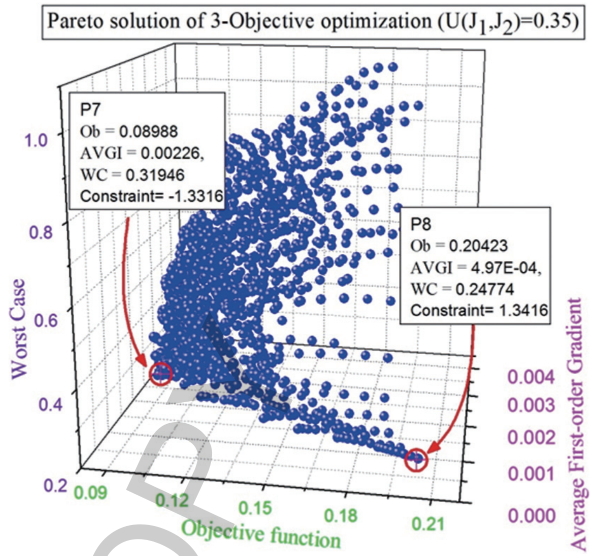


Fig. 6. The three objective optimization ($\mathrm{U}(J_1, J_2) = 0.35$).

show both the full scale and zoomed-in versions of the pareto fronts obtained for the three cases. In the figures and tables, solutions P1, P4 and P7 refer to the global optimum; P3 and P6 have the 'worst case' performance; while P2, P5 and P8 describe the best average gradient index solutions.

Unlike other reported work [16] that uses a stochastic optimization method to find the global optimum and then employs Monte Carlo method to explore the space around the global minimum, combined with WCO and the gradient index to judge the robustness of the solution, the method introduced here takes a holistic approach and explores the whole searching space. The kriging model allows comparison amongst several local minima (maxima) that may be more robust than the global optimum. Another major advantage of the procedure proposed in this work is the fact that it can be linked with any commercial electromagnetic design software giving more freedom to the designer.

7. Conclusions

The worst-case method (WCO) assisted by the prediction provided by the kriging model with AWEI has been proposed to solve robust optimization problems considering uncertainties of variables. The particular contribution of this paper is in using kriging prediction, rather than the computationally expensive finite element modelling, in assessing the robustness of the final design. The second contribution involves enhancing the worst case methodology through introducing the concept of average gradient index performance. Using this approach a conventional optimization problem, with constrains and uncertainties in variables, has been transformed into three-objective optimization with a relevant pareto front. The proposed algorithms have been verified by both numerical tests and a practical electromagnetic design problem described by TEAM 22 benchmark.

References

[1] P. Alotto, P. Molifino and G. Molinari, Optimization of electromagnetic devices with uncertain parameters and tolerances in the design variables, *COMPEL* **20**(3) (2001), 808–812.

[2] L. Egiziano, P. Lamberti, G. Spagnuolo and V. Tucci, Robust design of electromagnetic systems based on interval Taylor extension applied to a multiquadric performance function, *IEEE Trans Magn* **44**(6) (Jun 2008), 1134–1137.
[3] A. Ambrisi, A. Formisano and R. Martone, Tolerance analysis of NMR magnets, *IEEE Trans Magn* **46**(8) (Augest 2010), 2747–2750.
[4] G. Steiner, A. Weber and C. Magele, Managing uncertainties in electromagnetic design problems with robust optimization, *IEEE Trans Magn* **40**(2) (March 2004), 1094–1099.
[5] G.L. Soares, R.L. Adriano, C.A. Maia, L. Jaulin and J.A. Vasconcelos, Robust multi-objective TEAM22 problem: A case study of uncertainties in design optimization, *IEEE Trans Magn* **45**(3) (March 2009), 1028–1031.
[6] P. Alotto, C. Magele, W. Renhart, G. Steiner and A. Weber, Robust target functions in electromagnetic design, *COMPEL* **22**(3) (2003), 549–560.
[7] D.H. Jung and B.C. Lee, Development of a simple and efficient method for robust optimization, *Int J Numer Mech Eng* **53** (2002), 2201–2215.
[8] F.G. Guimaraes, D.A. Lowther and J.A. Ramirez, Multiobjective approaches for robust electromagnetic design, *IEEE Trans Magn* **42**(4) (April 2006), 1207–1210.
[9] X.J. Meng, S.H. Wang, J. Qiu, Q.H. Zhang, J.G. Zhu, Y.G. Guo and D.K. Liu, Robust multilevel optimization of PMSM using design for six sigma, *IEEE Trans Magn* **47**(10) (2011), 3248–3251.
[10] S. Xiao, M. Rotaru and J.K. Sykulski, Robust global optimization of electromagnetic designs utilizing gradient indices and kriging, *COMPUMAG 2013*, Budapest, Hungary, (30 June–4 July 2013).
[11] Z.Y. Ren, D.H. Zhang and C.S. Koh, Reliability evaluation of the electromagnetic device based on the second order sensitivity analysis, *Sixth International Conference ICEF*, Dalian, China, (2012).
[12] S. Xiao, M. Rotaru and J.K. Sykulski, Exploration versus exploitation using kriging surrogate modelling in electromagnetic design, *COMPEL* **31**(5) (2012), 1541–1551.
[13] S. Xiao, M. Rotaru and J.K. Sykulski, Adaptive weighted expected improvement with rewards approach in kriging assisted electromagnetic design, *IEEE Trans Magn* **49**(5) (2013), 2057–2060.
[14] A. Sykulski, N. Adams and N. Jennings, On-line adaptation of exploration in the one-armed bandit with covariate problem, *Proc of 9th Int Conf on Machine Learning and Applications* (2010).
[15] L. Lebensztajn, C.A.R. Marreto, M.C. Costa and J.-L. Coulomb, Kriging: A useful tool for electromagnetic device optimization, *IEEE Trans Magn* **40**(2) (2004), 1196–1199.
[16] Z.Y. Ren, M.T. Pham and C.S. Koh, Robust global optimization of electromagnetic devices with uncertain design parameters: Comparison of the worst-case optimization methods and multi-objective optimization approach using gradient index, *IEEE Trans Magn* **49**(2) (2013), 851–859.
[17] X. Song, M. Rotaru and J.K. Sykulski, Strategies for balancing exploration and exploitation in electromagnetic optimisation, *COMPEL* **32**(4) (2013), 1176–1188.
[18] P.G. Alotto, U. Baumgartner, F. Freschi, M. Jaindl, A. Kostinger, C. Magele and W. Repetto, TEAM workshop problem 22, http://www.compumag.org/jsite/images/stories/TEAM/problem22.pdf.
[19] L. Wu, S. Yang, Y. Bai and J. Machado, Efficient robust optimization based on polynomial chaos and tabu search algorithm, *International Journal of Applied Electromagnetics and Mechanics (IJAEM)* **39**(1–4) (2012), 145–150.