Hybrid Optimization in Electromagnetics Using Sensitivity Information from a Neuro-Fuzzy Model

Kashif Rashid, Jaime A. Ramírez, and Ernest M. Freeman

***Abstract*—The use of sensitivity information from a neuro-fuzzy model for the purpose of optimization is investigated in this paper. This approach permits the application of classic deterministic or hybrid optimization methods in establishing the global minimum of any approximated objective function using neuro-fuzzy modeling. For nondifferentiable functions this approach is of great benefit. An analytical problem and the TEAM 22 benchmark problem are investigated. Results using the genetic algorithm method and the sequential quadratic programming method in sequence show the usefulness of the formulation.**

***Index Terms*—Artificial intelligence, fuzzy neural networks, optimization methods, electromagnetic analysis.**

I. INTRODUCTION

IN engineering the objective function of complex multi-objective problems is commonly nonlinear, continuous, multivariate and constrained, which often is not readily differentiable. Such problems are typically optimized using stochastic methods (simulated annealing, etc.) [1]–[3], and deterministic methods (conjugate gradient methods, etc.) if at all applicable, tend to become trapped in local minima [4]. Stochastic methods however, have the disadvantage of requiring a much higher number of function evaluations in order to identify the global minimum. A means to overcome this high computational burden is to provide an approximation of the objective function such that future evaluations can be established at reduced cost [5]–[9]. These global response surface methods (GRSM's) are particularly suited for tackling multi-minima problems [5]. Although, the search is undertaken faster at a lower computational cost, the global minimum is often not exactly found and the result is not always reproducible in the same number of finite steps due to the very nature of the random search process. This has led to the development of hybrid strategies, where the optimization process is started with a stochastic method and after some iterations is switched to a deterministic method.

This paper examines the development of hybrid optimization strategies using sensitivity information derived from the fuzzy rule set of an approximated objective function using a neuro-fuzzy model [10], [11].

Manuscript received October 25, 1999. This work was supported by EPSRC, UK and in part by CNPq (Grant nos. 300353/97–9 and 451689/99) and by PUC-Minas, Brazil.

K. Rashid and E. M. Freeman are with the Department of Electrical & Electronic Engineering, Imperial College of Science, Technology & Medicine, Exhibition Road, London SW7 2BT, UK.

J. A. Ramírez is with PUC-Minas, Programa de Pós-Graduação em Eng. Elétrica, Av. Dom José Gaspar, 500, 30535-610 Belo Horizonte, MG, Brazil.

Publisher Item Identifier S 0018-9464(00)06699-1.

II. THE OPTIMIZATION MODEL

A trained neuro-fuzzy model, one which can predict the output of the model for any combination of input parameters, can be used to as an alternative to future experimental or finite element based simulations [12]. This empirical model provides an approximation to the real objective function and can be optimized using either stochastic or deterministic methods. Fig. 1 shows the optimization model in which a neuro-fuzzy model is designed and optimized iteratively.

Initially the system under investigation is sampled in parameter-space at iteration i. The location and number of samples is dictated by the sampling strategy; grid based, random, genetic or dimensional sampling, along with the desired number of points. The sampling matrix is used to initiate a function call to the function generator, which can be a computational source, a physical device or a combination of the two. The sampled matrix, describing the input–output relationship between model variables, is used to train a neuro-fuzzy model. Generally, the sampled matrix is partitioned into two sets; training and checking data. The neuro-fuzzy model is designed using training data and validated using checking error to ensure the model is not trained to over fit the model data. This empirical model then effectively replaces the actual function generator in the optimization process, obviating the need to run further computationally expensive simulations.

Thus, stochastic methods can be applied with relatively little cost since approximated functions evaluations are made at a fraction of the cost in comparison to the real function generator. Moreover, by coupling a deterministic method, the start point of which is the locality of the global minimum identified by the stochastic method, a hybrid approach can be implemented. An efficient and accurate search results since the time spent running a stochastic search is minimized and the pitfalls associated with poorly assigned starting points in a deterministic method are also overcome. The process is repeated with further function calls made to refine the area of interest until the convergence criterion is met. The final result is the perceived optimum.

Note that model sampling incurs the greatest time and computational cost which could be alleviated with the use of parallel processors or overcome if data were available directly from a design knowledge base. Also note, this paper deals with evenly partitioned neuro-fuzzy models in which the number and type of membership functions assigned to each input can be pre-defined by the user.

III. Results

A. Problem I

Firstly, an analytical test problem with a known solution is tested [13]:

$$\text{minimize : } \quad F(x) = 9x_1^2 + x_2^2 + 9x_3^2 \tag{1}$$

$$\begin{aligned} \text{subject to : } \quad & x_1x_2 - 1 \geq 0; \quad -10 \leq x_1 \leq 10 \\ & 1 \leq x_2 \leq 10; \quad -10 \leq x_3 \leq 1 \end{aligned} \tag{2}$$

The optimum is specified by [0.577 1.732 0] giving $F_{\min}$ = 6. The optimum is sought by iterative refinement using grid sampling at the first iteration followed by either dimensional sampling (see [9] for further details) or genetic sampling in subsequent iterations.

A neuro-fuzzy model with 4 Gaussian membership functions (MF's) assigned to each input is designed at each stage (gauss, 4), giving rise to 64 rules. The model is optimized using a hybrid approach. The genetic algorithm is initiated with a population of 600 chromosomes over 40 generations after which the SQP is invoked to isolate the global minimum. The progress of the optimum is given in Table I.

The convergence criterion is marked by the error column, which assesses the infinity norm between the optimum in consecutive iterations. Note that dimensional sampling necessitates 260 real function calls whilst genetic sampling requires only 185 samples. In each case the process terminates once the convergence criterion is ($< 5 \times 10^{-3}$).

B. Problem II

As an example of a numerical problem, the Team 22 SMES benchmark problem is investigated [14]. The aim of this problem is to optimize the Super-conducting Magnetic Energy Storage configuration, see Fig. 2, with respect to three individual objectives, each of which must be minimized. These are to ensure the stray field is minimal (f1), the stored energy is near 180MJ (f2) and that the quench condition ensuring superconductivity of the SMES arrangement is met (f3). The problem is composed of eight parameters in total, though in this case only the three variable discrete case is considered.

The combined cost function is given below:

$$F = k_1 f_1(B_{\text{stray}}) + k_2 f_2(\text{Energy}) + k_3 f_3(|B|_{\text{calc}}) \tag{3}$$

where k_1, k_2 and k_3 are weights describing the relative importance of each of the sub-objectives. It has been proposed that $k_1 = k_2 = 1$ and $k_3 = 5$ are suitable settings for the weights in order to accentuate the valleys of interest [6]. Thus (3) is given by:

$$F = \frac{B_{\text{stray}}^2}{B_{\text{norm}}^2} + \frac{|\text{Energy} - E_{\text{ref}}|}{E_{\text{ref}}} + 5(|B|_{\text{pred}} - |B|_{\text{calc}})^2 \tag{4}$$

where

$$E_{\text{ref}} = 180 \text{ MJ} \quad B_{\text{stray}}^2 = \sum_{i=1}^{22} \left| B_{\text{stray}_i} \right|^2 \quad \text{and}$$

$$|J| = (-6.4|B|_{\text{pred}} + 54.0)\text{A/mm}^2 \quad \text{with}$$

$$|B|_{\text{pred}} = 4.9219\,\text{T}.$$

Table II shows the ranges of the free variables and Table III the values of the fixed ones.

The SMES model was modeled using [15] and optimized iteratively using the schema presented in Fig. 1. At the first iteration grid sampling with five points on each axis is performed yielding a data set of 125 points. Initial modeling indicates that a model composed of 4 gauss MF's with T = 75% has the lowest RMSE, assuming not all the data is used for training. This model is automatically selected for further refinement, see Table IV.

The trained model is optimized using a hybrid approach. Firstly, a genetic algorithm is executed with a population of 500 chromosomes over 50 generations and then the SQP method is used to accurately isolate the optimum. This point subsequently becomes the focal point of the search in the following iteration, in which the search space is collapsed and additional samples are taken to refine the area around the perceived optimum. Again genetic sampling is employed at subsequent iterations for isolating future sample points. A population of evolving chromosomes is used to identify the troughs and valleys in an n-(= 4) dimensional hyper-surface. This approach has the benefit of implicitly inferring domain knowledge into the sampling process, which is not available in deterministic (grid, dimensional) or stochastic (random) sampling strategies, which are as a result prone to redundancy. Furthermore, constraints are also implicitly accounted in the sampling process, as is evident from the results of the previous example. An effective set of samples, one in which the conditions are evolving to lower depths yet maintaining a degree of diversity, is governed by the number of generations taken before the samples are taken. This issue is problem dependent and can be fixed by fixing the number of generations before sampling or by ascribing the minimum variance of fitness between the best and worst samples in any generation. In this example, sampling is undertaken after 10 generations from an evolving population of 500. The number of samples selected at each iteration is set to 50. Results for iteratively refined optimization employing (gauss, 4) neuro-fuzzy models are shown in Table V.

The stopping criterion for the search must be specified by the user. This can take several forms, the simplest of which is to examine the infinity norm of the optimum between successive iterations, as in the previous example. An alternative, as shown in the grey cells in Table V, is to examine the infinity norm between models designed using T = 75% and T = 100% at each iteration. That is, to compare the difference between the best generalized neuro-fuzzy model to one likely to have over fit the model data. When the two models converge, it is evident that further sampling in the subsequent iteration would have little impact on the search for the optimum and as such may prove computationally expensive with little benefit. In either case, the search terminates once the error factor falls below the stipulated stopping criterion, in this case 5×10^{-3}. Hence, the optimum is isolated after 5 iterations, having required 325 points in total. Note that the final result in Table V is for the continuous case. Results for the discrete case, including actual values from the function generator are given in Table VI.

In Table VI two discrete solutions are given: $d1$ and $d2$. $D1$ is the best discrete solution from the neuro-fuzzy model whereas $d2$ is the better solution from the actual function generator. Note

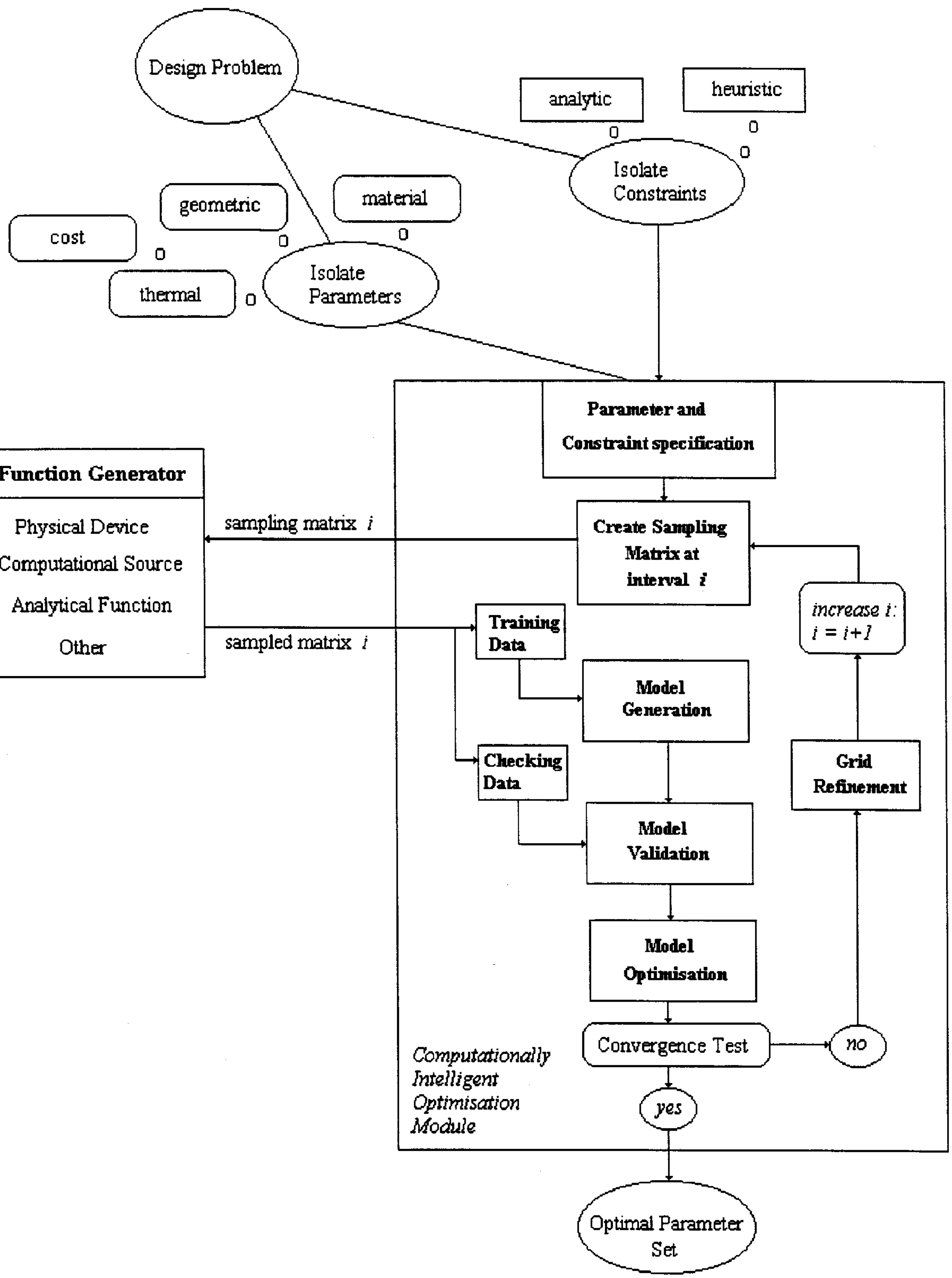

Fig. 1. Neuro-fuzzy modeling with hybrid optimization.

that F, the value of the approximated cost function is the global minimum in the continuous case and is therefore lower than either of the discrete solutions. From the actual function generator however, both the discrete solutions are lower than the value for the continuous case. This is due to the implicit error inherent in the final trained neuro-fuzzy model, which suggests that there is room for improvement in the model. However, given the values of the optimum in all cases, this will not necessarily be worth the additional effort. Also it worth noting that although mesh design is consistent with respect to this particular problem, variation should be expected when comparing other optimization methods using finite-element-based solutions.

TABLE I
VARIABLE RANGES

itn	Pts	x_1	x_2	x_3	F	error
Dimensional Sampling						
1	125	0.6334	1.5789	0.0872	14.2670	1.0000
2	27	0.4288	2.3322	-0.1511	-1.0156	0.7533
3	27	0.5432	1.8410	-0.2251	6.3705	0.4911
4	27	0.5639	1.7732	-0.0238	5.7584	0.2014
5	27	0.5709	1.7517	-0.0183	5.8979	0.0216
6	27	0.5717	1.7491	-0.0185	5.9453	0.0025
Genetic Sampling						
1	125	0.6334	1.5789	0.0872	14.2670	1.0000
2	20	0.5821	1.7180	-0.0122	6.0749	0.1391
3	20	0.5752	1.7387	-0.0056	6.0277	0.0206
4	20	0.5759	1.7363	-0.0047	6.0162	0.0024

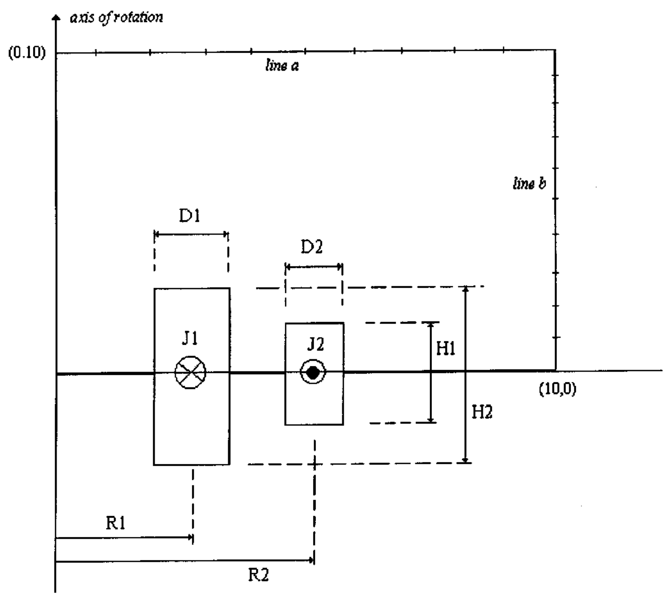

Fig. 2. The SMES model.

TABLE II
VARIABLE RANGES

Variable	R_2	H_2	D_2
Min	2.6	0.408	0.1
Max	3.4	2.2	0.4
step size	0.01	0.007	0.003

TABLE III
FIXED VARIABLES

variable	R_1	H_1	D_1	J_1	J_2
units	m	m	m	MA/m^2	MA/m^2
value	2.0	0.8	0.27	22.5	-22.5

TABLE IV
NEURO-FUZZY MODEL QUALITY AT ITERATION 1

rmse	gauss,3	gauss,4	gbell,3	gbell,4
T=50	1.0702	0.9858	1.0267	1.1266
T=67	1.1126	0.8589	1.1275	0.7827
T=75	0.7170	0.5137	0.7180	0.5390
T=80	13.6440	11.1860	13.8910	12.2270
T=100	0.3896	0.0477	0.4007	0.0538

TABLE V
ITERATIVE OPTIMISATION

itn	pts	R_2	H_2	D_2	F_{min}	RMSE error
1	125	2.6000	1.5154	0.2019	-1.3147	0.5137
		3.4000	1.4521	0.1000	2.5326	0.8000
2	50	2.8948	0.6415	0.3672	-3.9017	0.5899
		2.8778	0.5666	0.3889	-0.5059	0.0749
3	50	3.0976	0.4897	0.3683	-0.9870	0.7760
		3.0967	0.7178	0.2811	-1.7321	0.2281
4	50	3.1463	0.6576	0.3004	-3.3283	0.6914
		3.1005	0.6984	0.2836	-1.2629	0.0458
5	50	3.0533	0.4973	0.4000	-0.0837	0.6506
		3.0528	0.4949	0.4000	-0.3171	0.0024

TABLE VI
SMES OPTIMISATION RESULTS

Gauss 4,75	Continuous	Discrete	
	c1	d1	d2
R2	3.0533	3.05	3.05
H2	0.4973	0.499	0.492
D2	0.4000	0.400	0.400
F	-8.3652e-2	-8.2585e-2	-7.7397e-2
RMSE	0.6506	0.6506	0.6506
Bstray	1.1420e-6	1.1627e-6	9.3632e-7
Energy	9.0396e7	9.0347e7	8.9878e7
max|B|	4.8818	4.8925	4.8644
f1	1.2689e-1	1.2919e-1	1.0404e-1
f2	4.3962e-3	3.8505e-3	1.3527e-3
f3	1.6083e-3	8.6031e-4	3.2985e-3
F = f1+f2+f3	1.3289e-1	1.3390e-1	1.0869e-1
F = f1+f2+5f3	1.3933e-1	1.3734e-1	1.2188e-1

IV. DISCUSSION

The two problems presented demonstrate the merit of a hybrid optimization approach for tackling multi-minima problems. The search proves efficient, robust and accurate, since the run-time of a stochastic method is minimized and the deterministic method no longer requires a start point to be specified. This facility is made possible with the availability of sensitivity information from the neuro-fuzzy model.

It is also shown that iterative-based design aids fast- automated function approximation and optimization. Clearly, the main computational cost is associated with model sampling. As such, efficient sampling methods should be employed which minimize the number samples whilst maximizing the amount of information obtained. Genetic sampling proves particularly suitable in this vein by implicitly inferring domain knowledge during the sampling process and thus minimizing redundancy by only focusing on the areas of interest. Note that, if available, parallel processors could be employed to speed up the sampling process. That is, the sequential function call could be converted to a concurrent one. It is worth noting however, that evenly partitioned neuro-fuzzy models are suitable for problems with few dimensions (<5) since the number of model parameters, rules and therefore the required number of samples become prohibitively high. Generally, more complex models can be similarly tackled (Fig. 1) but relying on clustered neuro-fuzzy models, which attempt to represent the precise nature of the system under investigation by clustering model data into a number of key rules [16], at the design stage.

V. Conclusion

This paper has shown how sensitivity information extracted from a fuzzy rule set of an evenly partitioned neuro-fuzzy model can be employed to facilitate the application of hybrid optimization methods. This has the benefit of improving the accuracy and computational efficiency of the search process. Furthermore, iterative model design and optimization not only facilitates this process but also provides the means, at a little cost, for fast-automated function approximation and optimization. Hence, if an approximated function of sufficient accuracy can be realized, the method is suitable for optimising complex nonlinear non-differentiable multi-minima problems, which can be numerical or analytical.

References

[1] G. F. Uler, O. A. Mohammed, and C. S. Koh, "Design optimization of electrical machines using genetic algorithms," *IEEE Trans. Magn.*, vol. 31, no. 3, pp. 2008–2011, 1995.

[2] K. Hameyer and R. Hanitsch, "Numerical optimization of the electromagnetic field by stochastic search and MEC-model," *IEEE Trans. Magn.*, vol. 30, no. 5, pp. 3431–3434, 1994.

[3] C. Magele, K. Preis, W. Renhart, R. Dyczij-Edlinger, and K. Richter, "High order evolution strategies for the global optimization of electromagnetic devices," *IEEE Trans. Magn.*, vol. 29, no. 2, pp. 1775–1778, 1993.

[4] R. Fletcher, *Practical Methods of Optimization*. New York, NY, USA: John Wiley & Sons, 1980.

[5] P. Alotto, A. Caiti, G. Molinari, and M. Repetto, "A multiquadrics based algorithm for the acceleration of simulated annealing optimization procedures," *IEEE Trans. Magn.*, vol. 32, no. 3, pp. 1198–1201, 1996.

[6] T. Ebner, C. Magele, B. R. Branstatter, and K. R. Richter, "Utilizing feed forward neural networks for acceleration of global optimization procedures," *IEEE Trans. Magn.*, vol. 34, no. 5, pp. 2928–2931, 1998.

[7] T. Ishikawa and M. Matsunami, "An optimization method based on radial basis function," *IEEE Trans. Magn.*, vol. 33, no. 2, pp. 1868–1871, 1997.

[8] A. Abakar, J. L. Coulomb, and Y. Marechal, "Radial basis function network for acceleration of genetic algorithm optimization procedures," in *8th Intl. IGTE Symp. on Num. Field Calculation in Electrical Eng.*, Graz., Austria, 1998, pp. 33–37.

[9] K. Rashid, J. A. Ramirez, and E. M. Freeman, "The application of computational intelligence techniques to data modeling and optimization in electromagnetics," in *8th Intl. IGTE Symp. on Num. Field Calculation in elec. Eng.*, Graz., Austria, 1998, pp. 162–167.

[10] ——, "Derivative extraction from a neuro-fuzzy model," *Compel Journal*, vol. 19, no. 3, 2000, to be published.

[11] ——, "A general approach for extracting sensitivity analysis from a neuro-fuzzy model," in *Proceedings Compumag'99 Conference*, Sapporo, Japan, 1999.

[12] J. S. R. Jang and C. T. Sun, *Neuro-fuzzy and soft Computing*, NJ, USA: Prentice Hall, 1997.

[13] W. Hock and K.. Schittkowiski, *Test Examples for Nonlinear Programming Codes*. New York, NY, USA: Spring-Verlag, 1980.

[14] "TEAM Benchmark Problem 22 Definition: ,", http://www.igte.tu-graz.ac.at/team.

[15] MagNet6.4, Infolytica Corp., Montreal, Canada, 1999.

[16] S. Chiu, "A cluster estimation method with extension to fuzzy model identification," *IEEE Trans. Fuzzy Systems*, vol. 2, no. 3, pp. 1240–1245, 1994.