INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2001; **52**:219–238 (DOI: 10.1002/nme.215)

Parameter identification with weightless regularization†

Tomonari Furukawa*,‡

School of Aerospace, Mechanical and Mechatronic Engineering, J04, University of Sydney, NSW 2006, Australia

SUMMARY

Although the regularization increased the popularity of parameter identification due to its capability of deriving a stable solution, the significant problem is that the solution depends upon the regularization parameters chosen. This paper presents a technique for deriving solutions without the use of the parameters and, further, an optimization method, which can work efficiently for problems of concern. Numerical examples show that the technique can efficiently search for appropriate solutions. Copyright © 2001 John Wiley & Sons, Ltd.

1. INTRODUCTION

While forward computer analysis assists many industrial fields finding the result of actual phenomena from its cause, inverse analysis plays an important role in fields where the cause is to be derived from its result [1]. Among inverse problems, a typical inverse problem in industry is to identify the continuous parameter set of a continuous deterministic mathematical model, given the corresponding set of measured data. Most often, the parameter identification problem is converted into an objective function consisting of the measured data and the whole or partial model, and is solved by minimizing the function using an optimization method. In accordance, various calculus-based optimization methods have been used, depending upon the characteristics of the function ([2, 3] and references therein). The problem of this approach is, however, that the solution often cannot be obtained if measurement data and/or the model contain large errors [4] as this makes the objective function complex.

One of the approaches to overcome this problem is to add a regularization term [5], which normally consists of a function multiplied with weighting factors such as a regularization parameter, to the objective function. This term makes the functional smooth, so that a conventional calculus-based optimization can obtain an appropriate parameter set in a more stable fashion. Nevertheless, the solution obtained depends upon the selection of the weighting

*Correspondence to: Tomonari Furukawa, School of Aerospace, Mechanical and Mechatronic Engineering, J04, University of Sydney, NSW 2006, Australia
†This paper was partially presented at the Third International Conference on Inverse Problems in Engineering, Port Ludlow, WA, U.S.A., in June of 1999.
‡E-mail: tomo@acfr.usyd.edu.au

Received 22 May 1999
Revised 8 July 1999

Copyright © 2001 John Wiley & Sons, Ltd.

factors, and most of the research reports leave the selection for further studies, showing results only with a couple of selections [6, 7].

Techniques for finding the best value of a weighting factor has been proposed only by several researchers to the best of the author's knowledge. Conventional techniques include Morozov discrepancy principle [8], which obtains a regularization parameter based on minimization of an error criterion, and the generalized cross validation [9], which is derived from a statistical method. Later, Kitagawa [10, 11] proposed a technique based on the sensitivity of the regularization term with respect to the regularization parameter. Reginska [12] considered the maximiser of the L-curve, defined by Hansen [13], as the optimal parameter. Kubo *et al.* [14] also proposed a technique using singular value decomposition, while Zhuang and Zhu [15] proposed a multi-time-step method for inverse problems involving systems consisting of partial differential equations. The comparison of some of the techniques can be found in References [16, 17]. Despite their good performance to some degree, the fundamental question common in all the techniques is whether the automatic determination of a single solution by computation is necessary, as the solution of the inverse problem will never be known in nature unlike the forward analysis. In addition, an additional parameter must often be introduced to find the best regularization parameter, the solution being again dependent on the additional parameter.

Meanwhile, multi-objective optimization methods have been proposed for solving multi-objective design optimization problems [18–21]. These methods allow the design parameters to be optimized without weighting factors on design criteria such as weight and energy consumption. The solution of this vector functional formulation is henceforth represented as a space, namely the solution space, rather than a point, and the methods try to find a set of admissible solutions in the solution space.

Due to the derivation of multiple solutions and the possible complexity of the objective functions, the methods are mostly based on the evolutionary algorithms (EAs) [22], which execute robust search from multiple search points for single objective optimization. Many EAs are, however, very robust at the expense of efficiency in contrast to the conventional calculus-based methods, so that they are not inefficient for parameter identification problems of concern, which has a relatively simple formulation; i.e. composed of a continuous deterministic objective function with continuous search space. Moreover, since these algorithms find only a fixed number of solutions in the solution space, the solutions are sparse and not well distributed in the solution space.

In this paper, a technique for solving a regularized parameter identification problem without weighting factors is first proposed. In this technique, regularization terms are each formulated as another objective function, and the multi-objective optimization problem is solved by a multi-objective optimization method. Furthermore, a multi-objective optimization method termed Multi-Objective Continuous Evolutionary Algorithm (MCEA) is proposed to find the solutions for this class of problems efficiently. The algorithm is also formulated such that its solutions can describe the solution space to be derived.

Section 2 deals with the overview of the parameter identification, and the proposed weightless regularized identification technique is presented in Section 3. The proposed multi-objective optimization method is described in Section 4, and Section 5 presents numerical examples in order to investigate its superiority to the conventional techniques and its effectiveness as a parameter identification technique. In the first three subsections of Section 5, the performance of the proposed technique is tested with explicitly defined objective functions, and the last

Copyright © 2001 John Wiley & Sons, Ltd.

subsection deals with the parameter identification of a viscoplastic material for practical use. The final section summarizes conclusions.

2. PARAMETER IDENTIFICATION

2.1. Problem formulation

Suppose that we have a set of experimental data $[\mathbf{u}_i^*, \mathbf{v}_i^*]$, where $\mathbf{u}_i^* \in U$ and $\mathbf{v}_i^* \in V$, and the corresponding model $\hat{\mathbf{v}}$ having parameters $\mathbf{x} \in X$, the experimental data can be related to the model by

$$\hat{\mathbf{v}}(\mathbf{u}_i^*, \mathbf{x}) + \mathbf{e}_i = \mathbf{v}_i^* \tag{1}$$

where $\mathbf{e}_i$ represents the sum of the model errors and measurement errors:

$$\mathbf{e}_i = \mathbf{e}_i^{\text{mod}} + \mathbf{e}_i^{\text{exp}} \tag{2}$$

The parameter identification is typically defined to identify the continuous vector in engineering problems $X \subseteq R^n$, given a set of continuous experimental data, $U, V \subseteq R^n$. In order to solve it, a parameter identification problem is often converted to the minimization of a continuous functional:

$$f(\mathbf{x}) \to \min_{\mathbf{x}} \tag{3}$$

where $f : R^n \to R$. The parameter set minimizing such an objective function is to be found within a search space:

$$\mathbf{x}_{\min} \leqslant \mathbf{x} \leqslant \mathbf{x}_{\max} \tag{4}$$

where $[\mathbf{x}_{\min}, \mathbf{x}_{\max}] = X$. In this formulation, the solution of the identification problem is said to exist if there is at least one minimum within range (4), and the solution is said to be unique if there is only one minimum within the range.

As an example for functional (3), consider the popular method of least squares, the objective function of which is often represented as

$$f(\mathbf{x}) = \sum_{i=1}^{n} \|(\hat{\mathbf{v}}(\mathbf{u}_i^*, \mathbf{x}) - \mathbf{v}_i^*)\|^2 \tag{5}$$

It is clearly seen that the objective function consists of the model and the measurement data, thus the shape of the objective function depending upon them. The difficulty of the parameter identification is therefore that the objective function can become complex if the model and measured data contain considerable errors. It is more apparent when the number of measured data is small.

On the other hand, the majority of optimization methods can consistently find a global minimum only if the objective function is near-convex. Otherwise, the solution may diverge or vibrate depending on the initial search point chosen *a priori*. The approach for overcoming this problem is to introduce an additional term to the objective function in order to make it near-convex. This gives rise to the regularization described in the next section.

Copyright © 2001 John Wiley & Sons, Ltd.

2.2. Regularization

The complexity of the objective function in other words means that even a small change of the parameters may lead to a significant change to the functional to be minimized, and stabilization techniques are often termed as the regularization. In the Tikhonov regularization [23], which is the most popular regularization technique, the objective function is transformed into

$$\Pi(\mathbf{x}) = f(\mathbf{x}) + \alpha\Lambda(\mathbf{x}) \rightarrow \min_{\mathbf{x}} \tag{6}$$

where α and $\Lambda(\mathbf{x})$ are often termed as the Tikhonov regularization parameter and the Tikhonov regularization function. Assume that the solution is known to be adjacent to $\mathbf{x}^*$, the regularization term may be given by

$$\Lambda(\mathbf{x}) = \|\mathbf{K}(\mathbf{x} - \mathbf{x}^*)\|^2 \tag{7}$$

where $\mathbf{K}$ is a weighting matrix. Matrix $\mathbf{K}$ is most often set simply to the unity matrix unless some information is available.

It is therefore clear that the solution relies on the selection of parameter α, and the significant problem is that it is not easy to find the best parameter. In addition, if a conventional optimization method is used to solve this problem, the introduction of the additional term may still make the objective function non-convex, the solution resulting in vibration or divergence. The next section will present a technique to overcome the problem.

3. WEIGHTLESS REGULARIZATION

3.1. Problem formulation

The only way for finding solutions that do not depend upon the weighting factors is to remove them from the formulation, and we hereby propose a multi-objective formulation. With the unity weighting matrix of $\mathbf{K}$, Tikhonov regularization parameter α is the only weighting factor, and the objective of the problem is thus expressed as

$$\mathbf{f}(\mathbf{x})^{\mathrm{T}} = [f(\mathbf{x}), \|\mathbf{x} - \mathbf{x}^*\|^2] \rightarrow \min_{\mathbf{x}} \tag{8}$$

where $\mathbf{f}(\mathbf{x})^{\mathrm{T}} : R^n \rightarrow R^2$. If the weighting matrix is diagonal,

$$\mathbf{f}(\mathbf{x}) = [f(\mathbf{x}), \|x_1 - x_1{}^*\|^2, \ldots, \|x_n - x_n{}^*\|^2] \tag{9}$$

where $\mathbf{f}(\mathbf{x}) : R^n \rightarrow R^{1+n}$. The regularized parameter identification formulated as a multi-objective optimization problem is conclusively characterized as that:

- the problem is multi-objective,
- the objective function $f(\mathbf{x})$ is continuous but can be complex such as non-convex,
- the search space is continuous.

The first important question is what can be the solutions of this multi-objective optimization problem, and the next subsection will present the Pareto-optimality accordingly.

Copyright © 2001 John Wiley & Sons, Ltd.

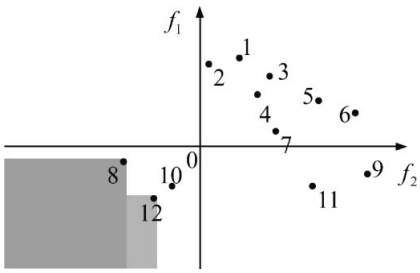

Figure 1. Pareto-optimal set.

3.2. Problem solution

While the single-objective optimization tries to look for a single solution, multi-objective optimization needs to derive a solution space, and this space is the natural solution to be obtained for the weight-independent regularised parameter identification formulated as a multi-objective optimization problem. Any point in the space satisfies Pareto-optimality, which was introduced in the field of economics a century ago, so the objective of the multi-objective optimization problem is converted to finding a set of Pareto-optimal solutions as equivalently as possible to the solution space.

Consider a problem where we have m objective functions, $f_k : R^n \to R,\ k = 1, \ldots, m$:

$$\mathbf{f}(\mathbf{x})^{\mathrm{T}} = [f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})] \to \min_{\mathbf{x}} \tag{10}$$

A decision vector $\mathbf{x}_u \in R^n$ is said to be Pareto-optimal if and only if there is no vector $\mathbf{x}_v \in R^n$ for which $\mathbf{v} = \mathbf{f}(\mathbf{x}_v) = (v_1, \ldots, v_n)$ dominates $\mathbf{u} = \mathbf{f}(\mathbf{x}_u) = (u_1, \ldots, u_n)$, i.e. there is no vector x_v such that

$$v_i \leqslant u_i, \quad \forall i \in \{1, \ldots, n\} \wedge v_i < u_i, \quad \exists i \in \{1, \ldots, n\} \tag{11}$$

Figure 1 illustrates an example where $\mathbf{x}_8$ and $\mathbf{x}_{12}$ satisfy Equation (11). The optimization method to find well-distributed Pareto-optimal solutions in an efficient and robust way will be presented in Section 4, and the next subsection deals with how to select a final solution from the Pareto-optimal set.

3.3. Determination of a solution

The great advantage of the proposed formulation is that we can select a single solution from a set of Pareto-optimal solutions after optimization rather than optimise a single solution with a specified weighting factor. Figure 2 shows the proposed process of the selection of a single solution from a parameter identification problem where objective function f_1 with regularization term f_2 is minimised to identify three parameters $[x_1, x_2, x_3]$. First, the user selects a Pareto-optimal solution in function space by considering the distribution of the Pareto-optimal solutions and the importance of the regularization term to the problem or by using a technique for finding the best regularization parameters. Note that the user may use a function space additionally defined to see other decision criteria such as the error and

Copyright © 2001 John Wiley & Sons, Ltd.

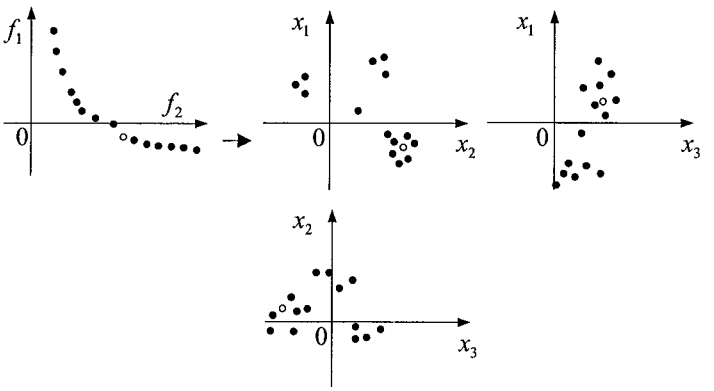

Figure 2. Process of deriving a single solution.

L-curve. The solution is then viewed in each two-dimensional parameter space to check visually whether the solution meets the expectation of the user. If the solution is not the expected one, the user selects a different Pareto-optimal solution, and the same process is repeated until the desired solution is obtained.

The number of graphs with three parameters is thus three and that with n parameters will be $\frac{1}{2}n(n-1)$. If the number of parameters is considerably large, the visual decision making is no longer possible, and the final solution must be selected automatically. The selection may be achieved mathematically or by artificial intelligence techniques such as expert systems, fuzzy logic or neural networks. We shall not discuss this further as it is out of scope of the paper.

3.4. Parameter identification process

Figure 3 shows the process of multi-objective optimization, while the process in case of the single-objective optimization is shown in Figure 4 for comparison. It is clearly seen that the optimization process in the single-objective optimization is within the loop and thus has to be repeated with a different set of weighting factors whereas the multi-objective optimization requires only one execution. This means that multi-objective optimization is much superior to the single-objective optimization in efficiency, provided that one execution in terms of the multi-objective optimization results in a similar Pareto-optimal set after a number of single-objective optimizations each with different weighting factors.

4. MULTI-OBJECTIVE OPTIMIZATION

4.1. Fundamentals

As a method to find a well-distributed Pareto-optimal set robustly and efficiently, MCEA proposed here is basically represented by the following four characteristics; the

Copyright © 2001 John Wiley & Sons, Ltd.

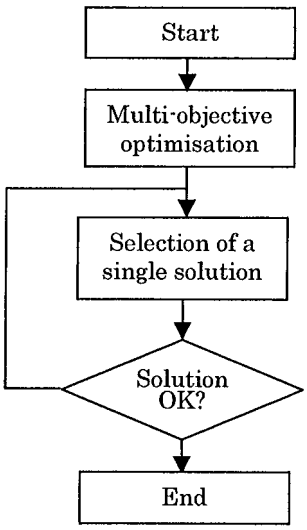

Figure 3. Deriving a solution in multi-objective optimization.

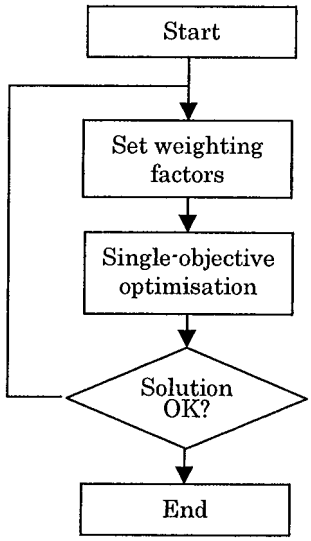

Figure 4. Deriving a solution in single-objective optimization.

method:

- searches with multiple points such that it can find multiple Pareto-optimal solutions,
- adopts probabilistic direct search algorithms based on evolutionary computation for robustness,
- implements the continuous representation of the points, continuous search formulation and continuous evaluation for efficiency,
- stores all Pareto-optimal solutions historically generated to grasp the whole solution space.

Figure 5 shows the fundamental structure of the proposed method. First, a population of individuals, each represented by a continuous vector, is initially (generation $t=0$) generated at random, i.e.

$$P^t = \{\mathbf{x}_1^t, \ldots, \mathbf{x}_\lambda^t\} \in (R^n)^\lambda \tag{12}$$

where λ represent the population size of parental individuals [24]. Each vector thus represents a search point, which corresponds to the phenomenological representation of individual.

4.2. Reproduction

The definition of the recombination and mutation becomes the probabilistic distribution of the phenomenological measures accordingly. In the recombination, parental individuals breed offspring individuals by combining part of the information from the parental individuals, thereby creating new points inheriting some information from the old points. The recombination

Copyright © 2001 John Wiley & Sons, Ltd.

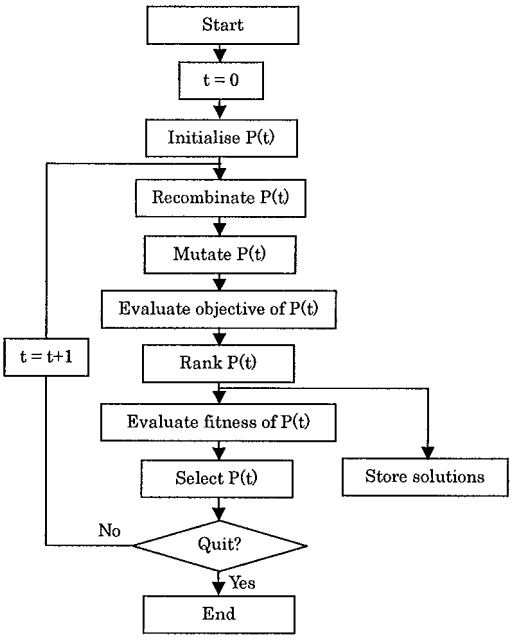

Figure 5. Fundamental MCEAs.

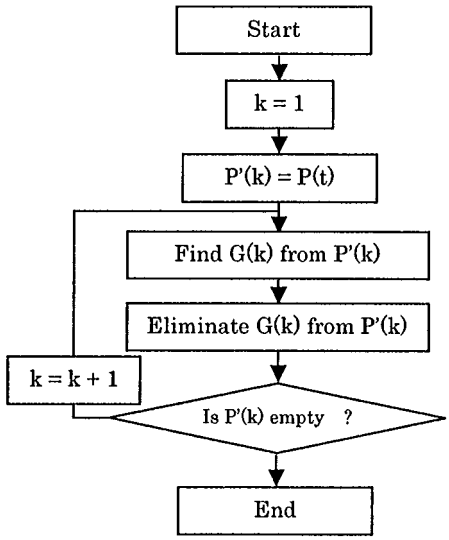

Figure 6. Ranking process of individuals.

operation is then defined as

$$\begin{aligned} \mathbf{x}'_\alpha &= (1-\mu)\mathbf{x}_\alpha + \mu\mathbf{x}_\beta \\ \mathbf{x}'_\beta &= \mu\mathbf{x}_\alpha + (1-\mu)\mathbf{x}_\beta \end{aligned} \tag{13}$$

where parameter μ may be defined by the normal distribution with mean 0 and standard deviation σ:

$$\mu = N(0, \sigma^2) \tag{14}$$

or simply a uniform distribution:

$$\mu = \mathrm{rand}(\mu_{\min}, \mu_{\max}) \tag{15}$$

The mutation can also be achieved simply by implementing

$$\mathbf{x}'' = \mathrm{rand}(\mathbf{x}_{\min}, \mathbf{x}_{\max}) \tag{16}$$

with a small possibility [25]. Note that the mutation may not be not necessary for parameter μ with normal distribution since it can allow individuals to alter largely with small possibility, when the coefficient μ is large.

4.3. Evaluation, ranking and selection

As the Pareto-optimal set satisfying Equation (11) is to be found as solutions, the ranking process of individuals is composed of an elimination rule. In the rule, the calculation of

Copyright © 2001 John Wiley & Sons, Ltd.

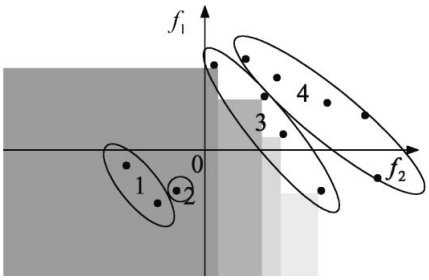

Figure 7. Ranks of individuals. Shadowed areas represent search areas of the points in group No. 3.

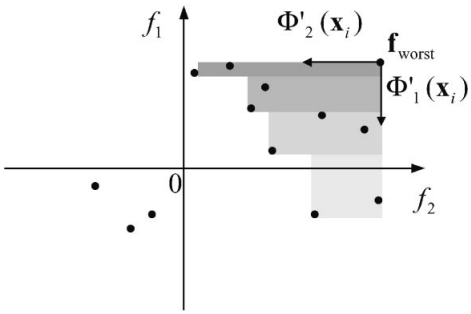

Figure 8. Evaluation of individuals. Shadowed areas each represent search areas of the points in group No. 3.

objective function at all the points $f(\mathbf{x}_i)$, $i=1,\ldots,n$, is first conducted, and the Pareto-optimal set is ranked No. 1. The points with rank No. 1 are then eliminated, and Pareto-optimal set in the population is ranked No. 2. All the subsequent ranks are generated stepwise in the same fashion until all the points are ranked [18]. The points in rank No. k, $G(k)$, are defined as

$$G(k)=\{\mathbf{x}_i \mid \text{rank}(\mathbf{x}_i)=k,\ \forall i\in\{1,\ldots,n\}\} \tag{17}$$

for further convenience, and the ranking process of the individuals is summarised in Figure 6. The ranked points are illustrated in Figure 7.

Figure 8 illustrates the evaluation of the fitness of each individual. The evaluation process starts with finding the best and worst objective function value of each point:

$$f_{\text{best}\,j}=\min\{f_j(\mathbf{x}_i)\mid \forall i\in\{1,\ldots,n\}\} \tag{18}$$

and

$$f_{\text{worst}\,j}=\max\{f_j(\mathbf{x}_i)\mid \forall i\in\{1,\ldots,n\}\} \tag{19}$$

If we temporarily define the fitness as

$$\Phi'_j(\mathbf{x}_i)=\frac{f_{\text{worst}\,j}-f_j(\mathbf{x}_i)}{f_{\text{worst}\,j}-f_{\text{best}\,j}} \tag{20}$$

we can get the normalized conditions:

$$0\leqslant\Phi'_j(\mathbf{x}_i)\leqslant 1 \tag{21}$$

and this allows us to treat the fitness of each function with the same scale. The fitness of points with the same rank has to be the same, and the true fitness of each objective function is thus defined as

$$\Phi_j(\mathbf{x}_i)\equiv\Phi_j^{G(k)}(\mathbf{x}_i)=\max\{\Phi'_j(\mathbf{x}_i)\mid \mathbf{x}_i\in G(k)\} \tag{22}$$

the fitness of each individual can be conclusively calculated as

$$\Phi(\mathbf{x}_i)=\sum_{j=1}^{m}\Phi_j(\mathbf{x}_i) \tag{23}$$

Copyright © 2001 John Wiley & Sons, Ltd.

which has the range

$$0 \leqslant \Phi(\mathbf{x}_i) \leqslant m \tag{24}$$

The selection operator favourably selects individuals of higher fitness to produce more often than those of lower fitness. As $\Phi(\mathbf{x}_i) \geqslant 0$ is satisfied by this equation, the proportional selection [18], which is reported to be faster in convergence than the other popular selection of the ranking selection [26], can be directly used in the proposed algorithm. In this selection, the reproduction probabilities of individuals are given by their relative fitness:

$$P_s(\mathbf{x}_i) = \frac{\Phi(\mathbf{x}_i)}{\sum_{j=1}^{\lambda} \Phi(\mathbf{x}_j)} \tag{25}$$

These evolutionary operations form one generation of the evolutionary process, which corresponds to one iteration in the algorithm, and the iteration is repeated until a given terminal criterion is satisfied.

4.4. Historical storage of Pareto-optimal sets

So as to grasp the configuration of the whole solution space, the resultant Pareto-optimal solutions are stored outside the loop of the evolutionary operations. The whole Pareto-optimal solutions obtained in the first generation are saved in this storage. From the second generation, the newly created Pareto-optimal solutions in the loop are compared to the stored Pareto-optimal solutions, and the new set of Pareto-optimal solutions is saved in the storage. This strategy allows the Pareto-optimal solutions created in the past to be kept as solutions and yield a good chance to increase the number of solutions, thus making the solution space easier to see. The storage of the solution independent of the current population also may contribute to the good distribution of the resultant solutions.

4.5. Comparison with other methods

The proposed method has been characterized by the multi-objective formulation, continuous evolutionary search formulation and the historical storage for the robust and efficient search of well-distributed Pareto-optimal solutions. Some conclusions in the superiority of the proposed method can be easily deduced from the past research, but numerical investigations are necessary for the others.

In the search algorithms, two major EAs, genetic algorithms (GAs) [27] and evolution strategies (ESs) [28], originally uses binary search with proportional/ranking selection and continuous search with ranking selection, respectively. In the previous reports [29, 30], those with the binary points and the ranking selection search more robustly than those with continuous points and the proportional selection at the expense of fast convergence, and vice versa. Continuous EAs (CEAs) proposed by the authors [23], incorporating continuous representation of points and proportional selection, therefore demonstrated its convergence approximately 10 times faster than that of GAs and ESs [31]. MCEAs, taking over them from CEAs should also be faster than the multi-objective versions of GAs and EAs without loss of generality, and this shall not be further mentioned in the following numerical examples.

The evaluation of fitness for multi-objective optimization is a completely new approach proposed in the paper. The ability of the approach to find appropriate Pareto-optimal solutions

Copyright © 2001 John Wiley & Sons, Ltd.

Table I. Parameters for MCEA.

Parameter	Value
No. of generations	2500
Population	10
Mutation rate	0.02

thus needs to be demonstrated and will be presented in the next section. The effectiveness of historical storage in the increase of the number of solutions is very likely but not certain so will also be investigated to make sure that the proposed method is better than all the other multi-objective optimization method in performance.

5. NUMERICAL EXAMPLES

5.1. Regularized parameter identification in two-dimensional parameter space

In order to confirm its appropriateness for finding Pareto-optimal solutions and the increase of solutions over generations, MCEA was first used to identify two parameters by minimizing a simple objective function where the exact set of solutions is known and can be seen visually in two-dimensional space. In this example, let the function be given by the simplest quadratic function:

$$f_1(\mathbf{x}) = \|\mathbf{x}\|^2 = \sum_{i=1}^{n} x_i^2 \tag{26}$$

where $n=2$, the set of parameter $\mathbf{x} \in R^2$ is subject to inequality constraint (4) with $\mathbf{x}_{\min}^{\mathrm{T}} = [-5,-5]$ and $\mathbf{x}_{\min}^{\mathrm{T}} = [5,5]$. The solution of the problem is clearly $\mathbf{x}^{*\mathrm{T}} = [0,0]$, but, to make a regularized parameter identification problem, we set that the solution is known to be adjacent to $[0.2, 0.4]$ thereby adding a Tikhonov regularization term as another objective function:

$$f_2(\mathbf{x}) = \|\mathbf{x} - \mathbf{z}\|^2 = \sum_{i=1}^{n} (x_i - z_i)^2 \tag{27}$$

where $\mathbf{z}^{\mathrm{T}} = [0.2, 0.4] \in R^2$. The problem therefore becomes to minimize functions (26) and (27). The exact Pareto-optimal solution for this problem can be determined analytically and is given by

$$X = \{\mathbf{x} \mid \mathbf{x} = r\mathbf{z},\ r \in [0,1]\} \tag{28}$$

and we can thus investigate the performance of the proposed technique with the exact solution. Values of major parameters for MCEA used to solve the problem are listed in Table I.

Figures 9(a)–(c) show the computed Pareto-optimal set in $f_1 - f_2$ space at 50th, 500th and 2500th generations, respectively, together with the exact solution. Figure 9(a) first implies that the good approximate solutions have been already obtained after 50 generations. It is then easily seen that computed solutions at larger generations are closer to the exact line, and this indicates that the proposed method converges appropriately to the exact solution. In addition,

Copyright © 2001 John Wiley & Sons, Ltd.

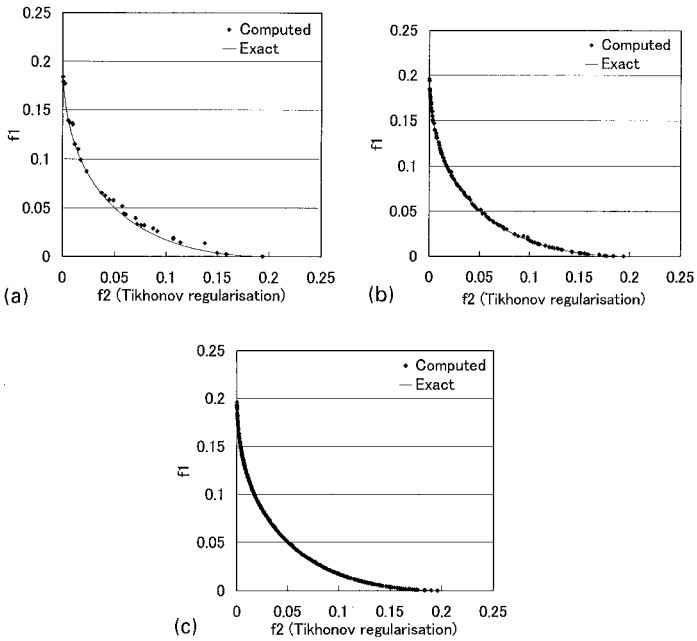

Figure 9. Pareto-optimal solutions of Example I in function space: (a) 50 generations; (b) 500 generations: (c) 2500 generations.

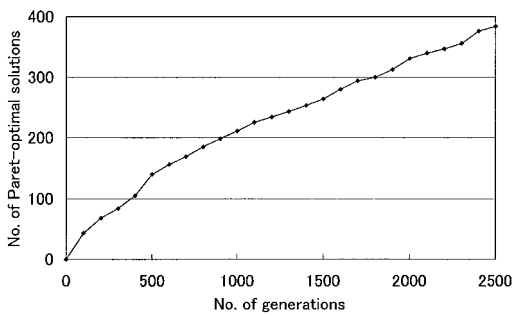

Figure 10. Pareto-optimal solutions with respect to generations in Example I: (a) 50 generations; (b) 500 generations: (c) 2500 generations.

the number of computed solutions increases with respect to the number of generations as shown in Figure 10, and this helps one to imagine the shape of the solution space.

Figures 11(a)–(c) show the corresponding Pareto-optimal solutions in x_1–x_2 space at 50, 500 and 2500 generations, respectively. The value of objective function f_1 is the distance of each computed solution from [0,0], and f_2 from $[0.2, 0.4]$. The appropriateness of Pareto-optimality of the computed solutions can be therefore confirmed visually in Figure 11(a). Moreover, one can easily see that the solutions are settling down to the exact solution with

Copyright © 2001 John Wiley & Sons, Ltd.

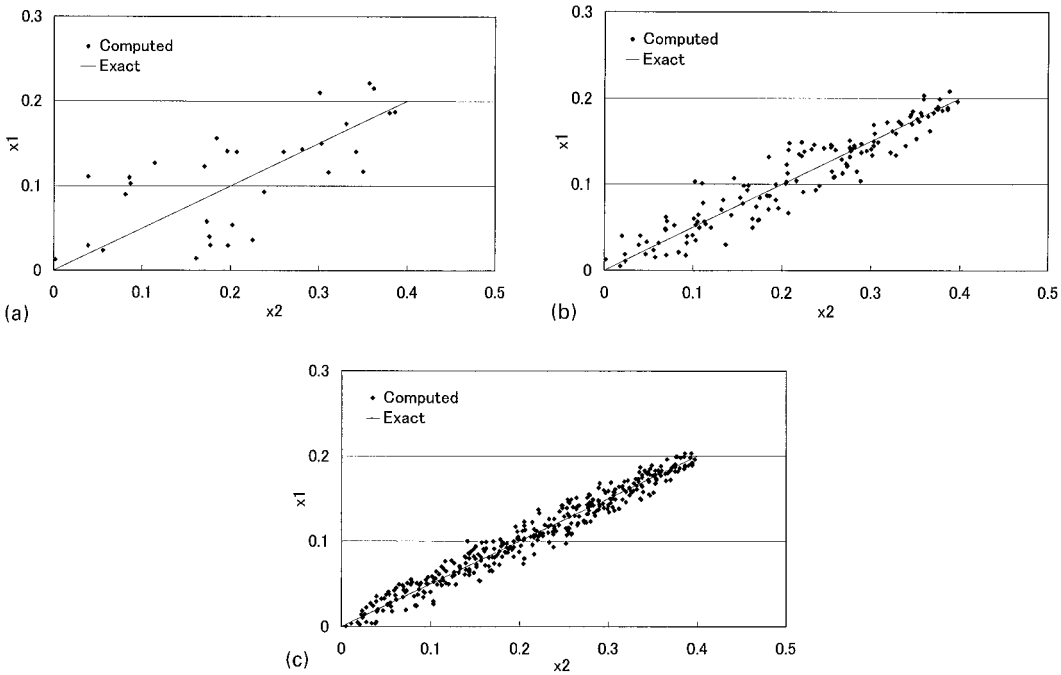

Figure 11. Pareto-optimal solutions of Example I in parameter space: (a) 50 generations; (b) 500 generations: (c) 2500 generations.

the increase of the number of generations and get very close to the exact solution at 2500 generations in Figure 11(c).

The selection of a final single solution is trivial with this example and is not further mentioned. However, we can obviously feel that the selection of a solution by seeing all solutions in the function space (Figure 9(c)) and in the parameter (Figure 11(c)) is better than the automatic derivation of a single solution with a dimensionless weighting factor.

5.2. *Regularized parameter identification within general parameter space*

To see the capability of MCEA in general multi-dimensional parameter space, the second example deals with five parameters ($n=5$) where the parameter space is constrained by inequality (4) with $\mathbf{x}_{\min}^{\mathrm{T}}=[-5,-5,-5,-5,-5]$ and $\mathbf{x}_{\max}^{\mathrm{T}}=[5,5,5,5,5]$. The objective function and Tikhonov regularization term to be minimized are given by functions (26) and (27), respectively, where $\mathbf{z}^{\mathrm{T}}=[0.3,0.4,0.5,0.6,0.7]\in R^5$, and the exact Pareto-optimal set is therefore given by Equation (28). Again, parameters listed in Table I were used for MCEA.

Figure 12 shows the Pareto-optimal set computed in function space at 50, 500 and 2500th generations. Also with this problem, the proposed method has appropriately found good approximate solutions, those at larger generations being closer to the exact curve with the increase of its number over generations as shown in Figure 13. Figures 14(a) and 14(b) show the resultant Pareto-optimal solutions at 50, 500 and 2500th generations in x_1–x_4 and x_2–x_4

Copyright © 2001 John Wiley & Sons, Ltd.

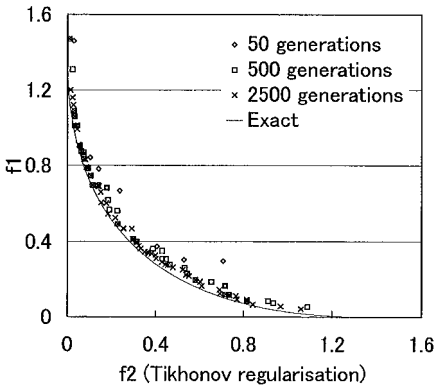

Figure 12. Pareto-optimal solutions of Example II in function space.

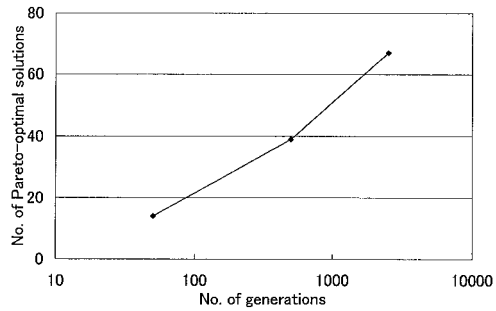

Figure 13. Pareto-optimal solutions with respect to generations.

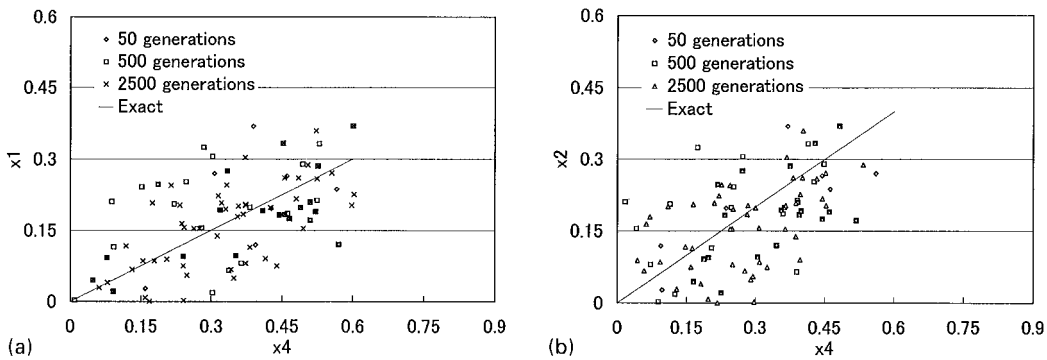

Figure 14. Pareto-optimal solutions of Example II in parameter space: (a) x_1–x_4 space; (b) x_2–x_4 space.

space respectively. It can be seen that the solutions are not as close as those in the last subsection, but they are clearly coming closer to the exact line.

In order to investigate the efficiency of the proposed multi-objective formulation compared to others, only objective function (26) was minimized with a single-objective optimization method. MCEA can be used as a single-objective optimiser simply by implementing only one function, so that MCEA was used for this optimisation. All the algorithms at the programming level are therefore the same, and the direct comparison is hence possible. Note that the use of MCEA for single-objective optimization results in the use of CEA.

Figure 15 shows the minimal value of objective function (26) of both the multi- and single-objective optimizations. The figure clearly indicates that there is only small difference in performance between both the optimizations, and this tendency did not change even when different sets of initial populations were used. This may be caused by the fact that the same individuals often occupy the population in single-objective optimization while multi-objective optimization keeps variety over generations. In addition to finding the best value of objective function (26) comparable to single-objective optimization, multi-objective optimization

Copyright © 2001 John Wiley & Sons, Ltd.

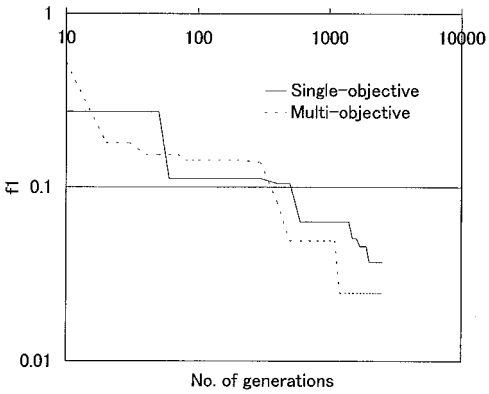

Figure 15. Best objective function values in multi- and single-optimization for Example II.

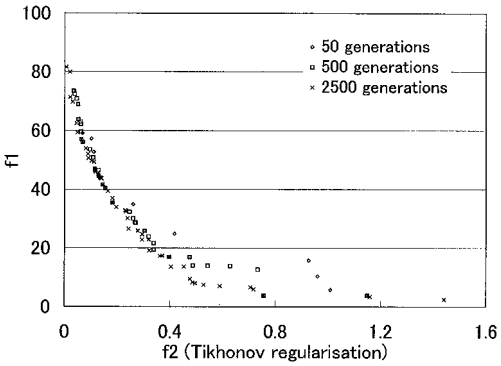

Figure 16. Pareto-optimal solutions of Example III in function space.

searched other Pareto-optimal solutions with various states of Tikhonov regularization (67 solutions at 2500th generation) while the single-objective optimizer searched for one solution, and we may conclude that the multi-objective optimization is superior to single-objective optimization.

5.3. Regularized parameter identification with a multimodal objective function

With the understanding of the appropriate performance of the proposed technique for identification with a simple objective, the identification with a complex function, which is more realistic to engineering problems, has been investigated. In this Example III, the objective function with five parameters has an additional term to Equation (26) and is given by

$$f_1(\mathbf{x}) = \|\mathbf{x}\|^2 + 50 - \sum_{i=1}^{5} 10\cos(\omega x_i) \tag{29}$$

The cosine terms clearly makes the function multimodal with a number of local minima and the function was used as a good example for a multimodal continuous function [32]. Again, Equation (27) was used as the Tikhonov regularization term, and Table I as MCEA parameters.

Figure 16 shows the resultant Pareto-optimal solutions in function space at 50th, 500th and 2500th generations. It is again seen that the distribution of the Pareto-optimal solutions is becoming smoother as the number of generations increases. The coarse distribution of solutions with small f_1 is yielded by its complexity. The Pareto-optimal solutions at 50th, 500th and 2500th generations in parameter spaces x_1–x_4 and x_2–x_4 are shown in Figures 17(a) and 17(b). In parameter space x_1–x_4, three groups of solution are seen at 50th generation. After the number of groups is reduced to two at 500th generation, a new solution is found near $[0, 0]$ at 2500th generation. Meanwhile, three groups of solution are seen at 50th generation in parameter space x_2–x_4 where one group consists of only one solution, and they result in two of them at 2500th generation.

Copyright © 2001 John Wiley & Sons, Ltd.

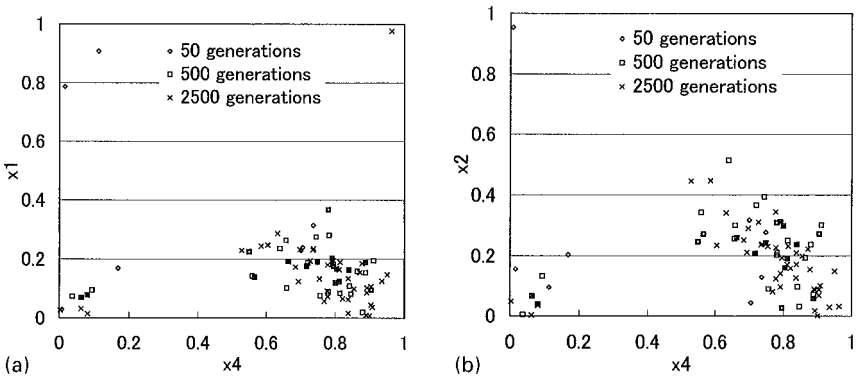

Figure 17. Pareto-optimal solutions of Example III in parameter space: (a) x_1–x_4 space; (b) x_2–x_4 space.

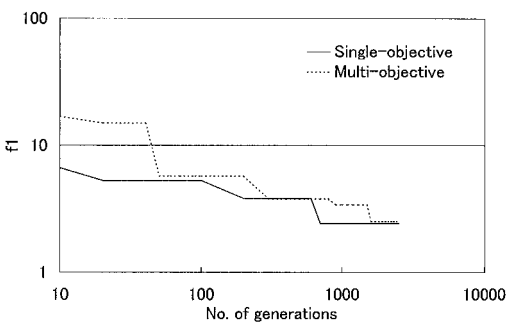

Figure 18. Best objective function values in multi- and single-optimization for Example III.

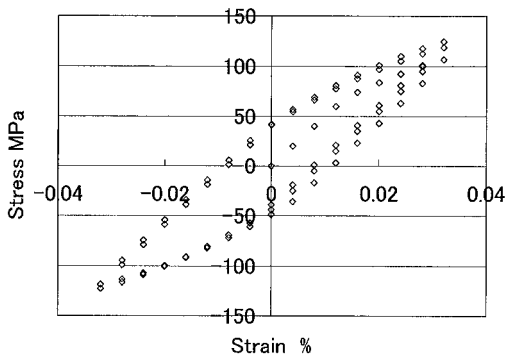

Figure 19. Pseudo-experimental data created with Chaboche model.

While MCEA found 53 solutions, the searching capability of MCEA only for f_1 was also compared to that of CEA for single optimization in the same manner, and the result of the comparison is shown in Figure 18. There is again little difference in performance between both the optimizations.

5.4. *Regularized parameter identification of viscoplastic material models*

Finally, the proposed technique was applied to a practical parameter identification problem of material models. In the mechanical tests of material, stress–strain data can be derived as experimental data. Let stress and strain be represented by σ and ε, the problem in the robust least squares formulation is given by

$$f_1(\mathbf{x}) = \sum_i \|\hat{\sigma}(\varepsilon_i^*; \mathbf{a}) - \sigma_i^*\|^2 \tag{30}$$

Copyright © 2001 John Wiley & Sons, Ltd.

Table II. Parameters for Chaboche model.

Parameters	R_0	K	n	H	D	h	d
Exact values	50	50	3	5000	300	100	0.6
Known values	40	50	2.5	4500	300	100	0.6

where $[\varepsilon_i^*, \sigma_i^*]$ are a set of experimental stress–strain data and $\sigma = \hat{\sigma}(\varepsilon; \mathbf{a})$ is a material model having $\mathbf{a}$ as material parameters. Measurement errors in the mechanical tests are relatively small, but the difficulty of solving this problem is created by the complex description of material model.

The material model used in the numerical example is Chaboche model, which is popularly used because of its capability for accurate description of the major material behaviours of viscosity and cyclic plasticity. The model under stationary temperature and uniaxial load conditions is of the form

$$\dot{\varepsilon}^{\mathrm{vp}} = \left\langle \frac{|\sigma - \chi| - R}{K} \right\rangle^n \mathrm{sgn}(\sigma - \chi) \tag{31a}$$

$$\dot{\chi} = H\dot{\varepsilon}^{\mathrm{vp}} - D\chi|\dot{\varepsilon}^{\mathrm{vp}}| \tag{31b}$$

$$\dot{R} = h|\dot{\varepsilon}^{\mathrm{vp}}| - \mathrm{d}R|\dot{\varepsilon}^{\mathrm{vp}}| \tag{31c}$$

where state variables $[\varepsilon^{\mathrm{vp}}, \chi, R]$ are the viscoplastic strain, kinematic hardening and isotropic hardening, $[K, n, H, D, h, d]$ are inelastic material parameters, and $\langle . \rangle$ is McCauley bracket. The stress–strain relationship cannot be explicitly written as described in Equation (30), but, given strain ε as a control input and the initial condition of state variables $[\varepsilon^{\mathrm{vp}}|_{t=0}, \chi|_{t=0}, R|_{t=0}] = [\varepsilon_0^{\mathrm{vp}}, \chi_0, R_0]$, the viscoplastic stress with respect to time can be derived iteratively, and stress can be ultimately calculated using

$$\sigma = E(\varepsilon - \varepsilon^{\mathrm{vp}}) \tag{32}$$

where E is the elastic modulus. In the model, parameters often unknown are inelastic material parameters $[K, n, H, D, h, d]$ plus the initial condition of isotropic hardening variable R_0, the parameters to be identified resultantly becoming $\mathbf{x} = [R_0, K, n, H, D, h, d]$.

To facilitate the analysis of identification, the numerical example uses only pseudo-experi mental data of cyclic plasticity shown in Figure 19, created from Chaboche model with a set of parameters described in Table II as 'exact values', so we will identify only parameters which influence cyclic plasticity; i.e., $\mathbf{x} = [R_0, K, n]$, setting the assumption that the others are exactly known. Material models are modelled by considering what kind of meaning the parameters included have, and therefore we often know roughly what values the parameters should take. We take values also in the table as 'known values' for parameters $[R_0, K, n]$. This gives Tikhonov regularization term expressed by Equation (27) as the second objective function f_2. The parameters used for optimisation were again those in Table I.

Figure 20 depicts the Pareto-optimal solutions in function space obtained from the parameter identification after 2500 generations. A total of 487 solutions are obtained, and it is easily

Copyright © 2001 John Wiley & Sons, Ltd.

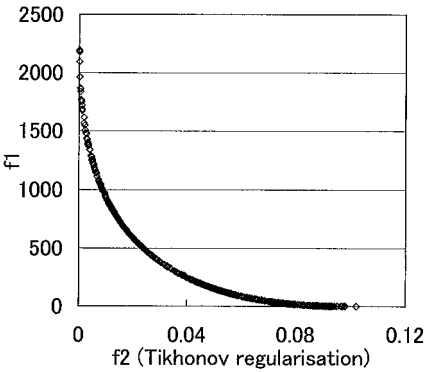

Figure 20. Best objective function values for parameter identification.

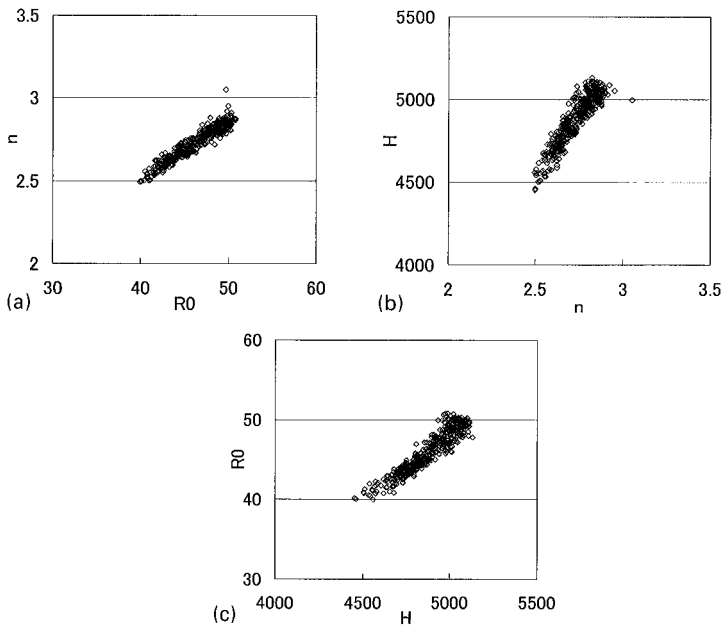

Figure 21. Pareto-optimal solutions of Example III in parameter space: (a) R_0–n space; (b) n–H space; (c) H–R_0 space.

seen in the figure that the solutions are well distributed. Respectively shown in Figures 21(a), 21(b) and 21(c) are the solutions in $R_0 - n$, $n - H$ and $H - R_0$ parameter spaces. All the figures show that the solutions are distributed along the straight line linking the 'exact values' and 'known values'. This result indicates that the proposed technique could find appropriate Pareto-optimal solutions for this problem.

Copyright © 2001 John Wiley & Sons, Ltd.

6. CONCLUSIONS

A weightless regularized identification technique and, further a multi-objective optimization method of MCEA, which can search solutions efficiently for this class of problems, have been proposed. The use of multi-objective method allows the whole solution set of the problem rather than a single solution to be derived by one optimization. The user can therefore select a single solution later by investigating the whole solution set in function and parameter spaces.

After the Pareto-optimality of solutions derived by MCEA was confirmed with a simple example visually, the proposed technique was applied to identification problems including material parameter identification, and the technique could find appropriate solutions in all the problems. The searching capability of the technique was also compared to that of a single-objective optimization method, and its superiority has been demonstrated. Conclusively, the overall effectiveness of the proposed technique for parameter identification has been confirmed.

Importance for selecting a single solution from Pareto-optimal solutions has been discussed in the paper but not its process. Developing a process is another big step, and the author is currently working on it as the advantage of the multi-objective approach has been confirmed in this paper. Several projects are in fact going on to find such a process [33].

Another important further study is the application of the technique to actual engineering problems. The author is also currently implementing the technique to the parameter identification of inelastic material models [31, 34]. The models contain 5–30 parameters, and its determination is above the human ability. The result of the identification will be reported in further papers.

REFERENCES

1. Kubo S. Inverse Problems. Baifu-kan, 1993 (in Japanese).
2. Kunzi HP, Tzschach HG, Zehnder CA. *Numerical Methods of Mathematical Optimization*. Academic Press: New York, 1971.
3. Press WH, Flannery BP, Teukolsky SA, Vetterling WT. *Numerical Recipes in C*. Cambridge University Press: Cambridge, 1988.
4. Bard Y. *Nonlinear Parameter Estimation*. Academic Press: New York, 1974.
5. Baumeister J. *Stable Solution of Inverse Problems.* Vieweg: Braunschweig, 1987.
6. Mahnken R, Stein E. Gradient-based methods for parameter identification of viscoplastic materials. In *Inverse Problems in Engineering Mechanics*, Bui T, Tanaka M (eds), 1994; 137–144.
7. Honjo Y, Kudo N. Matching objective and subjective information in geotechnical inverse analysis based on entropy minimization. In *Inverse Problems in Engineering Mechanics*, Tanaka M, Dulikravich GS (eds), Elsevier Science: Amsterdam, 1998; 263–271.
8. Morozov VA. *Methods for Solving Incorrectly Posed Problems*. Springer: New York, 1984.
9. Groetsche CW. *The Theory of Tikhonov Regularization for Fredholm Integral Equation of the First Kind*. Pitman: Boston, 1984.
10. Kitagawa T. A deterministic approach to the optimal regularization—the finite dimensional case. *Japan Journal of Applied Mathematics* 1987; **4**:371–379.
11. Kitagawa T. A numerical method to estimate the optimal regularization parameter. *Journal of Information Processing* 1988; **11**:263–270.
12. Reginska T. A regularization parameters in discrete ill-posed problems. *SIAM Journal of Scientific Computation* 1996; **17**:740–749.
13. Hansen PC. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review* 1992; **34**(4): 561–580.
14. Kubo S, Takahashi T, Ohji K. Characterization of the Tikhonov regularization for numerical analysis of inverse boundary value problems by using the singular value decomposition. In *Inverse Problems in Engineering Mechanics*, Tanaka M, Dulikravich GS (eds), 1998; 337–344.
15. Zhang X, Zhu J. An efficient numerical algorithm with adaptive regularization for parameter estimations. In *Inverse Problems in Engineering Mechanics*, Tanaka M, Dulikravich GS (eds). Elsevier Science: Amsterdam, 1998; 299–308.

Copyright © 2001 John Wiley & Sons, Ltd.

16. Kitagawa T. Methods in estimating the optimal regularization parameters. In *Inverse Problems in Mathematical Engineering*, Yamaguti M *et al.* (eds). Springer: Berlin, 1991; 37–42.
17. Kitagawa T. A comparison between two classes of the method for the optimal regularization. In *Inverse Problems*, Kubo S (ed.). Atlanta Technology Publications, Atlanta, 1992; 25–35.
18. Goldberg D. *Genetic Algorithms in Search*, *Optimization and Machine Learning*. Addison-Wesley: Reading, MA, 1989.
19. Fonseca CM, Fleming PJ. Genetic algorithms for multi-objective optimisation: formulation, discussion and generalisation. In *Proceedings of the Fifth International Conference on Genetic Algorithms*, Forrest S (ed.). Morgan Kaufmann: San Mateo, CA, 1993; 416–423.
20. Fonseca CM, Fleming PJ. An overview of evolutionary algorithms in multiobjective optimization. *International Journal of Evolutionary Computation* 1995; **3**(1):1–16.
21. Coello CA. A comprehensive survey of evolutionary-based multi-objective optimization techniques. *International Journal of Knowledge and Information Systems* 1999; **1**(3):269–308.
22. Baek T, Schwefel H-P. An overview of evolutionary algorithms for parameter optimization. *International Journal of Evolutionary Computation* 1993; **1**(1):1–23.
23. Tikhonov AN, Arsenin VY. *Solutions to Ill-posed Problems*. Wiley: New York, 1977.
24. Furukawa T, Dissanayake G. Genetic algorithms using real-valued strings. *Proceedings of the 71st JSME Annual Meeting*, vol. 930-71; 1993; 509–510.
25. Kowaltczyk T, Furukawa T, Yoshimura S, Yagawa G. An extensible evolutionary algorithm approach for inverse problems. In *Inverse Problems in Engineering Mechanics*, Tanaka M, Dulikravich GS (eds.). Elsevier Science: Amsterdam, 1998; 541–550.
26. Baker JE. Adaptive selection methods for genetic algorithms. In *Proceedings of the First International Conference on Genetic Algorithms and Their Applications*, Grefenstette JJ (ed.), 1985; 101–111.
27. Holland JH. *Adaptation in Natural and Artificial Systems*. The University of Michigan Press: Michigan, 1975.
28. Schwefel H-P. *Numerical Optimization of Computer Models*. Wiley: New York, 1981.
29. Hoffmeister F, Back T. Genetic algorithms and evolution strategies: similarities and differences. *Technical Report*, University of Dortmund, Germany, Sys-1/92, 1992.
30. Furukawa T, Yagawa G. Genetic algorithms for real search space and their use for nonlinear inverse problems. *Transactions of the Japan Society of Mechanical Engineers* 1995; **61**(586):1409–1415.
31. Furukawa T, Yagawa G. Inelastic constitutive parameter identification using an evolutionary algorithm with constitutive individuals. *International Journal for Numerical Methods in Engineering* 1997; **40**:1071–1090.
32. De Jong K. An analysis of the behaviour of a class of genetic adaptive systems. *Ph.D. Thesis*, University of Michigan, 1975.
33. Jeong MJ, Yoshimura S, Furukawa T, Yagawa G, Kim YJ. A data interpretation for multi-objective multi-solution problems using clustering algorithm. *Proceedings of Computational Engineering Conference* 2000; **5**:231–234.
34. Furukawa T, Yagawa G. Implicit constitutive modelling for viscoplasticity using neural networks. *International Journal for Numerical Methods in Engineering* 1998; **43**:195–219.

Copyright © 2001 John Wiley & Sons, Ltd.