# SIMULATION-BASED DESIGN OPTIMISATION: METHODOLOGY AND APPLICATIONS

Peter Stehouwer and Dick den Hertog

Centre for Quantitative Methods
P.O. Box 414, 5600 AK Eindhoven, The Netherlands

E-mail: dot@cqm.nl

## ABSTRACT

This paper describes a systematic approach for finding optimal design parameter settings in situations where expensive and time consuming computer simulations are used to evaluate product or process characteristics. Contrary to the usual optimisation approaches our approach is not iterative. The approach consists of four steps. The first step defines the design parameter space and generates a set of suitably chosen simulation runs. For that purpose we developed a new method based on Latin Hypercube Designs that can handle non-box design spaces. In the second step, the designer executes the proposed simulation runs using one or more CAE tools. The third step involves the application of response surface modelling techniques to the obtained simulation results. This step results in a compact model description for each of the simulated characteristics. In the final step, these compact models are used for integral design optimisation and robust design using non-linear programming and Monte Carlo sampling, respectively. For each of these steps we will give some mathematical backgrounds. In more detail we describe the application of this systematic approach to a multidisciplinary design optimisation problem. We end up with discussing the advantages of this approach over iterative approaches.

## INTRODUCTION

The ever-increasing pressure on the development time of new products and processes has changed the design process over the years drastically. In the past, design merely consisted of experimentation and *physical prototyping*. In the last decade, physical computer simulation models such as circuit design tools and finite element analysis models are widely used in engineering design and analysis. The current reliability and stability of these Computer Aided Engineering (CAE) tools has enabled the *virtual prototyping* of complex products and processes.

Still designers are confronted with the problem of finding settings for a, possibly large, number of design parameters that are optimal with respect to several simulated product or process characteristics; see Figure 1. These characteristics, called *response parameters*, may originate from different engineering disciplines. Since there are still many possible design parameter settings and computer simulations are often time consuming, the crucial question becomes how to find the best possible setting with a minimum number of simulations.

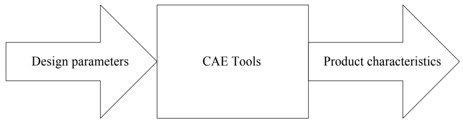


Figure 1: Design problem.

Usually in such a situation, designers use their intuition and experience. They carry out a number of simulation runs and choose the design that gives the best results. This intuitive approach can be considerably improved by using *statistical methods* and *mathematical optimisation techniques*. These techniques enable the optimisation of large, complex design problems.

Many papers describe *iterative optimisation* procedures to obtain optimal parameter settings in situations where expensive and time consuming computer simulations are used to evaluate product or process characteristics. For examples we refer to Conn and Toint (1996), den Hertog (1996), Powell (1994,1996), Schoofs (1987), Toropov (1992), and Toropov, Filatov and Polynkine (1993). Disadvantage of these approaches is that little insight is obtained in the behaviour of the responses in terms of the design parameters. Moreover, when the optimisation problem changes (e.g., a stricter bound on a response parameter), the iterative procedure has to be restarted.

This paper presents a systematic non-iterative approach in which both optimal design parameters are found and insight is given by means of so-called compact models for each response parameter. When the optimisation problem changes, new optimal settings can be calculated without carrying out new simulations.

The usage of this approach is illustrated with some examples of real-life engineering applications. In detail we describe a multidisciplinary design application from Philips.

## METHODOLOGY

The presented design optimisation approach elaborates and extends on both Response Surface Modelling (RSM) and what is called Design and Analysis of Computer Experiments (DACE); see Myers (1999) and Sachs, Welch, Mitchel, and Wynn (1989). The approach consists of four steps and is supported by the Design Optimisation Tool COMPACT, developed by CQM.

Figure 2 gives these steps. In the sequel, each of these steps will be described.

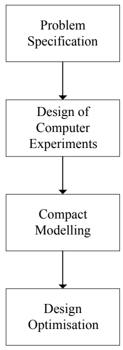


Figure 2: COMPACT approach.

Step 1: Problem Specification

In the problem specification step the design optimisation problem is formulated as to find settings for the design parameters such that the design and response parameters satisfy certain constraints and some optimality requirement is satisfied. This optimality requirement is expressed in terms of an *objective function*. Elements that need to be specified in Step 1 are:

- Design and response parameters
- Design and response parameter constraints
- Objective function

Note that the importance of choosing suitable design and response parameters should not be underestimated. An example that we frequently see in practice is the case were the maximum over a number of responses has to be minimised. In such a situation it is better to model the individual responses in stead of taking the maximum as a single response.

Step 2: Design of Computer Experiments

The second step generates a set of suitably chosen design parameter settings or *design points* that must lie within the feasible design region, i.e., the part of the design parameter space that satisfies all bounds on design parameters defined in Step 1.

The problem of choosing the initial design points is called *Design of Experiments (DOE)* (Montgomery, 1984). Classical DOE mainly focuses on physical experimentation in which experiments are subject to noise. Classical DOE schemes have the following drawbacks when used for computer experimentation

- In computer experimentation noise does not play a role, since running a computer simulation twice generally yields exactly the same results. Therefore, no information is gained from the repeated simulation of the same design point such is often done in classical DOE.
- Also due to the presence of noise, in physical experimentation it is often optimal to have design points lie on the borders of the design region. In computer experimentation other parts of the design region are often equally interesting.
- A drawback of most classical experimental design methods is that they are only applicable for rectangular design regions.

For these reasons we do not propagate the use of classical simulation schemes for computer experimentation.

In computer experimentation simulation schemes for computer experimentation must be *space filling* and *non-collapsing.* These notions will be explained next.

*Space filling schemes*– As we will explain, the goal of Step 3 is to obtain response surface models for all response parameters that predict well for the entire design region. To accomplish this, one has to choose the design points such that as much information as possible is captured from the simulation tool. Intuitively this is the case when the design points are spread throughout the design region as evenly as possible, i.e., the simulation scheme is *space filling*. Hereby we assume that no information about the function underlying the simulation model is available.

*Non-collapsing schemes* – Initially it is usually not known which design parameters are important and which are not. A simulation scheme is called non-collapsing if, in case one or more of the design parameters appear to be unimportant, every point in the scheme still gives information about the influence of the other design parameters on the response. In this way none of the time consuming computer experiments may become useless.

From our design optimisation practice we know that it often occurs that the feasible design region is non-box. For example, this may happen if points outside this region have no physical interpretation and cannot be simulated. Moreover, it is always better to use prior knowledge on uninteresting or infeasible parts of the design space when making your simulation scheme.

For these reasons simulation schemes for computer experimentation should be

- space filling,
- non-collapsing, and
- able to handle non-box design regions.

The approach we developed for generating such schemes searches for the most space filling simulation scheme within the class of so-called *Latin Hypercube Designs (LHD).* It extends the approach presented by Morris and Mitchell (1995) and will be outlined below. A detailed technical treatment will be published elsewhere.

The approach consists of two steps. The first step constructs an LHD. The second step searches for the best possible LHD with respect to space-fillingness. This combinatorial search is performed using Simulated Annealing (Aarts and Korst, 1989). Next we describe these steps.

*Construction step* – In box regions, LHDs can be constructed as follows. Let *n* be the number of design points that we are willing to simulate. Divide each design parameter dimension in *n* equidistant levels. The design points are obtained by selecting for each dimension a (random) permutation of the levels and combining these permutations into a simulation scheme.

To construct LHDs for non-box region we take more levels than points and randomly generate an LHD on the finer level grid. If the obtained LHD is infeasible, the process is repeated while increasing the number of levels. A LHD for a non-box region will be called a *constrained* LHD.

LHDs do not suffer the collapse problem: if one or more of the design parameters appear to be unimportant, every point in the design still gives information about the influence of the other design parameters on the response. In this way none of the time consuming computer experiments may become useless.

*Search step* – There is still much freedom in assigning levels to dimensions and therefore numerous feasible LHDs exist. Assigning levels can for instance be done randomly. In the search step we need to quantify the notion 'space filling'. Intuitively, a simulation scheme is space filling if the points are spread out and do not cluster in one portion of the experimental region. This does not have to mean that the experimental region is actually full. For instance, if one places a small number of points in a large number of dimensions, then space is rather empty but the points can still be spread around in a sensible way. As a measure for the space-fillingness of a simulation scheme we take the *minimal distance* between two of its design points. The larger this minimal distance the better the simulation scheme. A design for which the minimal distance is maximal is called a *maximin distance simulation scheme*.

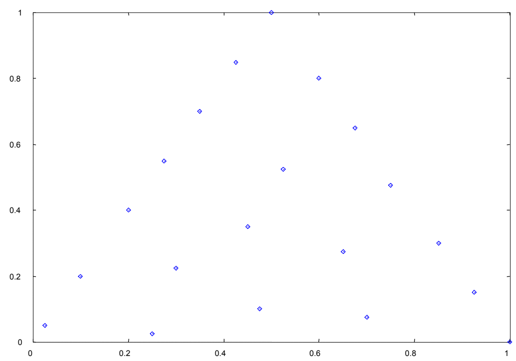


Figure 3: Computer generated space-filling LHD.

The experimental design method implemented in COMPACT can generate LHD-based maximin distance simulation schemes for linearly constrained (both rectangular and non-rectangular) design regions. The method uses simulated annealing to find an LHD with maximal minimal distance. Figure 3 gives an example of a constrained 2D simulation scheme generated by COMPACT.

Finally, in case there exist simulation experiments that are already executed, COMPACT can generate a maximin distance simulation scheme that accounts for the set of previously simulated design points.

Step 3: Compact modelling

After Step 2 the designer executes the proposed simulation runs using one or more CAE tools. The third step aims at obtaining good and compact model descriptions for each of the response parameters in terms of the design parameters. These models are based on the results of the simulations proposed in Step 2.

COMPACT builds so-called Response Surface Models (RSM) describing the relationships between design and response parameters. In the current version of COMPACT these may be either first or second order polynomial models or Kriging models (Sachs, Welch, Mitchel, and Wynn, 1989). Cross-validation techniques are used to assess how well an RSM fits the underlying relationship. To judge whether a first order, a second order, or a Kriging model is more appropriate different models can be compared.

Polynomial model terms can be pruned on the basis of the relative importance of model terms obtained by scaling all design parameters. As model selection criterion we again use cross-validation.

If, given the current simulation data, the most appropriate RSM does not fit the underlying relationship satisfactory, extra design points have to be generated and simulated. To that end there are two options:

1. Generate and simulate an additional set of design points while keeping the design region fixed.
2. Shrink the design region and generate and simulate an additional set of design points for this smaller design region. Discard simulation results of design points that lie out of the smaller region. Note that properly shrinking of the design region can be done on the basis of the trends observed in the current RSM.

Both these options concern reprocessing of the Steps 2 and 3, i.e., generation and simulation of additional design points and building a new RSM.

Regression is the mathematical technique used to fit polynomial RSMs. In the case of linear regression each response parameter is expressed as a linear function of the design parameters. Quadratic regression expresses the response parameters as quadratic functions of the design parameters. COMPACT provides both linear and quadratic regression. The regression analysis finds the particular function that fits the experiment data the most accurately (least squares method). The regression functions are used for determining the optimal design. In case of Kriging models, fitting is done using direct search methods.

*Compact model validation* – Validation methods like for instance statistical significance testing loose their interpretation in the context of computer experimentation and may therefore be misleading. Perhaps the best way to assess model prediction capabilities is it to run the simulation tool at a number of extra points and to calculate the *Root Mean Squared Error* (RMSE). However, taking into account that the simulations are typically very time-consuming, this is no serious option.

We have good experiences with the use of (*leave-one-out) cross-validation* to assess the prediction capabilities of a compact model. Given the simulated data of $n$ design points this technique loops through all $n$ design points. In every cycle of the loop the current design point is held apart and a model is constructed. The prediction capabilities of this model are then assessed on the design point held out. Using the squared errors on the points held out the *cross-validation RMSE* is calculated. Cross-validation is computationally rather demanding. However, compared with the time of a typical simulation run its computation time is negligible.

Step 4: Design optimisation

Steps 1-3 result in an RSM for each of the response parameters. In Step 4 these RSMs can be used for prediction, optimisation and detailed analysis. Note that performing an extra simulation experiment to check the final design often finishes Step 4. This section described the design optimisation functionality supported by COMPACT.

*Prediction* – New designs can be specified and infeasibilities can be analysed. Using an RSM, response values can be predicted for specific designs. COMPACT provides several graphical interfaces including 3D surface plots. Moreover, designs can be compared.

*Robust design* – COMPACT provides Monte Carlo sampling techniques to analyse the robustness of a certain design to random errors in the design parameters. For every design parameter a suitable probability distribution can be chosen. Histograms and scatter plots can be generated for individual responses and the objective function.

*Design optimisation* – Design optimisation consists of finding values for the design parameters that satisfy all constraints specified in Step 1 and minimise the objective function. COMPACT facilitates design optimisation by using powerful mathematical optimisation techniques.

Note that the design optimisation problems can have thousands of response parameters. This leads to linear (LP) or non-linear programming (NLP) problems with thousands of constraints. Hence, since the designer wants to carry out the what-if optimisation interactively, the (N)LP solvers should be fast. In COMPACT the generalised reduced gradient code CONOPT (Drud, 1992) is used for NLP problems and LAMPS (LAMPS, 1992) is used for LP problems.

*Sensitivity analysis* – It has often proven to be very useful to investigate the impact of the value of a particular lower or upper bound on the value of the objective function of the optimal design. Such *bound sensitivity analysis* involves re-calculation of the optimal design for a number of successive bound values. Bound sensitivity analysis is provided for all available types of bounds.

COMPACT also provides the calculation of the impact of the value of one of the design parameters on the value of the objective function of the optimal design. Such *design parameter sensitivity analysis* involves re-calculation of the optimal design for a number of successive design parameter values.

## CASE: A DOUBLY CURVED PANEL

In the last five years we participated in over 40 design optimisation projects in all kinds of engineering disciplines with a problem complexity ranging from 2-50 design parameters and 1-1000 responses. To illustrate the presented approach and the use of COMPACT we give a simple, but real-life, example from our design practice.

In the design of metal panels at Philips only the geometry of the panel can be varied. The panel described in this paper has a doubly curved surface and its geometry is defined by three parameters that define the curves. The panel design has to be optimised with respect to two aspects. On the one hand, the *thermal deformation* has to be minimal. Thermal deformation is defined as the deformation under local thermal load. On the other hand, the *buckling load* has to be maximal. The buckling load is defined as the minimal pressure on the convex side of the panel by which it buckles. To evaluate panel designs, Philips uses a finite element model of the panel. Thermal deformation and buckling load are calculated using a thermal analysis and a linear eigenvalue analysis, respectively. Computation time of a typical simulation run is about one hour of CPU time on a CAE workstation.

The finite element model consists of a number of plate elements with the correct thickness and material properties. The boundaries of the panel are imposed.

In the thermal analysis the panel is locally heated and the corresponding maximal panel displacement is calculated. The thermal deformation is now expressed as the *thermal load ratio*, i.e., the ratio of the predefined allowed deformation and the actual deformation. The higher this ratio, the better the panel design.

In the eigenvalue analysis a low pressure is put onto the convex side of the panel. Based on the stress and deformation in this state, the buckling load is calculated.

In this way values can be obtained for the buckling load and the thermal load ratio for each combination of the three geometry parameters. We applied the structured approach supported by COMPACT in this design optimisation project. The remainder of this section describes the four steps.

Step 1: Problem Specification

The panel design has three design parameters: C1, C2, and C3. Together these define the doubly curved panel surface. For each of these parameters we determined upper and lower bounds; see also Table 5. An extra constraint was added on height difference between the centre and the corners of the panel given by C1+C2+C3. This difference has to be between 20.0 and 32.0 [mm]. There are two response parameters: buckling load and thermal load ratio. The objective is to maximise the buckling load while keeping the thermal load ratio greater or equal than one.

Step 2: Design of Computer Experiments

Taking into account the available simulation budget, we decided to generate a constrained LHD with 13 panel designs. These panels were simulated using the finite element model. The highest value of the buckling load with a thermal load ratio larger than one equals 205.2 [N\m2]. Since the simulation scheme was space filling, the designer has already reached a reasonable degree of exploration at this point. However, this can be improved upon…

| | design parameter | | |
|---|---|---|---|
| | C1 | C2 | C3 |
| 0001 | 15.11 | 8.60 | 5.22 |
| 0002 | 12.26 | 11.17 | 1.79 |
| 0003 | 13.40 | 5.60 | 2.65 |
| 0004 | 10.54 | 6.03 | 6.93 |
| 0005 | 11.69 | 9.03 | 3.93 |
| 0006 | 16.83 | 5.17 | 4.79 |
| 0007 | 9.40 | 7.31 | 5.65 |
| 0008 | 11.11 | 8.17 | 3.08 |
| 0009 | 17.40 | 7.74 | 2.22 |
| 0010 | 12.83 | 6.46 | 7.79 |
| 0011 | 15.69 | 9.46 | 6.50 |
| 0012 | 9.97 | 9.88 | 4.36 |
| 0013 | 13.97 | 6.88 | 3.50 |

Table 1: Constrained LHD.

| | response parameter | |
|---|---|---|
| | Buckling load | Thermal load ratio |
| 0001 | 177.00 | 1.22 |
| 0002 | 176.90 | 1.04 |
| 0003 | 98.20 | 0.68 |
| 0004 | 95.87 | 0.86 |
| 0005 | 144.30 | 0.99 |
| 0006 | 117.40 | 0.93 |
| 0007 | 106.50 | 0.83 |
| 0008 | 126.90 | 0.83 |
| 0009 | 165.90 | 1.03 |
| 0010 | 121.40 | 1.02 |
| 0011 | 205.20 | 1.39 |
| 0012 | 150.40 | 1.01 |
| 0013 | 130.50 | 0.88 |

Tabel 2: Simulated panel data.

Step 3: Compact modelling

Based on the simulation results shown in Table 2 we fitted a compact model as follows. We started by fitting a fully quadratic model with all interactions on the data set. After that, we used cross-validation to assess the compact model prediction capability. We quantified the relative importance of each of the model terms by scaling all design parameters on [0,1]. After deletion of the most unimportant term we again determined the prediction capability using leave-one-out cross-validation. This process was repeated as long as the cross-validation Root Means Squared Error (RMSE) decreases.

| | coefficients |
|---|---|
| Buckling load | |
| const | -33.3085 |
| C1 | 3.4337 |
| C2 | 6.6151 |
| C3 | -3.1533 |
| Thermal load ratio | |
| const | -0.4475 |
| C1 | 0.0275 |
| C2 | 0.0539 |
| C3 | 0.1050 |

Table 3: Coefficients linear terms.

| | coefficients | |
|---|---|---|
| | C2 | C3 |
| Buckling load | | |
| C1 | 0.6016 | |
| C2 | | 0.8663 |
| C3 | | |
| Thermal load ratio | | |
| C1 | 0.0028 | |
| C2 | | |
| C3 | | -0.0046 |

Table 4: Coefficients interactions.

The cross-validation RMSE values for the compact models for buckling load and thermal load ratio are 2.26 and 0.01, respectively. For this application the accuracy is sufficient.

Step 4: Design optimisation

Based on the compact models obtained in Step 3 we used COMPACT to apply design optimisation. We maximised the buckling load such that all design constraints are satisfied and the thermal load ratio is greater than or equal to one. The maximal buckling load was 239.6 [N/m2] with a thermal load ratio of 1.49; see also Table 5. This is an improvement of almost 17 percent in buckling load compared to the best panel found in Step 2. Fortunately, there exist a panel design for which both buckling load and thermal load ratio are very good. It appears from Table 5 that the optimal panel takes all the available panel height of 32.0 [mm].

| | lower bound | optimal design | upper bound |
|---|---|---|---|
| C1 | 9.40 | 17.40 | 17.40 |
| C2 | 5.17 | 11.17 | 11.17 |
| C3 | 1.79 | 3.43 | 7.79 |

Table 5: Optimal panel design.

In this application panel height is preferably as low as possible. To investigate the trade-off between maximal allowed panel height and optimal buckling load we varied the upper bound on panel height while maximising the buckling load. The result is given in Figure 4. The upper curve gives the buckling load; the lower curve gives the corresponding thermal load ratio multiplied by 100. It appears that, when decreasing the maximal allowed panel height form 32 to 20, both buckling load and thermal load ratio halve.

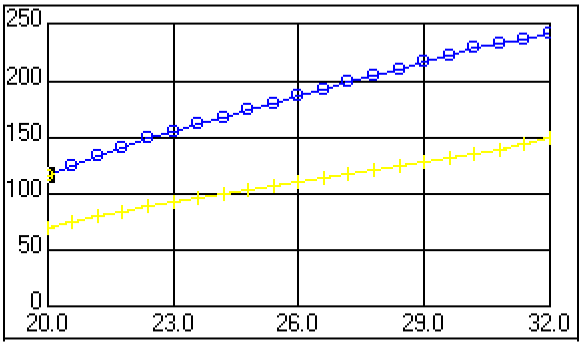


Figure 4: Sensitivity bound panel height.

## CONCLUSION

In this paper we have described a systematic approach for design optimisation. We experienced that using this approach supported with the tool COMPACT improves your designs, leads to a better understanding of the effect of design parameters on responses, and reduces the number of simulations and therefore decreases the amount of computation time on expensive CAE workstations. As a result of that, there is more time for the designer to concentrate on, for instance, 'what-if' analysis.

Moreover, the advantage of this approach over iterative approaches is that once the compact models are built, redesigns can be done quickly and effectively without the need for extra, often complex and time consuming, computer simulations. In some cases it turned out to be particularly useful to communicate the compact models to the local development centres for redesign purposes.

Finally, we experienced that by structuring the design process it becomes stable in the sense that repeating the work in roughly the same way will produce roughly the same (good) results.

## ACKNOWLEDGEMENT


The authors wish to thank Frank Lutgens of Maastricht University and Arjen Vestjens from Centre for Quantitative Methods for their major contribution to the experimental design part. Moreover, we wish to thank Vincent Bouwman, Jan-Willem ter Weeme, and Kouchi Zhang from Philips Eindhoven for their input on the physical part of the panel optimisation problem.


## REFERENCES


Conn A.R., Toint, Ph.L. (1996), "An algorithm using quadratic interpolation for unconstrained derivative free optimization", in: Nonlinear Optimization and Applications, G. di Pillo and F. Giannes (eds.), Plenum Publishing, pp. 27-47.

Johnson, M.E., Moore, L.M., and Ylvisaker, D. (1990), "Minimax and maximin distance designs", Journal of Statistical Planning and Inference, Vol. 26, pp. 131-148.

Drud, A.S. (1992), "CONOPT – a GRG code for large-scale nonlineair optimization – reference manual", ARKI Consulting and Development A/S, Bagsvaerd, Denmark, 1992.

Hertog, D., den (1996), "Sequential Regression Programming", Optix User Guide (confidential).

LAMPS User Guide, Advanced Mathematical Software Ltd., London England, 1992.

Morris, M.D. and Mitchell, T.J. (1995), "Exploratory designs for computer experiments", Journal of Statistical Planning and Inference, Vol. 43, pp. 381-402.

Montgomery D.C., (1984), "Design and Analysis of Computer Experiments (Second Edition)", John Wiley & Sons, New York.

Myers, R.H. (1999), "Response surface methodology – current status and future directions", Journal of Quality Technology, Vol. 31, pp. 30-74.

Powell, M.J.D. (1994), "A direct search optimization method that models the objective and constraint functions by linear interpolation", Advances in optimization and numerical analyses, Proceedings of the sixth workshop on optimization and numerical analyses, Oaxaca, Mexico, Vol. 275, Kluwer Academic Publishers, Dordrecht, pp. 51-67.

Powell, M.J.D. (1996), "A direct search optimization method that models the objective and constraint functions by linear interpolation", Presentation at the SIAM conference 1996, Virginia.

Sachs, J., Welch, W.J., Mitchel, T.J., and Wynn, H.P. (1989), "Design and analysis of computer experiments", Statistical Science, Vol. 4, pp. 409-435.

Schoofs, A.J.G. (1987), "Experimental design and structural optimization", Ph.D. Thesis, Technical University of Eindhoven.

Toropov, V.V. (1992), "Multipoint approximation method in optimization problems with expensive function values", in: Computational System Analysis 1992, Sydow, A. (ed.), Elsevier, pp. 207-212.

Toropov, V.V., Filatov, A.A., and Polynkine, A.A. (1993), "Multiparameter structural optimization using FEM and multipoint explicit approximations", Structural Optimization, Vol. 6, pp. 7-14.