# Simulation-based Design Optimization Methodologies applied to CFD

John Parry*, Robin Bornoff*, Peter Stehouwer**, Lonneke Driessen**, Erwin Stinstra**
*Flomerics Ltd, 81 Bridge Road, Hampton Court, Surrey KT8 9HH, UK
**CQM BV, Vonderweg 11, Building HCZ-3, PO Box 414, 5600 AK Eindhoven, The Netherlands
Email: john.parry@flomerics.co.uk, stehouwer@cqm.nl

## Abstract

Finding the optimal physical design for an electronic system is extremely time-consuming. In this paper we describe a sequential global optimization methodology that can lead to better designs in less time, and illustrate its use by optimizing the design of a heat sink for a simple system. The results show the need for a global approach, the insights that can be gained through automated design optimization, and illustrate the efficiency of the reported methodology in finding the optimum design.

## Keywords

Analytical and numerical simulations, simulation-based optimization, computational fluid dynamics, CFD, heat sink design

## Introduction

As electronic products become more sophisticated and design margins tighten, defining the thermal management strategy early in the design cycle is vital to ensure a cost-effective design for the level of heat dissipation, and high field reliability. Optimizing the cooling system for an electronic product can involve juggling many design variables, such as airflow rate, fan and vent locations, heat sink size, etc.

Numerical tools are increasingly used in the physical design of electronic products to qualify and improve the design and reduce time to market. However, despite the computing power available, exploring all possible design alternatives is extremely time-consuming.

Much research has already been carried out in the field of design optimization and numerous methods exist. Methods that explicitly take into account the high cost involved with a function evaluation can roughly be divided into two groups, sequential methods and non-sequential methods. Non-sequential methods are aimed at modeling the whole design space with help of dedicated Design of Experiments techniques and Response Surface Models. See for example Stehouwer and den Hertog [1]. Use of a sequential method is more suitable for multi-dimensional design spaces containing infeasible areas. References for some sequential optimization methods are Brekelmans et al. [2], Nocedale and Marazzi [3], Conn et al. [4], and Toropov et al. [5].

The paper discusses a simulation-based sequential global optimization methodology developed for use with a Computational Fluid Dynamics (CFD) tool, FLOTHERM®, and illustrates the application of the technology by optimizing the design of a plate fin heat sink. The main benefits of using automated design optimization are that better designs can be obtained in less time, and insights are gained into the performance of the design.

## Overview of the Optimization Problem

The general picture is given in Figure 1. On the left is a set of *Design Parameters* that have to be set optimally. Typical examples are component locations, number of fins on a heat sink, etc. On the right are the relevant quantities calculated by the analysis tool, called *Response Parameters*. Examples are junction temperatures, pressure drops, fan flow rates, etc. The design optimization problem is to find a set of values for the design parameters such that the design parameters satisfy certain constraints, the response parameters satisfy certain constraints, and some *Objective*, being a function of the response parameters, is optimized.

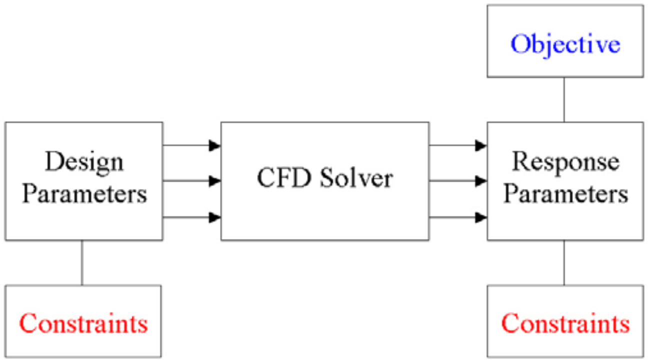


**Figure 1. The design optimization problem**

The principle characteristics that make this class of problem quite complex are:

- Function evaluations (i.e. CFD runs) are time-consuming
- Non-linear objective function and non-linear constraints
- The existence of local minima
- Numerical noise introduced by mesh changes
- Small residual errors in the solution due to the use of finite convergence criteria
- Absence of derivative (i.e. gradient) information
- Presence of integer design parameters like the number of fins on a heat sink
- Collision constraints (i.e. objects are prevented from colliding with other objects in the design space.

The optimization methodology described here deals with the complexities listed above. Compared to pre-existing sequential algorithms the novel aspects are:

- The global optimization approach
- Dealing with integer design parameters, and
- Dealing with object collision constraints.

The approach consists of two steps: an explorative search of the design space, followed by a local optimization. These steps are outlined next.

### Explorative Search (ES)

The possible existence of local minima makes it necessary to apply a global optimization strategy. First we generate an initial set of designs or *Design Points* within the feasible design space that will be run by the CFD tool.

The purpose of this set of design points is twofold:

- Most importantly, the design space is explored to locate interesting areas to be further explored in the local optimization step.
- It provides a set of base points or support vectors that are subsequently exploited during the execution of the sequential optimization algorithm.

To create a set of well chosen starting points, a Latin Hypercube Design (LHD) is generated. Stehouwer and den Hertog [6] describe the construction of such LHDs for continuous valued design parameters with constraints. These LHDs possess the following characteristics, which are important for simulation-based experiments:

- Space Filling – designs are spread throughout the design space as evenly as possible
- Non-Collapsing – each design parameter has a unique value, i.e. different for every design point
- Non-box – able to handle non-linearly constrained design spaces with infeasible regions

Figure 2 gives an example of the design points generated using a space-filling non-collapsing LHD for a non-linearly constrained two-dimensional design space.

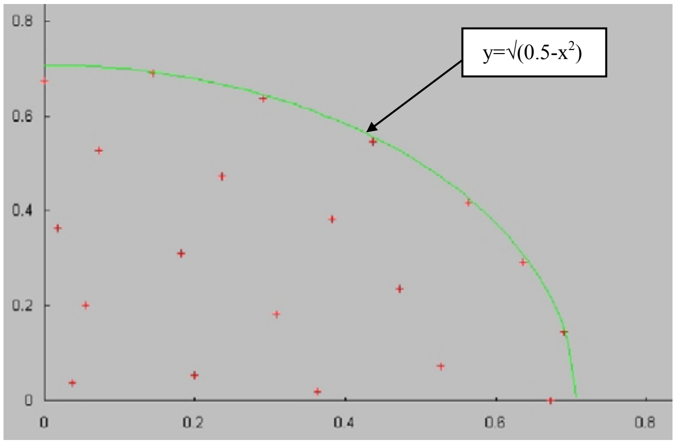


**Figure 2. A 2D space-filling LHD.**

### Local Optimization (LO)

From a given starting point a so-called local optimization approach is used to converge to the nearest local optimum. Traditional optimization methods based on derivatives fail, because derivative information is not available, and is too expensive to estimate through finite differencing. Moreover, these methods may get trapped in non-physical local minima introduced by numerical noise. For example, the change in the result obtained when an object is moved slightly may be due more to changes in the grid distribution than to any physical effect. Hence a sequential optimization approach to deal with time-consuming CFD runs without reliance on derivative information had to be developed, as briefly described below.

Iteratively, the following steps are executed:

1. Local linear approximations of the CFD model outputs are obtained with the help of weighted regression techniques.
2. The resulting approximate (or compact) models are then optimized within a *Trust Region* centered on the current best design to find the best feasible objective improving point (i.e. a step towards the optimum).
3. If the geometry of the design points that determine the local approximations are located in such a way that they result in a bad approximation of the actual model, then we evaluate a geometry improving point (to improve the accuracy of the approximate model) instead of an objective improving point.
4. At each iteration, a new local linear approximation is built, and either a new design point is evaluated (objective or geometry improving), or the trust region is decreased, depending on the progress of the optimizer.

The focus of the approach is on getting good solutions with a limited number of function evaluations. The termination criteria are, among others, the size of this trust region and a maximum number of simulation steps to be performed. For a more thorough description of a similar approach we refer to Brekelmans et al. [2].

To aid the search for the global optimum, a multi-start strategy can be used, meaning that several local optimizations are started sequentially. The starting points generated in the explorative search step are visited in order of diminishing objective value. Given the computational expense of the CFD runs, there is a trade-off between precisely converging on the local optimum found from one starting point and moving to the next starting point in the hope of finding a better optimum.

### Illustrative Example

The example considered here illustrates the use of the optimization technology with CFD to help optimize the design of a heat sink. The example is deliberately simple to aid understanding, being a simple channel. The problem definition is as follows:

A 20mm x 20mm thermally enhanced board-mounted component, powered at 10W, has to be cooled below the design limit for the junction temperature of 95ºC (100ºC less a 5ºC safety margin), based on a local ambient temperature of 45ºC. The objective of the optimization is to find the cheapest heat sink that is required to achieve this by natural convection for the given system configuration.

The absolute size constraints on the heat sink size are imposed by the system. Whereas the system design imposes no restriction on the number of fins, their thickness, the type of heat sink, nor it's fabrication method, based on a choice of the latter (e.g. extruded), sensible design parameter ranges can be set based on experience. If the optimum design is found to be at the limit of the range set, these can be considered further.

In this example, both the base thickness and the fin height are considered as design parameters. However, to ensure that the heat sink fits into the system the overall height of the heat sink is restricted by providing a constraint on these design parameters:

*Base Thickness + Fin Height ≤ 55mm*

This results in a design space that contains an infeasible region. However, this does not pose a problem for the approach global optimization approach reported here.

**Table 1. Design Parameter Ranges**

| Design Parameter | Min value | Max value |
|---|---|---|
| Overall Height (Base + Fins) | 11mm | 55mm |
| Base Width | 30mm | 60mm |
| Base Length | 30mm | 60mm |
| Number of Fins | 5 | 50 |
| Fin Thickness | 0.5mm | 2.5mm |
| Base Thickness | 1.0mm | 5.0mm |

The optimization task presented here is challenging for several reasons:

- One of the parameters (number of fins) is an integer and therefore discontinuous
- The natural convection environment provides strong interaction between the heat sink geometry and the flow (radiation is ignored in this example)
- The range over which the parameters are being varied is quite large so the design points are widely spaced

The location of the heat sink is set as a function of the base width and base length being varied by the optimizer, such that the heat sink base remains centered on the component as its size changes, i.e.:

*Base X Location = Constant - 0.5 * Base Length*
*Base Y Location = Constant - 0.5 * Base Width*

**Table 2. Response Parameter Ranges**

| Response Parameter | Min value | Max value |
|---|---|---|
| Heat Sink Mass | N/A | N/A |
| Component Temp. Rise | N/A | 50ºC |

We cannot directly minimize cost. However, for a given fabrication process and attachment method, this correlates with the mass of the heat sink, so the objective is chosen to be:

*Objective = min(Heat Sink Mass)*

The number of cells between the fins is held constant at 3, previously found to give good results for laminar flow between the fins [7].

## Results

For this example we have performed the following optimization studies:

- local optimization from the base case (LO);
- local optimization from the best design obtained from an initial space-filling set of approx. 30 experiments (ES+LO); and
- global optimization, by starting a local optimization from each design in the ES (ES+GO).

The number of designs required in the ES depends on the number of design variables and how non-linear the resulting response surface is, which is generally not known *a priori*. As a rule of thumb, the number of designs is recommended to be 5 times the number of design variables, giving of 30 runs in this case.

The design parameters for the base case, and the best feasible design from each of these three optimizations are shown in Tables 3 to 6 below, together with the resulting response parameters values.

**Table 3. Base Case Design**

| | |
|---|---|
| Overall Height (Base + Fins) | 35.0mm |
| Base Width | 60.0mm |
| Base Length | 60.0mm |
| Number of Fins | 12 |
| Fin Thickness | 0.50mm |
| Base Thickness | 5.0mm |
| Component Temp. Rise | 33.0ºC |
| Heat Sink Mass | 80.6g |

**Table 4. LO from Base Case Design**

| | |
|---|---|
| Overall Height (Base + Fins) | 25.3mm |
| Base Width | 52.7mm |
| Base Length | 52.7mm |
| Number of Fins | 10 |
| Fin Thickness | 0.50mm |
| Base Thickness | 4.3mm |
| Component Temp. Rise | 44.9ºC |
| Heat Sink Mass | 49.1g |

**Table 5. LO from Best Design in the ES**

| | |
|---|---|
| Overall Height (Base + Fins) | 35.2mm |
| Base Width | 52.7mm |
| Base Length | 20.0mm |
| Number of Fins | 10 |
| Fin Thickness | 0.50mm |
| Base Thickness | 2.0mm |
| Component Temp. Rise | 49.6ºC |
| Heat Sink Mass | 15.3g |

**Table 6. GO from All Designs in the ES**

| | |
|---|---|
| Overall Height (Base + Fins) | 42.9mm |
| Base Width | 40.7mm |
| Base Length | 26.7mm |
| Number of Fins | 7 |
| Fin Thickness | 0.54mm |
| Base Thickness | 1.0mm |
| Component Temp. Rise | 49.9ºC |
| Heat Sink Mass | 15.0g |

For each of optimization the performance of the optimizer is shown in the graphs below. These show the heat sink mass for each step taken during the local optimization (SO step) and the resulting component temperature rise. Also reported is the mass of the current best feasible design (best mass). This is the lowest mass heat sink satisfying the constraint on the component temperature rise achieved up to this point in the optimization process.

In each graph the base case result is shown for SO Step 0 to provide a feasible initial design.

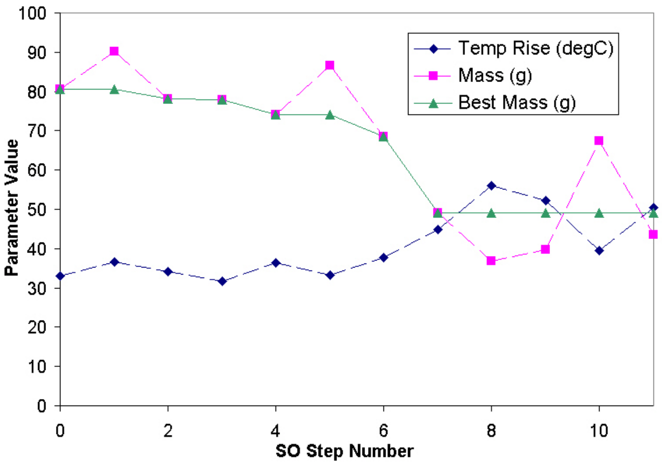


Figure 3. LO from Base Case

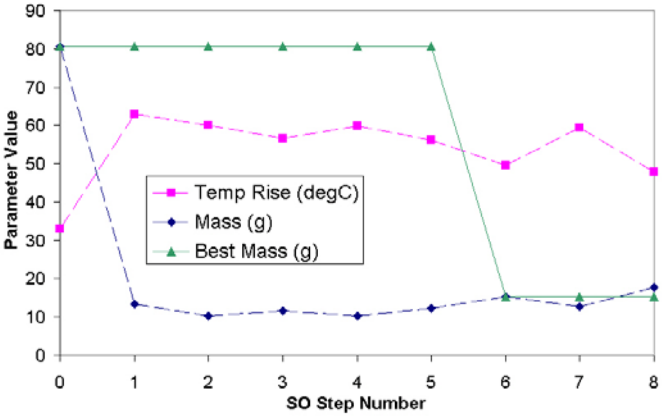


Figure 4. LO from Best Design in the ES

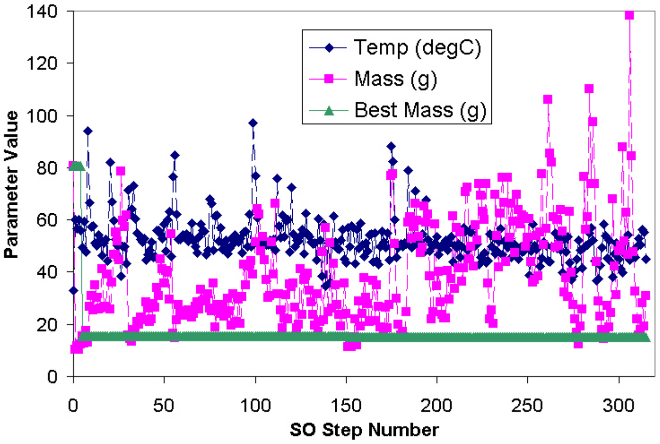


Figure 5. GO: LO from Every Design in the ES

### Discussion of Results

In Figure 3, SO Step 1 gives a higher mass heat sink. The current best design is the original, and so the 'best mass' remains unchanged. Subsequent steps find improved designs until the local optimum is found at SO Step 7. The termination criteria for the search algorithm are set to ensure that computational effort is not wasted by finding the local optimum with unnecessarily high precision, given that the global optimum may well be in some other part of the design space. However, this is not a limitation of the algorithm.

In Figure 4, SO Steps 1 to 5 all give a lower heat sink mass. However, the temperature rise for each of these steps is too high, so SO Step 0 remains the current best design until SO Step 6 where the local optimum is found. This design, given in Table 5 is a quite different design from that found when starting from the base case, given in Table 4, providing evidence for the presence of multiple optima within the design space, and hence the need to use a global optimization approach.

The first few SO Steps in Figure 5 replicate those in Figure 4, since the GO starts with a LO from the best design in the ES before moving on to the next best design, and so on. This remains the optimum design until SO Step 152 where a final small improvement in the design is achieved. The LO performed from the base case as part of the GO follows a different path from that shown in Figure 3, as shown in Figure 6. For comparison purposes, the base case results have again reported as SO Step 0.

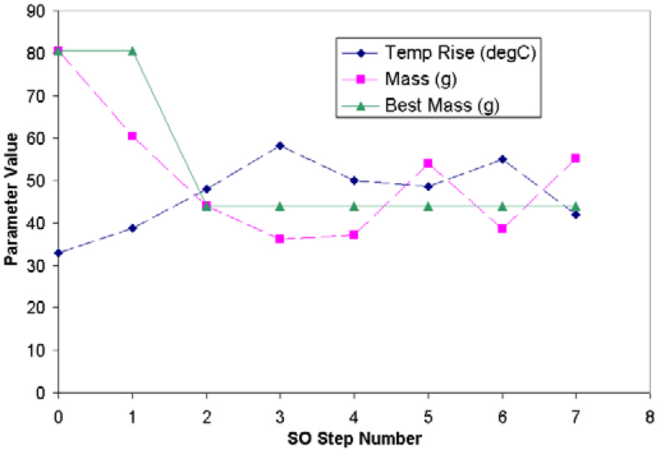
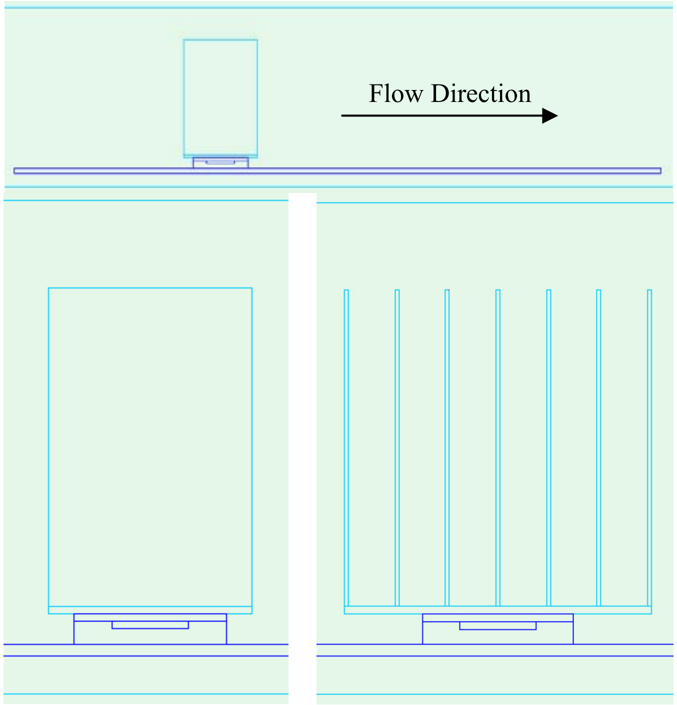


Figure 6. LO from Base Case During GO

The reason for this is that the optimizer 'learns' as is progresses. As more design points are evaluated within the design space these improve the linear approximate models used during the LO stage to guide the optimizer towards the optimum solution leading to faster convergence of the algorithm. The optimization shown in Figure 3 only has knowledge of the base case result, plus 6 other designs built around the base case to provide a minimum number of design points to define initial trust region for the number of design parameters. The optimization reported in Figure 6 has knowledge of the 30 ES designs, plus all design points solved earlier in the GO. Comparing Figures 3 and 6 shows the benefit of this additional knowledge, as the optimizer finds a better local optimum in fewer steps. This provides further evidence of local minima within the design space.

The global optimal design found is shown in Figure 7 below. Note that this is not guaranteed to be the most optimal design within the design space. This optimal design has relatively few, tall, widely spaced fins. The high aspect ratio of the fins excludes extrusion as a fabrication process. This design, and the sub-optimal designs found by local optimization from the base case and the best design in the ES all have fin thickness at or close to the minimum value of 0.5mm, i.e. on the boundary of the design space. This indicates that the heat sink fabrication method needs to be reconsidered.



Side Detail End Detail

**Figure 7. Optimal Design**

One option would be to explore designs that are fabricable by extrusion. This could be achieved by imposing a constraint on the design variables, so that the fin aspect ratio is less than (say) 10:1:

$$\textit{Fin Thickness} \geq 0.1 * \textit{Fin Height}$$

In addition, the maximum number of fins could be reduced to make the design space smaller, and therefore less costly to traverse.

Another option would be to explore alternative fabrication methods such as folded fin designs, by considering fin thicknesses say down to 0.1mm, without the above constraint and without reducing the maximum number of fins. By exploring both options, the most cost effective solution could be found.

To pursue these options a subsequent ES step should be performed on the new design space. The space-filling characteristics of the LHD technique fit new design points around those already in the feasible design space.

How rapidly the GO finds the best design is a measure of the efficiency of the approach, and is important since it may well be impractical to perform a LO from every design in the ES. In this example, there is a moderate tendency for designs with the lowest objective value (heat sink mass) to result in higher component temperature rises, as shown in Figure 8. Despite this, the GO methodology is seen to be very efficient, reducing to within a few percent of the final optimum within just 6 SO steps.

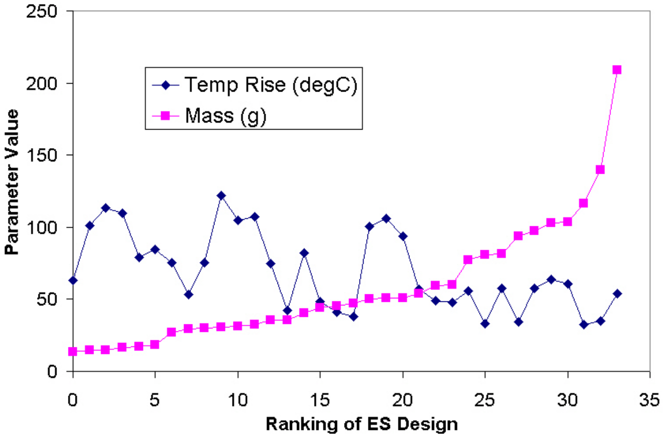


**Figure 8. ES Designs ranked by Objective Value**

## Conclusions

The use of a simulation-based design optimization methodology has been applied to the optimization of a heat sink design for a given application-specific environment, and the merits of using a global optimization strategy have been shown.

Heat sink selection is important from a business perspective, as a reduction in heat sink mass can represent a significant overall cost saving in high volume products.

Space constraints increasingly require heat sink designs to be tailored to their environment, and should therefore be optimized as part of the system design. To make the problem more tractable, localized mesh regions can be used to concentrate mesh around the heat sink. Much of the effort required to converge the solution for a new design results from the cell imbalances introduced when the mesh is changed and the dependent variables (pressure, velocities, temperature etc.) are interpolated onto the new mesh. Restricting the mesh changes to a localized region around the heat sink minimizes this.

Once the designs have been created for the ES, each design can be run independently and therefore concurrently, allowing spare capacity on the network to be utilized prior to the LO stage.

At the system-level there are many other design problems that would benefit from automated optimization, such as fan selection and fan positioning, vent sizing and vent positioning, card slot spacing within a sub-rack, etc. The optimization methodology reported here is equally applicable to all levels of packaging, from optimizing the thermal performance of a chip package for a given cooling strategy, to optimizing the ventilation system and equipment layout in a data center.

## References

1. den Hertog, D. and P. Stehouwer, “Optimizing color picture tubes by high cost nonlinear programming”, European Journal of Operational Research, Vol. 140, pp. 197-211, 2002
2. Brekelmans, R, L. Driessen, H. Hamers, and D. den Hertog, “A new sequential optimization approach to product and process design involving expensive simulations”, Proceedings of the First ASMO UK / ISSMO Conference on Engineering Design Optimization, Harrogate, UK, 2001
3. Marazzi, M. and J. Nocedal, “Wedge trust region methods for derivative free optimization”, J. Mathematical Programming Series A, Vol. 91, pp. 289-305, 2002
4. Conn, A. R., K. Scheinberg and Ph. L. Toint, “Recent Progress in Unconstrained Nonlinear Optimization Without Derivatives.”, J. Math. Programming, Vol. 79, pp. 379-414, April 1997.
5. Toropov, V. V., “Multipoint approximations based on response surface fitting: a summary of recent developments”, Proceedings of 1st ASMO UK/ISSMO Conference, pp. 371-380, Ilkley UK, 1999
6. Stehouwer, H. P. and D. den Hertog, “Simulation-based Design Optimisation: Methodology and Applications”, Proceedings of the First ASMO UK / ISSMO Conference on Engineering Design Optimization, pp. 349-355, Ilkley, UK, 1999
7. Dyson, J., “Laminar and Turbulent Flow and Heat Transfer between Parallel Surfaces”, Document Number FLOTHERM/CS1/0293 Issue 1.0, Flomerics Ltd, 1993