Global Optimization with Expensive Functions - Sample Turbomachinery Design Application

Caroline Sainvitu, Vicky Iliopoulou and Ingrid Lepot

Cenaero, Numerical Methods and Optimization Group
Bâtiment EOLE, Rue des Frères Wright, 29
B-6041 Gosselies, Belgium - caroline.sainvitu@cenaero.be

Abstract

This contribution presents some of the tools developed at Cenaero to tackle industrial multidisciplinary designs. Cenaero's in-house optimization platform, Minamo implements mono- and multi-objective variants of Evolutionary Algorithms strongly accelerated by efficient coupling with surrogate models. The performance of Minamo will be demonstrated on a turbomachinery design application.

1 Introduction

Nowadays, with the continuous increase in computing power, a widespread practice in engineering is that of simulation-based design optimization. Indeed, design of complex engineering systems, which is synonymous with the use of accurate high-fidelity simulationmodels (e.g. Computational Fluid Dynamics (CFD) analysis or Finite Element Method (FEM)), has become a reality. However, even with today's computational power, it is rarely conceivable to thoroughly search the design space using the high-fidelity simulations. Since optimization procedures are mandatory to quickly provide optimal designs, an adequate and general answer to optimization based on computationally expensive analysis lies in the exploitation of surrogate models. *Surrogate-Based Optimization (SBO)* essentially exploits surrogates or approximations instead of the expensive analysis results to contain the computational time within affordable limits (see [4, 12]), with occasional recourse to the high-fidelity model. The performance of such methods is known to be largely dependent on the following key elements: the initial sample set used to build the surrogate model(s), the underlying optimization algorithm(s), the surrogate model(s) training and the surrogate model(s) management schemes.

This paper is structured as follows. First, the SBO methodology implemented in Minamo is exposed with a focus on the design of experiments and the surrogate modeling. Subsequently, the performance of our in-house optimization platform is demonstrated on a turbomachinery design application.

M. Diehl et al. (eds.), *Recent Advances in Optimization and its Applications in Engineering*,
DOI 10.1007/978-3-642-12598-0_44,

2 Optimization Methodology

In most engineering design optimization, every evaluation of functions involved in the problem is expensive and their derivatives are, generally, unavailable or available at a prohibitive cost. Typical optimization techniques, like gradient-based methods[11], are not applicable or not efficient in such contexts. Despite their speed of convergence, these methods are indeed known to lack space exploration. They are appropriate if the derivatives are available or can be inexpensively approximated and if a good starting point is known. Moreover, they are restricted to mono-objective problems and only permit to solve multi-objective problems by using an aggregation of the objectives with pre-defined weights for each objective. Derivative-free algorithms [3] have been developed for local optimization of computationally expensive functions, but most of the time engineers are interested by a global optimum.
For these reasons, Minamo implements mono- and multi-objective Evolutionary Algorithms (EAs)sing real coded variables. These methods are stochastic, population-based search techniques and widely used as efficient global optimizers in the engineering world. Such zero-order optimization techniques are indeed robust and able to cope with noisy, discontinuous, non-differentiable, highly non-linear and uncomputable functions. Most importantly, they also permit to simultaneously handle multiple physics as well as multiple objectives. They are also less prone to getting trapped in local optima than other optimization algorithms as gradient-based methods. Moreover, EAs provide a list of optimal solutions from which the user/engineer can choose the best design according to his/her experience (see the two families of promising designs obtained in Section 3). However one drawback of EAs is that they may suffer from slow convergence due to their probabilistic nature. As a consequence, for engineering applications involving expensive high-fidelity simulations, the CPU time required for a pure EA is usually not practical. This highlights the importance to reduce the number of calls to these simulations. Therefore, the optimization process in Minamo is significantly accelerated by the use of cheap-to-evaluate surrogate models, also known as metamodels or response surface models.

2.1 Surrogate-Based Optimization

The heart of the proposed methodology consists of a surrogate modeling optimization strategy. As already underlined, SBO refers to the idea of accelerating optimization processes by exploiting surrogates for the objective and constraint functions. An SBO design cycle consists of several major elements as shown in Figure 1. It is worth underlying the major importance of the first step, namely the problem definition and optimization specification, which can include the parameterization, the definition of the bounds, the objectives and the constraints. The second step consists of building an initial database by

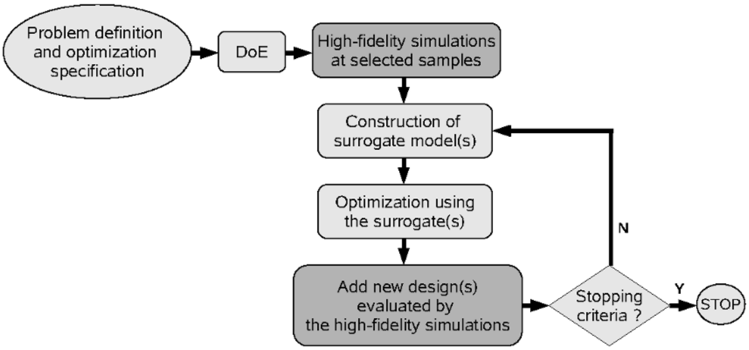

Fig. 1. Online surrogate-based optimization framework.

choosing a set of points in the design space and conducting high-fidelity simulations at these points. This exercise is called the *Design of Experiments (DoE)* Based on this DoE, surrogate models are constructed in order to build an analytical relationship between the design parameters and the expensive simulation responses. This phase provides cheap responses to be used by an optimizer. Using the surrogate models to evaluate the objective and constraint functions, an optimization is then carried out to identify the optimum, at least in the sense of the surrogates. The accurate simulation is used to evaluate the objective function and constraint values for this optimum in order to check the accuracy of the surrogates at the optimal solution. The new simulation result (and possibly simulation results at other infill points) is (are) added to the database which is therefore continuously improved with new design points, leading to increasingly accurate approximate models all along the design. This design loop is repeated until the maximum number of optimization cycles specified by the user is reached. In this contribution, an EA is employed to optimize the surrogate model(s) because this optimizer choice allows any kind of surrogate models without particular properties such as differentiability of the surrogates and also permits to deal with multiple objectives. It is important to note that our SBO scheme can incorporate the derivative information, when it is available, in different ways without any major modifications. For instance, the derivatives could be exploited directly in the construction of the metamodels. The periodic retraining of the surrogates ensures that the metamodels continue to be representative of the newly-defined search regions. Furthermore, in order to obtain a better approximate solution, a framework for managing surrogate models is used. Based on effectiveness of approximations, a *move-limit*procedure adapts the range of the variables along the design process, focusing the optimization search on smaller regions of the design space and exploiting local models. As the optimization proceeds, the idea is to enlarge or restrict the search space in order to refine the candidate optimal region. The main advantage of this is that it assures that the optimization does not generate new points in regions where the surrogates are not valid.

In order to guarantee diversity in the population, Minamo also exploits a *merit function*which is combined with the objective function of each candidate solution [15]. This function takes into account the average distance of a candidate with the other candidate solutions, and favors the solutions far away from their neighbours. A good approach for SBO seeks a balance between exploitation and exploration search, or refining the approximate model and finding the global optimum. Our strategy also allows the addition of several new design points evaluated in parallel at each cycle. Typically, the design point coming from the optimization of the surrogate(s) is added and other update points may be appended to the database as well. Using several research criteria per iteration allows to combine exploitation (optimization of the approximate function) and exploration (to systematically aim for a better global capture) within a single iteration, speeding up the restitution time of the optimization. In other words, although most of the optimizers based on the Kriging model use one single refinement criterion per iteration (such as the Expected Improvement criterion), Minamo is capable to proceed by iteratively enhancing with more than one point per iteration by using a balancing between model function minimization and uncertainty minimization. This process builds upon multiples high-fidelity simulations (e.g. CFD runs) in parallel.

The efficiency of our SBO algorithm is illustrated in the search of the global minimum of the Ackley function which is a well-known multimodal function. The left plot of Figure 2 depicts the function with 2 design parameters, while

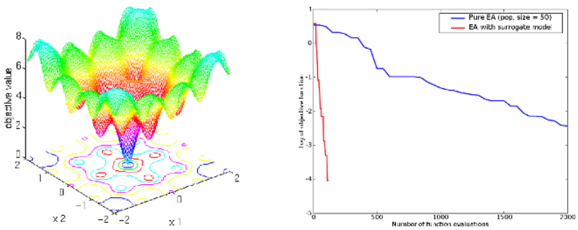

Fig. 2. Ackley function and convergence history comparison.

the optimization has been carried out on the same function but generalized to 5 dimensions within $[-2, 2]$ for every parameter. The optimization is first performed using the EA alone, with a population of 50 individuals for 40 generations (*i.e.* 2000 function evaluations). These results are compared with those obtained by the method combining the surrogate model with the EA. An initial database comprising 20 sample points is used and then only 100 design iterations are performed. The convergence histories are displayed in the right plot of Figure 2. The results clearly indicate that, for a given fixed number of actual function evaluations, the SBO approach drastically outperforms a pure EA optimization using actual function evaluations.

In Minamo, particular attention has been paid to handling simulation failures *i.e.* experiments where the simulation fails to converge. Indeed, when optimization is carried out using high-fidelity simulations, it is an inevitable fact that not all simulations provide reliable results (due to an inappropriate mesh, failed geometry regeneration, etc.). The best practice is to try to make the simulation chain as robust as possible, and let the optimizer take care of the simulation failures. In Minamo, the simulation failures are recorded for every sample point through a boolean response, called the success/failure flag. Two separate surrogate models are maintained simultaneously, namely the response model(s) (used for the evaluation of objective and constraint functions) and the failure prediction model (used for the evaluation of the simulation failure). The idea is to bias the search away from failed sample points by penalizing, via a constraint, regions containing simulation failures.

2.2 Design of Experiments

The *design of experiments* is the sampling plan in the design parameter space. This is a crucial ingredient of the SBO procedure, especially when the function evaluations are expensive, because it must concentrate as much information as possible. The qualities of surrogate models are mainly related to the good choice of the initial sample points. The challenge is in the definition of an experiment set that will maximize the ratio of the model accuracy to the number of experiments, as the latter is severely limited by the computational cost of each sample point evaluation. Minamo features various DoE techniques aiming at efficient and systematic analysis of the design space. Besides classical space-filling techniques, such as Latin Hypercube Sampling (LHS), Minamo's DoE module also offers Centroidal Voronoi Tessellations (CVT) and Latinized CVT (LCVT) [14]. A drawback of LHS is that sample points could cluster together due to the random process by which the points are generated. CVT efficiently produces a highly uniform distribution of sample points over large dimensional parameter spaces. However, a CVT dataset (in a hypercube) has the tendency for the projections of the sample points to cluster together in any coordinate axis. LCVT technique tries to achieve good dispersion in two opposite senses: LHS and CVT senses. The idea is to compute a CVT dataset and then apply a Latinization on this set of points. Latinizing a set of points means transforming it into another set of neighbouring points that fulfills the Latin hypercube property. The aim of this Latinization of CVT sample points is to improve the discrepancy of the set of points. LCVT technique has both lower discrepancy than pure CVT and higher volumetric uniformity than pure LHS (see Figure 3). The discrepancyis a measure of a point set's uniformity of projection onto all the coordinate axes. As uniformity increases, discrepancy decreases. All these space-filling techniques, independent of the design space dimensionality and of the type of surrogates, constitute good first choices to generate an *a priori* sample set in large dimensions. The DoE can be generated quickly by making use of massively parallel computers.

Since the computation of the response functions can typically take several

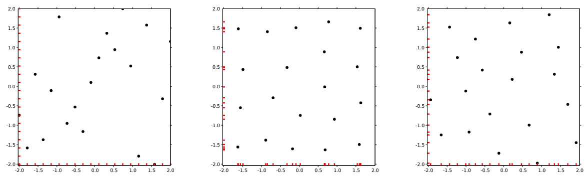

Fig. 3. LHS, CVT and LCVT, respectively, sample sets showing discrepancies of point projections (in red) onto coordinate axes.

hours on tens of computational cores, next to LCVT implementation, further research effort has been put to achieve a good accuracy of approximate models with a reasonable number of samples by incorporating function knowledge. In order to further tailor the sampling and to better capture the responses underlying physics, Minamo exhibits an *auto-adaptive DoE* technique. The idea is to locally increase the sampling intensity where it is required, depending on the response values observed at previous sample points. Such auto-adaptive techniques are also known as capture/recapture sampling or *a posteriori* sequential sampling (see [8, 9]). They incorporate information on the true function in sample distribution, explaining the term *a posteriori*. The aim is to automatically explore the design space while simultaneously fitting a metamodel, using predictive uncertainty to guide subsequent experiments. Our method consists in iteratively refining the sample dataset where the model exhibits its maximum of error, with the error indicator provided by a Leave-One-Out (LOO) procedure [10]. The use of adaptive sampling helps shorten the time required for the construction of a surrogate model of satisfactory quality. Figure 4 shows the performance of this sampling technique on a mathematical function with 2 design parameters. It allows to directly and correctly identi-

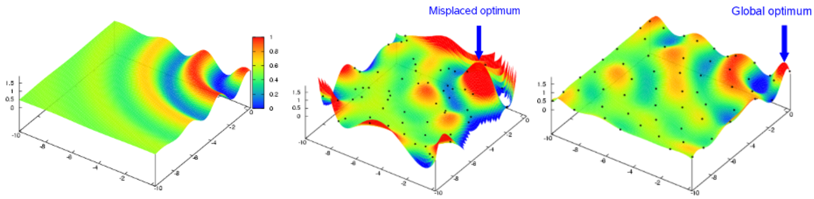

Fig. 4. The exact function, the model built using 60 LHS points and the one with 60 points generated by auto-adaptive LCVT sampling technique, respectively.

fied the region of the global optimum, whereas, using the same type of model and the same number of samples from LHS, the optimum is misplaced and the optimization will therefore be stuck in a local optimum of the original function.

2.3 Surrogate Modeling

The challenge of the surrogate modeling is similar to that of the DoE: the generation of a surrogate that is as good as possible, using as few expensive evaluations as possible. Polynomial fitting surfaces are generally not well-suited for high dimensional and highly multimodal problems. Several non-linear data-fitting modeling techniques can be used to build the surrogates, e.g. artificial neural networks, Radial Basis Functions (RBF) networks, Kriging or support vector machines [2]. Contrary to polynomial models, these techniques have the advantage of decoupling the number of free parameters with respect to the number of design parameters. Furthermore, they can describe complex and multimodal landscapes. The Minamo surrogate module offers several generic interpolators such as RBFetworks, ordinary and universal Kriging. In the training process, a trade-off must be attained between the accuracy of the surrogates and their computational cost. For our RBF network, the models are generated without the user's prescription of the type of basis function and model parameter values. Our method autonomously chooses the type of basis functions (between Gaussian or multiquadric) and adjusts the width parameter of each basis function in order to obtain an accurate surrogate model. RBF implementation is built on the efficient LOO procedure proposed by Rippa [13], while for our Kriging implementation, the parameters defining the model are estimated by solving the log-likelihood estimation problem using our EA as this problem is known to be multimodal.

3 Sample Turbomachinery Design Application

he performance of Minamo is demonstrated with the multi-point aerodynamic optimization of a non axisymmetric hub for a high pressure compressor single-row rotor blade. This work has been performed within the NEWAC project (NEW Aero engine Core concepts, project co-funded by the European Commission within the Sixth Framework Program for Research and Technological Development), aimed at technological breakthroughs for the field of aero engines efficiency and emissions. The objective was to identify the hub endwall parameter values that create a non axisymmetric hub endwall leading to significant global losses reduction with respect to the axisymmetric case at design point, while preserving the total-to-total pressure ratio close to stall.
Computer-Aided Design (CAD)systems have become an entire and critical part of the design process in many engineering fields. Therefore, it is of prime importance to exploit the native CAD system and CAD model directly within the design loop in order to avoid translation, manipulation/regeneration errors resulting from different geometry kernels. For the works presented in [6, 7], the CAPRI CAD integration middleware [5] has been exploited to provide direct CAD access without manual interventions in the CAD system during the optimization loops. Based on CAPRI, an object-oriented framework has

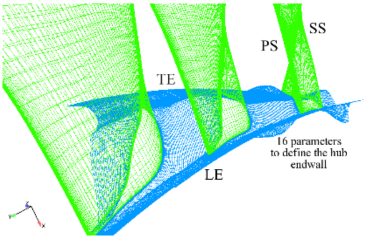

Fig. 5. Hub and blade surface mesh for the non axisymmetric hub optimizations.

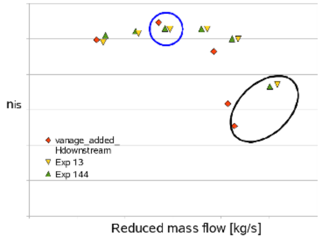

Fig. 6. Performance map with the baseline axisymmetric and optimized hub endwalls (individuals 13 and 144).

been developed for Minamo to: interact with the underlying CAD system transparently, modify the shape design variables, regenerate the CAD model and provide an updated native geometry representation to be used for the analyses.

The non axisymmetric hub surface has been parameterized under CATIA V5 and imported into the AutoGrid5 mesh generation tool for meshing purposes. The flow computations have been performed with $3D$ Reynolds-Averaged Navier-Stokes simulations using the elsA code developed at ONERA [1]. These tools have then been coupled with Minamo. Most importantly, this optimization chain can be easily applied to any blade/endwall geometry with only minor adjustments. The hub endwall has been parameterized with 16 design parameters, that can create circumferential $3D$ bumps and hollows that follow the blade curvature. The 2.2 million grid points mesh deformation at the hub is illustrated in Figure 5. Reference [6] has focused on the description of the optimization chain and methodology that have been set up, with presentation of the mono-point optimization results. Indeed, before the multi-point optimization was conducted, only one operating point was considered in order to gain first experience with limited computational cost and let the optimizer as free as possible to explore the search space. The objective was to maximize the isentropic efficiency of the compressor while imposing no additional operational or manufacturing constraints. The initial DoE was performed with LHS and held 97 sample points among which 74 experiments were considered as a success (≈ 4.5 times the number of parameters). The type of surrogate models used was RBF network. This first optimization allowed indentification of a non axisymmetric surface yielding an isentropic efficiency gain of about 0.4%. This increase may be seen as quite important, when considering that the geometry changes very locally, only at the hub endwall. However, the total-to-total pressure ratio decreased by 0.4%. This highlights one of the main drawbacks of the mono-point optimization that lead to the specification of a second robust optimization [7], now considering two operating points. The first operating point was again chosen close to peak efficiency (design point) and the second point was chosen closer to the stall region (stall point), in order to better represent the performance map of the compressor. The objective was to maximize the efficiency at the design point while preserving at least the same total-to-total pressure ratio at the stall point. The mass flow at design point was also constrained to remain within 0.5% of the reference axisymmetric flow value and some manufacturing constraints were also imposed (limited digging/removal of material). The number of success experiments for the DoE was 71 over the 97 experiments. The most interesting sample of this new DoE appeared to be the hereafter noted individual 13, which yielded an increase in terms of isentropic efficiency of about 0.39% with respect to the axisymmetric case, while it increased the total-to-total pressure ratio by 0.31% at stall without exceeding the limit on the mass flow at design point. A series of promising individuals were then found along the optimization phase in itself. Some of them were quite close in terms of performance and shape to the best DoE experiment. However, most interestingly, a second family of promising designs, quite different and somewhat smoother in terms of $3D$ surface definition, was found. This illustrates the ability of the EA to globally search the space and possibly offer a panel of solutions to the designer. Let us point out one design in this second family, individual 144, which yields an increase of efficiency of 0.35% with respect to the reference axisymmetric case, while increasing the total-to-total pressure ratio by 0.1% at stall without exceeding the limit on the mass flow at design point and satisfying the manufacturing constraints (this was not the case of individual 13). Interestingly also, individual 134 appeared quite close in shape to the interesting designs found from the mono-point optimization. The isentropic efficiency curves of the rotor with the optimized non axisymmetric hub endwalls and with the baseline axisymmetric hub are shown in Figure 6 for the two-point optimization results. The pressure contours on the blade suction side are displayed in Figure 7 and indicate that the main loss mechanism results from the shock and acceleration system along the blade suction side. The different non-axisymmetric endwall geometries de-

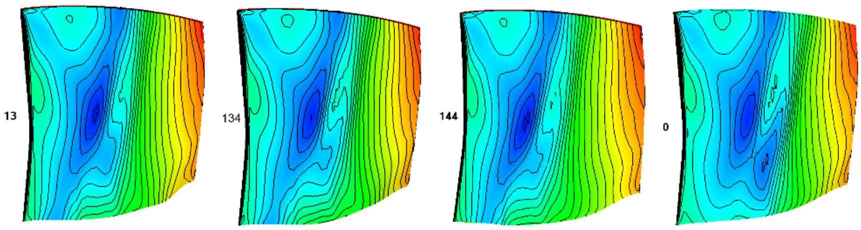

Fig. 7. Pressure contours on the blade suction side at design point for the two-point optimization - Non axisymmetric individuals 13, 134, 144 and axisymmetric case 0.

creased the losses until 50% of the blade span in the region just downstream the blade trailing edge at the hub compared to the reference axisymmetric hub geometry. The optimized designs decreased the losses downstream the shock, during the flow acceleration between 10 and 50% span.

4 Conclusion

This paper has presented our in-house optimization platform, Minamo, implementing an SBO scheme. Its capabilities have been demonstrated in a truly industrial framework with an aerodynamic design optimization. With Minamo, multi-physics multi-criteria designs tackling over a hundred parameters within a heavily constrained setting are successfully handled on a day-to-day basis.

References

1. Cambier L, Gazaix M (2002) elsA: an Efficient Object-Oriented Solution to CFD Complexity. Proceedings of the 40th AIAA Aerospace Sciences Meeting and Exhibit, Reno, USA
2. Chen V C P, Tsui K-L, Barton R R, Meckesheimer M (2006) A review on design, modeling and applications of computer experiments. IIE Transactions, Volume 38, Pages 273-291
3. Conn A R, Scheinberg K, Toint Ph L (1997) Recent progress in unconstrained nonlinear optimization without derivatives. Mathematical Programming, Volume 79, Pages 397-414
4. Forrester A I J, Keane, A J (2009) Recent advances in surrogate-based optimization. Progress in Aerospace Sciences, Volume 45, Issues 1-3, Pages 50-79
5. Haimes R, Follen G (1998) Computational Analysis Programming Interface. International Conference on Numerical Grid Generation in Computational Field Simulations, University of Greenwich, United Kingdom
6. Iliopoulou V, Lepot I, Geuzaine P (2006) Design optimization of a HP compressor blade and its hub endwall. ASME Paper GT2008-50293, Berlin
7. Iliopoulou V, Mengistu T, Lepot I, (2008) Non Axisymmetric Endwall Optimization Applied to a High Pressure Compressor Rotor Blade. AIAA-2008-5881, 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Victoria, British Columbia, Canada
8. Jin R, Chen W, Sudjianto A (2002) On Sequential Sampling for Global Metamodeling in Engineering Design. DETC-DAC34092, 2002 ASME Design Automation Conference, Montreal, Canada
9. Jones D R , Schonlau M, Welch W J (1998) Efficient Global Optimization of Expensive Black-Box Functions. Journal of Global Optimization, Volume 13, Number 4, Pages 455-492
10. Meckesheimer M, Barton R R, Simpson T W, Booker A J (2002) Computationally Inexpensive Metamodel Assessment Strategies. AIAA Journal, Volume 40, Number 10, Pages 2053-2060
11. Nocedal J, Wright S J (1999) Numerical Optimization. Springer, New York
12. Queipo N V, Haftka R T, Shyy W, Goel T, Vaidyanathan R, Tucker P K (2005) Surrogate-based analysis and optimization. Progress in Aerospace Sciences, Volume 41, Issue 1, Pages 1-28
13. Rippa S (1999) An algorithm for selecting a good value for the parameter c in radial basis function interpolation. Advances in Computational Mathematics, Volume 11, Number 2-3, Pages 193-210

14. Saka Y, Gunzburger M, Burkardt J (2007) Latinized, improved LHS, and CVT point sets in hypercubes. International Journal of Numerical Analysis and Modeling, Volume 4, Number 3-4, Pages 729-743
15. Torczon V, Trosset M W (1998) Using approximations to accelerate engineering design optimization. AIAA-98-4800 in the Proceedings of the 7th AIAA/NASA/USAF/ISSMO Symposium on Multidisciplinary Analysis and Optimization