Optimally weighted ensembles of surrogate models for sequential parameter optimization

Echtenbruck, M.M.

Author: Echtenbruck, M.M.
Title: Optimally weighted ensembles of surrogate models for sequential parameter optimization
Issue Date: 2020-07-02

Chapter 1

Introduction

A common method to perform an optimization on a function that can not be optimized analytically is to perform a search on this function by evaluating points on the function at strategically chosen positions. However, in real-world optimization tasks, the search is often constrained by time or cost for the function evaluations.

1.1 Optimization of Industrial Problems

A typical example of such optimization problems is a cyclone dust separator [1, 2]. As Slack et al. state, "the cyclone dust separator is perhaps the most widely used separation device to be found in industry. It owes its popularity to the low manufacturing and maintenance costs brought about by its simple design. There are no moving parts in the device itself, which can be constructed from a wide range of materials including refractories for high-temperature operation. Combined with moderate pressure drop and a range of throughputs and efficiencies, these advantages have made the cyclone the most attractive solution to separation in both gas-solid and liquid-solid systems" [3].

Cyclone dust separators exist in a large variety of shapes, but the most common design is the reverse-flow cyclone as depicted in Figure 1.1.
The fluid is induced into the cyclone through the inlet on the upper end of the cyclone. By the position of the inlet and the shape of the cyclone body, the fluid is forced on a circular path. The emerging centrifugal forces, caused through

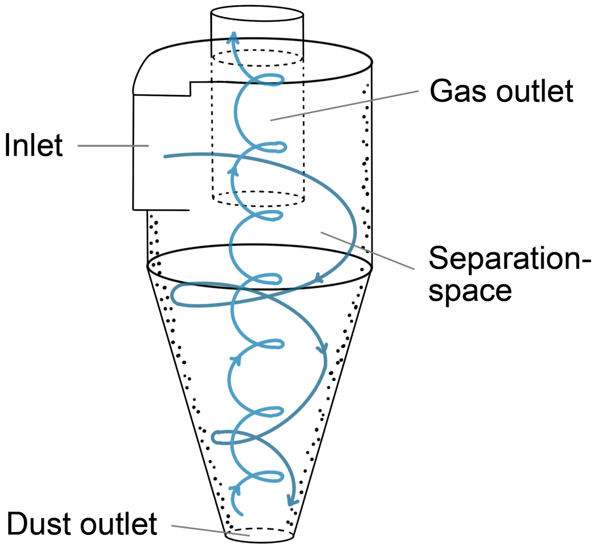

Figure 1.1: Depicted is a schematic representation of a cyclone dust separator with an illustration of the internal flow. The fluid enters the system through a lateral inlet at the upper end of the cyclone and moves then in a downward swirl along the outer areas of the cyclone. When reaching the lower end, the flow changes the direction and turns upward again, in a more narrow swirl, before it leaves the system through the gas outlet at the upper end of the cyclone. The dust particles are separated from the fluid through centrifugal forces, they are moved against the outer wall of the cyclone and then fall through the dust outlet at the lower end of the system.

the circular swirl of the fluid, separate the dust particles from the fluid and fall through the dust outlet at the lower end of the cyclone. However, the fluid, which is now separated from the dust particles, leaves the cyclone through the gas outlet on the upper end of the cyclone.

Main quality indicators for cyclone dust separators are collection efficiency and pressure drop. The collection efficiency, defined as the fraction of dust particles filtered from the fluid, reflects how well the cyclone dust separator performs its primary task. However, the pressure drop has the main impact on operational cost. This is aggravated by the fact, that these two criteria are conflicting, i.e., the settings allowing for the best collection efficiency may not correspond to the setting that enables the lowest pressure drop.

The quality indicators are heavily influenced by the geometry of the cyclone,

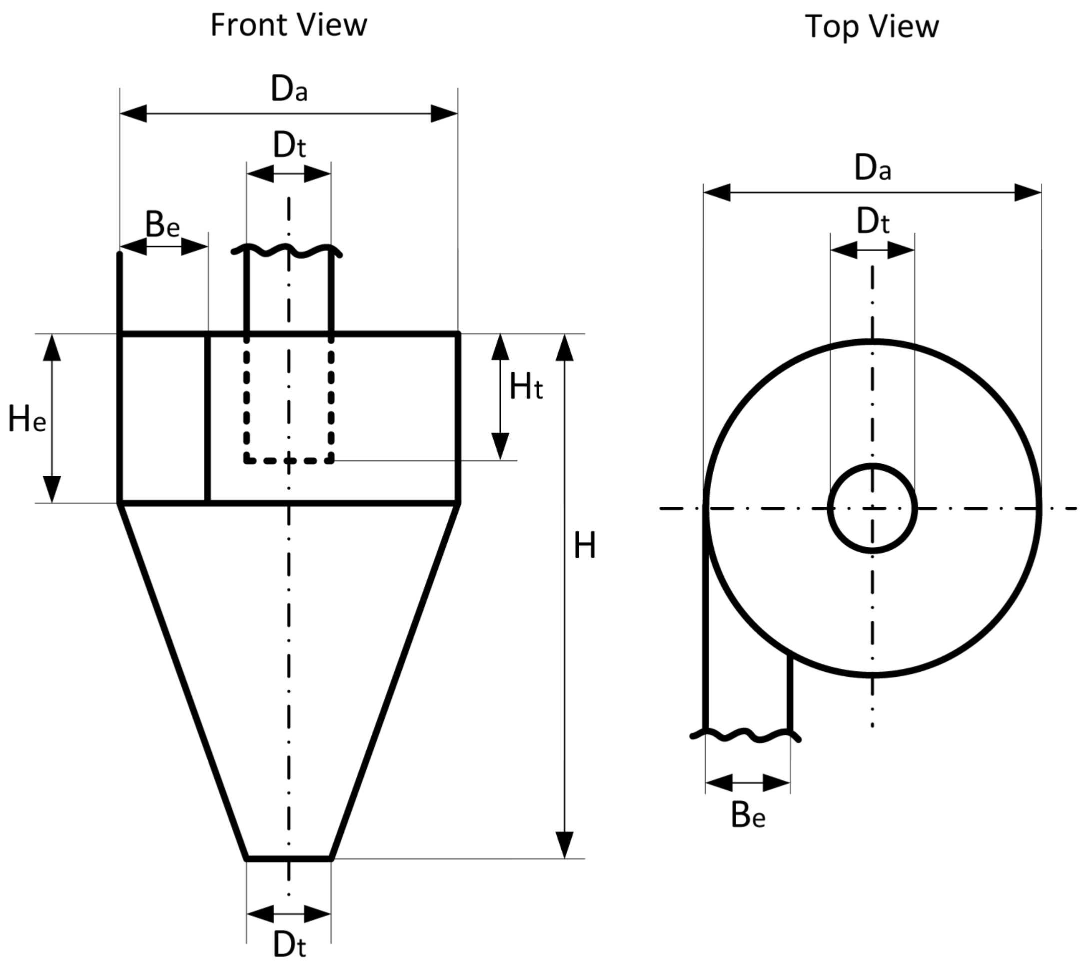

Figure 1.2: Depicted is a schematic representation of a cyclone dust separator (from [1]). Shown is the front view as well as the top view with all critical parameters indicated. These are height H and diameter D_a of the cyclone, diameter D_t and immersion H_t of the outlet pipe and width B_e and height H_e of the inlet.

which is determined by several design parameters, like the height or diameter of the cyclone. All critical parameters are depicted in Figure 1.2. The performance of a newly designed cyclone can be evaluated with Computational Fluid Dynamics (CFD) simulations. Such simulations are, depending on the required accuracy, extremely time-consuming.

All in all, this problem definition specifies an expensive, in terms of calculation time, black-box function whose function values, collection efficiency, and pressure drop, are determined by a set of parameters as depicted in Figure 1.2. In modern computer-aided design environments, simulation-based optimization is a standard tool for finding optimal parameter settings [4, 5, 6, 7]. However, the performance of such automated-design optimization tools is often challenged by the high demand in terms of computational effort. For instance, physics simulations including fluid dynamics or detailed energy flow computations in real-world structures easily entail several minutes or even hours of runtime for a single design [8]. Optimization requires the repeated execution of such time-consuming simulations, and this requires optimal use of the information gained from each simulation.

Further well-known engineering problems are, for example, the optimization of an airfoil shape for an aircraft wing [9, 10] or aerodynamic shape optimization in the automotive industry [11, 12]. For such problems, a CFD simulation has to be carried out in order to evaluate a specific shape with respect to the different design objectives and other constraints. Jameson et al. (2018) [13] state, that some of these simulations take up to 3 days, what makes a design optimization task impossible since in general lots of simulation evaluations would be needed.

1.2 Motivation and Aim

In global optimization of expensive black-box functions, it is a common technique to learn a surrogate-model, e.g., regression model, of the response function from available evaluations and to use this model to decide on the location of future evaluations. Sequential parameter optimization (SPO) is a well-known approach for solving black-box optimization problems with expensive function evaluations [14, 15]. It combines a sequential experimental design approach and tools for reducing expensive function evaluations on the original model by replacing them partly with fast approximate evaluations on surrogate models. SPO resembles Bayesian Global Optimization [16, 17], but it is less specific in the particular regression model as it does not necessarily make use of uncertainty quantification. SPO packages, such as the SPO Toolbox (SPOT), come with a large variety of surrogate models (base models) from which the user can choose.

Still, the choice of the surrogate model can have a significant influence on the solution quality and performance of the optimizer. Burnham et al. even state that the choice of the right surrogate model is the most critical question in making statistical inferences [18]. However, in order to make meaningful decisions on which surrogate model to select for a given problem, often expert knowledge is needed.

This includes knowledge about the objective function and the characteristics of the surrogate model likewise.

In many situations, preliminary knowledge about the function, or all available models, is not available. To overcome this problem, it would be beneficial if the algorithm could learn all by itself which surrogate model type suits the problem best, based on the given data. This can be done by evaluating different models on available training data and using a statistical model selection approach to select the most promising surrogate model.

But how to handle the situation when there is more than one strong model in the set? In such circumstances, it might be beneficial to combine inference output across several models. Such methods will be referred to as ensemble models. Different approaches to achieve this are known to literature. In Chapter 3 a short overview of previous work regarding ensemble models is given, and a taxonomy of ensemble models is defined. However, so far, only few work has been done on adopting ensemble methods specifically for sequential optimization, which holds its challenges for efficient ensemble modeling.

The overarching goal is to release the user from the burden to select the right surrogate model from a set of heterogeneous surrogate models. From this aim, the main research question of this thesis can be derived:

- Is it possible to create an ensemble building strategy that selects or combines heterogeneous surrogate models to achieve the best possible result and works reliably and as accurately as possible on arbitrary objective functions?

Since many real-world problems are constrained to a comparatively small number of function evaluations, the main focus is laid on regression models and SPO processes that are able to work with smaller numbers of function evaluations.

Black-box optimization refers to a problem definition where an optimum of a function is searched for that cannot be optimized analytically. Since this is, in general, the case for real-world problems, the focus is also laid on black-box optimization. In this thesis, without loss of generality, all optimization tasks are considered as minimization tasks.
Optimization processes that rely on the assistance of surrogate models are referred to as surrogate-based optimization (SBO). The process of training a surrogate model (also referred to as model) on available data is denoted as fitting the model to the data. After the model has been fitted to the data, the model provides an approximation of the underlying function, which is based on the inherent assumptions of the model and the available data. Determining an approximation

of unknown function value with the help of a model is denoted as a prediction.

1.3 Overview of this Thesis

This thesis is structured as follows:

Chapter 2 gives an introduction to all preliminary information needed for the understanding of this work. It starts by introducing the models that are used in this thesis. Then, a brief introduction to SPO as well as to the SPO framework which is used in this thesis is given. Moreover, the objective functions used for the experiments are introduced.

Chapter 3 gives an overview of previous developments in this area, different ensembles approaches and applications, and defines a taxonomy of ensembles. The advantages and disadvantages of the diverse approaches are considered with regard to the general goal of this work. The taxonomy of ensembles of surrogates that is specified in this chapter is an original work of this thesis and has not yet been published.

Chapter 4 regards the findings of Chapter 3 within the premises of the overall goal of this work and draws appropriate conclusions. A new ensemble building approach is then derived from these conclusions, implemented and thoroughly analyzed.

Chapter 5 performs the step from static modeling to SPO. The designed approach is adapted for and applied to SPO. Experiments are carried out, and results are analyzed to allow for further insights into the functioning of the method.

Chapter 6 summarizes the works of this thesis and discusses the methods presented and the results obtained. Also, possible future work and additional open questions in this research area are shown up and discussed.

1.4 Overview of Publications

Substantial parts of this thesis rely on works that have previously been published or are in the process of being published during the writing of this thesis. Some of these works are incorporated only contentwise, and others are in large parts adopted verbatim. For the sake of clarity, the way how each publication has been included in this thesis is outlined at the end of the respective chapters. Of concern are primarily the following publications ordered by the chapter of appearance.

Chapter 4
Martina Friese and Martin Zaefferer. Two challenges in surrogate-modeling: Merging surrogate-models into ensembles and dealing with structured or combinatorial search spaces. Contributed Talk at Surrogate-Assisted Multi-Criteria Optimization (SAMCO) Workshop, Lorentz Center, Leiden, NL, (2016)

Martina Friese, Thomas Bartz-Beielstein, Michael Emmerich, Building ensembles of surrogates by optimal convex combination. In *Proceedings of Bioinspired Optimization Methods and their Applications* , BIOMA 2016, Gregor Papa and Marjan Mernik (editors), Jožef Stefan Insitute, Ljubljana, Slovenia, pg.131-143 (2016)

Chapter 5
Martina Friese, Thomas Bartz-Beielstein, Thomas Bäck, Michael Emmerich, Weighted Ensembles in Model-based Global Optimization. In *AIP Conference Proceedings of LeGO 2018 - Int. Workshop on Global Optimization*, Leiden, The Netherlands, September 18-21, 2018. AIP Web of Science,pg.020003 (2019)

Martina Friese, Thomas Bartz-Beielstein, Thomas Bäck, Michael Emmerich, Optimally Weighted Ensembles of Surrogate Models for Sequential Parameter Optimization, Journal of Global Optimization, Special Issue LeGO Workshop 2019, (submitted for)

Chapter 6
Jörg Stork, Martina Friese, Martin Zaefferer, Thomas Bartz-Beielstein, Andreas Fischbach, Beate Breiderhoff, Tea Tušar, and Boris Naujoks. Open issues in surrogate-assisted optimization. In *High-Performance Simulation-Based Optimization*, Bookchapter, Thomas Bartz-Beielstein, Bogdan Filipič, Peter Korošec,

El-Ghazali Talbi (editors), pg.225-244, Springer (2020)