Structural and Multidisciplinary Optimization
https://doi.org/10.1007/s00158-018-1934-2

RESEARCH PAPER

Optimization via multimodel simulation

A new approach to optimization of cyclone separator geometries

Thomas Bartz-Beielstein[1] · **Martin Zaefferer**[1] · **Quoc Cuong Pham**[1]

Received: 31 August 2017 / Revised: 23 January 2018 / Accepted: 30 January 2018

Abstract
Increasing computational power and the availability of 3D printers provide new tools for the combination of modeling and experimentation. Several simulation tools can be run independently and in parallel, e.g., long running computational fluid dynamics simulations can be accompanied by experiments with 3D printers. Furthermore, results from analytical and data-driven models can be incorporated. However, there are fundamental differences between these modeling approaches: some models, e.g., analytical models, use domain knowledge, whereas data-driven models do not require any information about the underlying processes. At the same time, data-driven models require input and output data, but analytical models do not. The *optimization via multimodel simulation* (OMMS) approach, which is able to combine results from these different models, is introduced in this paper. We believe that OMMS improves the robustness of the optimization, accelerates the optimization-via-simulation process, and provides a unified approach. Using cyclonic dust separators as a real-world simulation problem, the feasibility of this approach is demonstrated and a proof-of-concept is presented. Cyclones are popular devices used to filter dust from the emitted flue gasses. They are applied as pre-filters in many industrial processes including energy production and grain processing facilities. Pros and cons of this multimodel optimization approach are discussed and experiences from experiments are presented.

Keywords Combined simulation · Multimodeling · Simulation-based optimization · Metamodel · Multi-fidelity optimization · Stacking · Response surface methodology · 3D printing · Computational fluid dynamics

1 Introduction

In many real-world optimization applications, models support the optimization algorithm by giving valuable information with respect to the optimized system. Modeling allows the estimation of system performance under new conditions as well as the comparison of different operating conditions and parameterizations, e.g., new geometries. This article shows how different model types, namely analytical, surrogate, *computational fluid dynamics* (CFD), and 3D printing models can be employed in an optimization process. Because every modeling approach has its advantages and disadvantages, a combination, which uses information from several models at the same time, can be beneficial.

Hence, the first contribution of this paper is to outline how heterogenous models and data sources can be integrated into a streamlined optimization process. The term "heterogenous" refers to a wide variety of distinguishing features of the models, such as run time, complexity, theoretical basis, computation versus laboratory experiments, and many more. The second contribution is the demonstration of how the proposed approach is applicable to the cyclone optimization problem. Finding an optimal geometry for cyclone dust collectors is a well studied, non trivial optimization problem of high practical relevance. We describe how this single application gives rise to diverse types of models and show how the derived information may be incorporated into the optimization algorithm in a meaningful way. The third contribution is a detailed discussion of our findings, which is of great importance for practitioners.

Loosely speaking, mathematical modeling is "the link between mathematics and the rest of the world" (Meerschaert 2013). Mathematical modeling can be performed using analytical and numerical models: *Analytical models*

✉ Thomas Bartz-Beielstein
thomas.bartz-beielstein@th-koeln.de

[1] Technische Hochschule Köln, Steinmüllerallee 1, 51643 Gummersbach, Germany

are mathematical models that have a closed form solution, i.e., the solution to the equations used to describe changes in a system can be expressed as a mathematical analytic function. Nelson (1995) refers to analytical models as "rough-cut models", i.e., mathematically solvable and typically less detailed models. *Numerical (simulation) models* are mathematical models that use some sort of numerical time-stepping methods such as Newton's method to simulate the model's behavior over time. In contrast to analytical models, solutions of simulation models are usually presented as tables or plots. Simulation is a widely used method for studying complex real-world systems, because many systems cannot be completely described by analytical models. Also, experimentation with the real system is often infeasible or expensive (Law 2007).

Nowadays, CFD simulation is a well established technique. It is also used in many studies, which describe the topic discussed in this publication: the optimization of cyclone separator geometries (Hoffmann and Stein 2007; Elsayed and Lacor 2010).

In contrast to simulations, models may also also be of a physical nature, that is, in the sense of real-world or laboratory experiments. *3D-printing* is a popular modeling technique of this type. It is commonly used to validate the results, e.g., a certain geometry, from CFD simulations. Recently, 3D-printing was integrated into the optimization via simulation loop (Preen and Bull 2014).

Another class of models that gained importance over the last decades are *surrogate models*, also known as *metamodels* (Jin et al. 2001; Bartz-Beielstein and Zaefferer 2017). They are built from and then used instead of the underlying real processes or simulation models. Popular metamodeling techniques include regression, radial basis functions, and Kriging (Santner et al. 2003; Kleijnen 2008).

Although the model based approach can be considered a success story, it also causes some problems. Several critical issues in simulation studies are related to errors (Nelson 1995). These errors can be due to bias (e.g., initial-condition effects) or to problems with the pseudorandom-number generators. If feasible, an analytical analysis is often preferable to simulation, because of the lack of sampling error. Simulation models can also be computationally demanding, because each simulation describes only one single setting. Therefore, several repeats with varying input data are necessary, whereas an analytical model allows the calculation of the exact characteristics of the system for several settings. Furthermore, an inappropriate level of model detail, failure to collect adequate system data, and using wrong performance indicators for comparisons are common pitfalls in both analytical and numerical simulation studies (Law 2007).

The increasing computational power and the availability of 3D printers provide tools for new modeling approaches. Several simulations can be run in parallel, e.g., long running CFD simulations can be accompanied by experiments with 3D printers, whereas the analytical model is evaluated as a baseline. Combinations of the following approaches are possible: (i) analytical models, (ii) numerical simulation, (iii) surrogate models, (iv) lab experiments, and (v) field experiments. The central question in this context is: *Are there any benefits in combining different modeling approaches and can the weakness of one approach be compensated by other approaches?* To answer this question, an approach to combine these heterogeneous results is necessary. This article presents a new approach to handle several simulation models in parallel, which will be referred to as *optimization via multimodel simulation* (OMMS). The OMMS approach can be used as the central part of the well-established optimization via simulation methodology (Fu 1994). To exemplify OMMS, a real-world application is used: cyclone dust collectors. This article presents results from an experimental study, which can be regarded as a proof-of-concept. For the experiments, we have chosen a combination of four different modeling approaches:

- (M_A) analytical,
- (M_C) CFD simulation,
- (M_S) surrogate (metamodels), and
- (M_P) 3D printing models.

Many textbooks describe methods for finding the best model, but do not discuss the combination of several models. Nelson (1995) states that textbooks "tend to give the impression that there is a unique best model of any real or conceptual system. This is not correct." More than one type of model will be used in practice.

The idea of using different models with different resolutions has been discussed in the literature for many years. Zeigler and Oren (1986) describe multiple levels of model aggregation (resolution, abstraction). These levels depend on the objectives, knowledge, and the available budget (resources, e.g., time). Fishwick and Zeigler (1992) present a formalism and a methodology for developing multiple, cooperative models of physical systems from qualitative physics. Barzier and Perry (1991) describe a two-level modeling approach for developing simulation models in the shipbuilding industry. Chaudhuri et al. (2015) describe a flapping wing optimization task. They use multiple surrogates, multiple infill criteria, and multiple points for the same experimental data set. Kazemi et al. (2016) use different machine learning approaches to create simple and reliable models for predicting granule size distributions. An iterative procedure assisted by cross validation was implemented to find the best model among thousands. The cyclone modeling, simulation, and optimization approach presented in our study is related to the work of Preen

and Bull (2014), who optimized vertical-axis wind turbines using miniaturized 3D-printed wind turbines.

Yang (2003) states that selection of one model can be better when the errors in prediction are small and that the model combination works better when the errors are large. Zerpa et al. (2005) demonstrate that a weighted average surrogate model performs better than individual surrogates. Instead of using individual surrogate models, Goel et al. (2007) propose using the best surrogate or a weighted average surrogate model from an ensemble of surrogates to identify regions of high uncertainty and to improve the robustness of the approximations. They report that there is no single surrogate model that consistently performs better than the others and that ensembles of surrogates "performed at par with the corresponding best surrogate model for all test problems" and the ensemble based approach yields robust solutions. Simpson et al. (2012) present a thoughtful review of several multimodel approaches. They state that "the use of multiple surrogates (i.e., a set of surrogates and possibly a weighted average surrogate) is very appealing in design optimization due to the fact that the best surrogate may not lead to the best result; and complementary because fitting many surrogates and repeating optimizations is cheap compared to cost of simulation." They also describe a multidisciplinary approach which is not directly comparable to OMMS, because independent models for different sub-systems are combined rather than integrating several models of the same system. Müller and Shoemaker (2014) use Latin hypercube sampling to generate initial design points, cross-validation to evaluate the model quality of several surrogate models, and the Dempster-Shafer theory approach for the combination of surrogate models (Dempster 1968). They demonstrate that surrogate model combinations containing radial basis function models performed best.

Furthermore, co-Kriging, which is a popular method that combines results from fine and coarse grained models, can be mentioned in this context (Forrester et al. 2007). Typically, co-Kriging tries to combine data from models which have different fidelity, e.g., a fine model that is expensive to compute and a less accurate, coarse model, which is cheaper to compute. In contrast to single-fidelity Kriging models, co-Kriging attempts to learn the correlation between the coarse and fine model, thus being able to exploit the larger amount of data derived from the coarse model to improve the representation of the expensive, fine model. This could be used for the meta-modeling step, especially when different levels of fidelity are available.

In general, there are two options to deal with multiple models: (i) selection of the best model and (ii) combination of results from several models. Most approaches try to select one model, whereas OMMS combines results from several models using stacked regression (Wolpert 1992; Bartz-Beielstein 2016). Our study presents an integrated simulation and experimentation methodology on various scales (or layers).

This paper is structured as follows: Cyclone dust collectors are briefly described in Section 2. Section 3 presents the OMMS loop. Section 4 compares results from different modeling approaches. Experimental results based on these modeling approaches are presented in Section 5. How to combine results from various models via ensemble building is shown in Section 6. Finally, Section 7 gives a conclusion and an outlook.

2 Cyclone dust collectors

Throughout this article, we will use the cyclone geometry optimization problem as an example to demonstrate the introduced OMMS approach. Cyclones are used in oil and gas, iron and steel, chemical and food industry to filter a maximal amount of dust from flue gas (Hoffmann and Stein 2007). They can be applied in extremely harsh and demanding environments, but show a relatively low collection efficiency compared to electrostatic dust collectors. An efficient cyclone requires the optimization of its geometry parameters, which are shown in Fig. 1. Even with today's modern tools, the complexity of cyclone behavior is such that experimental studies are necessary for a solid understanding of the phenomena governing their behavior. The cyclone geometry can be specified by the parameter vector, $\mathbf{x}_g$, with the following entries: inlet width b_e, body diameter D_a, diameter of the vortex finder D_t, diameter of the dust exit D_u, total height h, inlet height h_e, vortex finder immersion h_t, and cylinder height h_z. In addition to these geometry parameters, $\mathbf{x}_g$, the specification of the operating parameters, $\mathbf{x}_p$, is necessary. The geometry and process parameter sets are shown in Table 1. We will concentrate in this study on the collection efficiency as specified in Löffler (1988), which will be explained in Section 4.1.

3 Optimization via multimodel simulation in the loop

In the *optimization via simulation* setting, the goal is to perform runs of the simulation model in an efficient manner and to determine those input variables, which result in an optimal (or near optimal) solution (Fu 1994). The OMMS approach extends the standard optimization via simulation setting by integrating results from several model types. In contrast to mathematical models, which usually require some input values only, data-driven models require the specification of input and output values. To clarify the data

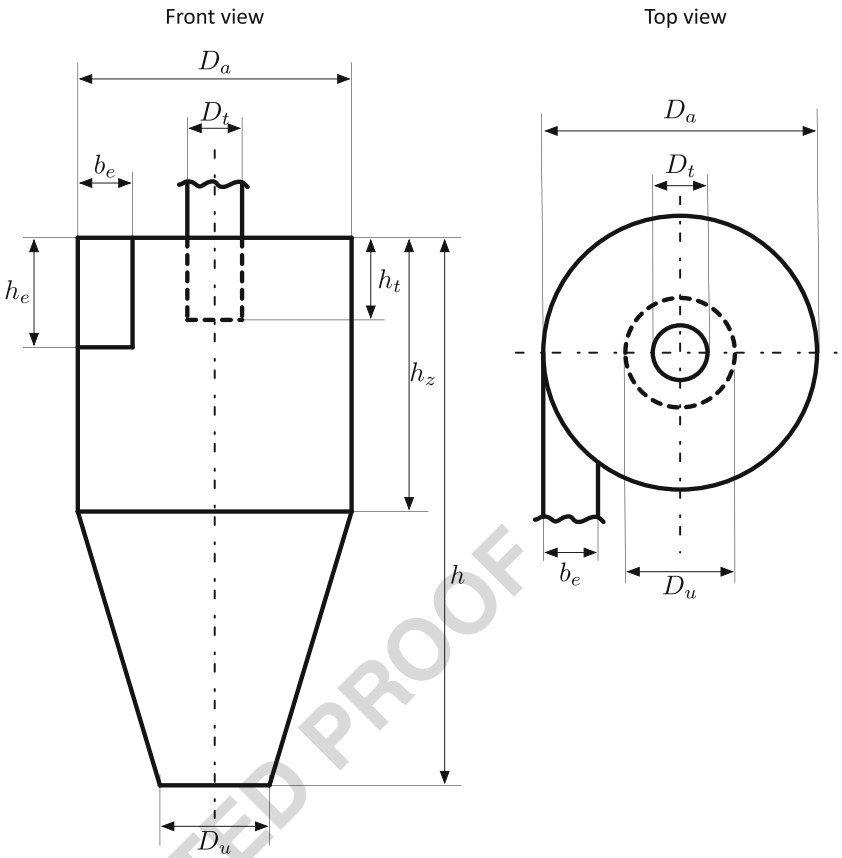

Fig. 1 Standard geometry of the cyclone considered in this study. The corresponding geometry parameters, $\mathbf{x}_g$, are described in Table 1

flow and model building process in the OMMS approach, the following model categories will be used:

- *X-models* use input parameters, e.g., geometry and process parameters.
- *XY-models* use the input parameters as well as the corresponding output values, e.g., collection efficiency.

So, the analytical (M_A), CFD (M_C), and 3D printing (M_P) models are considered as X-models, whereas the surrogate (M_S) models are XY-models.

The main idea of OMMS can be summarized as follows: First, data from different X-models are collected. Then, data-driven (XY-models) are trained, based on the available data. These models are combined in an ensemble approach based on stacked generalization. Finally, the combined ensemble is subject to an optimization run, after which a promising candidate solution is evaluated to receive new data, and the next iteration starts. The general concept of OMMS is further illustrated in Fig. 2. Here, we consider the optimization of the cyclone's geometry parameters, $\mathbf{x}_g$.

OMMS consists of the following steps:

(S-1) Select an initial design. Set $t = 1$, where t denotes the number of parameter sets. The first set of geometry parameters, $\mathbf{x}_g^{(t)}$, is generated.

(S-2) Specify the process parameters $\mathbf{x}_p$. They are not changed during the optimization.

(S-3) Select X-models (e.g., CFD, analytical). In addition to the geometry and process parameter sets, further parameters might be necessary for each separate model. These model specific parameters will be referred to as $\mathbf{x}_m$. For example, the CFD simulator requires the specification of parameters for heat transfer, surface properties, damping, collision, and radiation. These parameters are not used in other simulation models. They are not changed during the optimization. The set $\mathbf{x}^{(t)} = (\mathbf{x}_g^{(t)}, \mathbf{x_p}, \mathbf{x_m})$ will be used to build the X-models.

(S-4) Build X-models. To build these models no information about the dependent (output) variable y is needed. In this step, one or several models ($f_1, \ldots, f_p$) from the set of X-models, which comprehends 3D-printed objects, analytical model formulas, or CFD simulation models, are generated. The construction process results in several models, which use the same set of parameters $\mathbf{x}^{(t)}$.

(S-5) Evaluate models. The X-models are evaluated, i.e., each model generates an output: $f_j : \mathbf{x}^{(t)} \rightarrow y_j^{(t)}$. Note, some models generate a deterministic

JrnlID 158_ArtID 1934_Proof#1 - 02/02/2018

Q2 **Table 1** Nomenclature from Löffler (1988)

Parameter	Units	Values (L, M, S)	Type	Description	Optimized
b_e	mm	12.8; 9.92; 7.97	$\mathbf{x}_g$	Inlet width	Yes
D_a	mm	80.64; 116.48; 39.97	$\mathbf{x}_g$	Body diameter	Yes
D_t	mm	26.88; 29.12; 19.98	$\mathbf{x}_g$	Diameter of the vortex finder	Yes
D_u	mm	26.88; 39.04; 15.04	$\mathbf{x}_g$	Diameter of the dust exit	Yes
h	mm	160; 160; 160	$\mathbf{x}_g$	Total height of the cyclone	Yes
h_e	mm	38.4; 29.6; 19.98	$\mathbf{x}_g$	Inlet height	Yes
h_t	mm	0; 35; 44	$\mathbf{x}_g$	Vortex finder (outlet pipe) immersion	Yes
h_z	mm	44.8; 29.64; 59.95	$\mathbf{x}_g$	Cylinder height	Yes
r_a	mm	$D_a/2$	*	Cyclone radius	No
r_i	mm	$D_t/2$	*	Radius of the vortex finder	No
h_i	mm	$h - h_t$	*	Height of the imaginary cylinder CS	No
r_e	mm	$r_a - b_e/2$	*	Mean inlet pipe radius	No
F	–	F_e/F_i	*	Ratio between inlet and outlet area	No
F_e	mm^2	$h_e \times b_e$	*	Inlet area	No
F_i	mm^2	$\pi \times r_i^2$	*	Outlet area	No
v_e	ms^{-1}	20	$\mathbf{x}_p$	Inlet velocity	No
λ_g	–	0.005	$\mathbf{x}_p$	Load-free friction coefficient	No
μ	Pa s	1.8×10^{-5}	$\mathbf{x}_p$	Viscosity	No
ϱ_f	kg/m^3	1.2000	$\mathbf{x}_p$	Gas density	No
ϱ_p	kg/m^3	2700	$\mathbf{x}_p$	Particle density	No
c_{roh}	kg/m^3	0.061	$\mathbf{x}_p$	Raw gas concentration	No
B	–	c_{roh}/ρ_f	*	Mass load	No
v_i	ms^{-1}	$\dot{V}/(\pi r_i^2)$	*	Velocity vortex finder (outlet pipe)	No
$v_r(r_i)$	ms^{-1}	$\dot{V}/(2\pi r_i(h - h_t))$	*	Radial gas velocity on the outlet pipe	No
$v_{\varphi i}$	ms^{-1}	$(r_i r_e \pi)/(\alpha F_e + h_i r_e \pi \lambda)$	*	Tangential velocity at CS	No
$\dot{V}$	m^3/h	$F_e \times v_e$	*	Volumetric flowrate through the cyclone	No
λ	–	$\lambda_g(1 + 2\sqrt{B})$	*	Wall friction factor; friction coefficient	No

"Values (L, M, S)" refers to the values of the geometry parameters $\mathbf{x}_g$ for the Löffler, Muschelknautz E., and Stairmand high efficiency cyclones, respectively. The vortex finder immersion, h_t, is modified for every cyclone geometry. The type "$\mathbf{x}_p$" denotes operating parameters. Parameter values, which depend on other values, are labeled as "*" in the Type column. Parameters to be optimized are labeled in the last column

output, e.g., CFD models, whereas other, e.g., 3D-printed models, generate stochastic (noisy) outputs. Therefore, repeats should be considered for the stochastic models, to improve the quality of the measured values (Law 2007; Haftka et al. 2016).

(S-6) Collect results. Besides the set of pairs $\{(\mathbf{x}^{(k)}, y_j^{(k)})\}$, for $k = 1, \ldots, t$ and $j = 1, \ldots, p$, additional results $\{(\mathbf{x}^{(m)}, y_l^{(m)})\}$, for $m = 1, \ldots, s$ and $l = 1, \ldots, q$, e.g., from historical data or data from the literature, can be used in the construction of the metamodels.

(S-7) Select XY-models. XY-models use the parameter set, $\mathbf{x}^{(k)} = (\mathbf{x}_g^{(k)}, \mathbf{x_p})$ as well as the corresponding output values $y_i^{(k)}$ for model building, with $k = 1, \ldots, (t + s)$. In general the number of design points $(p + s)$ is required to be large enough to allow building reasonable models. These XY-models will be referred to as level-0 models.

(S-8) Build metamodel. Stacking is used to build a level-1 metamodel by combining information from several level-0 models. The level-1 metamodel will be referred to as $M_S^\star$. The level-1 metamodel is typically a relatively simple linear model. Instead of stacking, a weighted combination of level-0 models or co-Kriging can be used. If level-0 models of similar fidelity are combined, stacking is recommended. In a mixed case, a co-Kriging model could be integrated into a stacked metamodel (i.e., as a single level-0 model). Other ensemble techniques may also be applicable, e.g., bagging or boosting (Murphy 2012). However, stacking is very effective even when combining

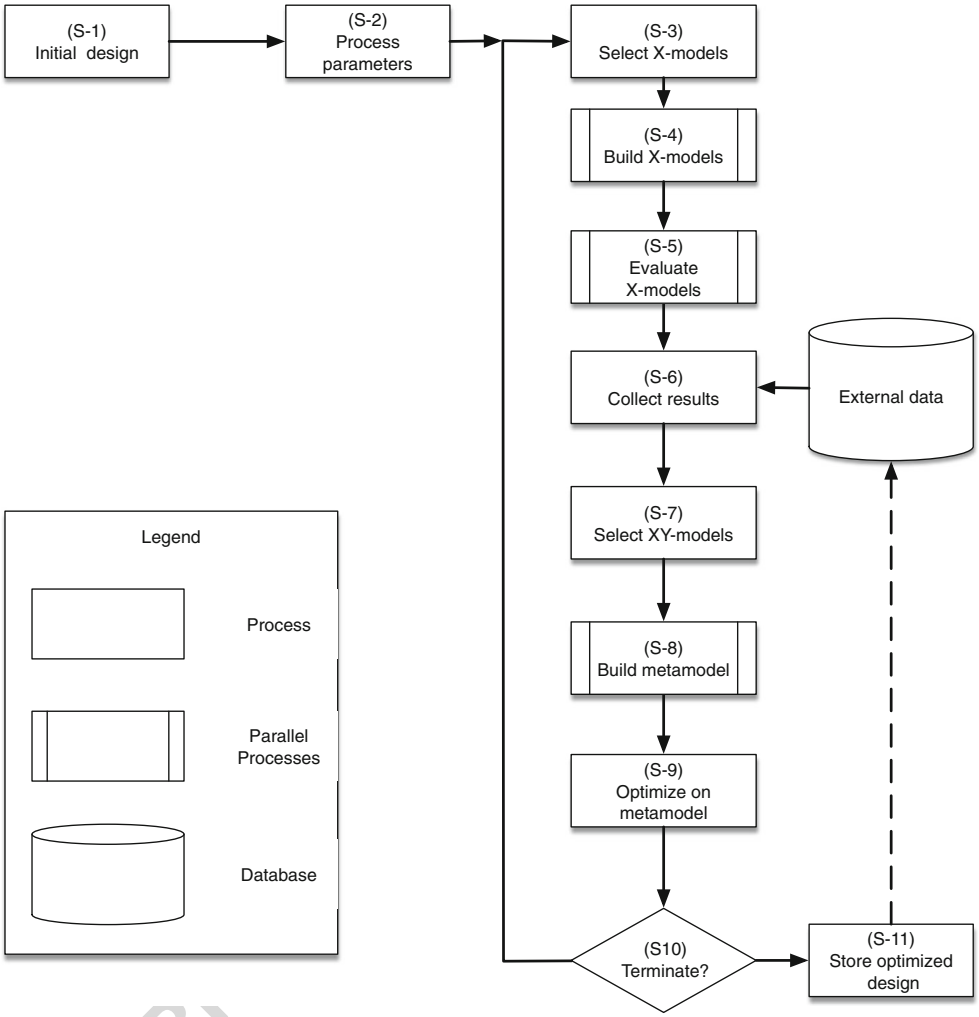

Fig. 2 Optimization via multimodel simulation in the loop. Several simulation models are used in parallel. Elements of the first set of models, i.e., during steps (S-3), (S-4), and (S-5), can be one or several CFD simulators, analytical models, or experiments based on 3D-printed objects. Results from these different models are collected and optionally combined with additional results, which were stored in a database. The second set of models is built during step (S-8). Models from the second set are classical surrogate models, e.g., neural networks, linear regression models, or Kriging models. Because simulation results, i.e., y-values are available at this stage of the multimodel simulation process, a broader set of models can be used than during the first steps (S-3) to (S-5). Results from these models can be combined in several ways. We describe an approach that is based on stacked generalization (Wolpert 1992). Optimization is performed on the stacked model (S-9). The termination criterion depends on the user's requirements, e.g., when the budget of evaluations is expended (S-10)

only few, strong learners and it provides additional information, e.g., the contribution of each of the combined models to the ensemble.

(S-9) Optimize on the metamodel. The model $M_S^\star$ is used as a surrogate for performing the optimization step. The optimization results in a new set of promising geometry parameters, which will be evaluated in the following step. The counter for the number of parameter sets t is incremented and the new design can be referred to as $\mathbf{x}_g^{(t)}$. Instead of increasing t by one, several new design points, e.g., from models with different run times, can be added to the parameter set.

(S-10) Check the termination criterion. Usually, if the budget, i.e., simulation time, is exhausted or the desired solution quality is reached, the process is stopped and the result is presented.

(S-11) Store the optimized design, $\mathbf{x}_g^\star$. Optionally, it can be added to a database.

4 Modeling approaches

To exemplify the OMMS approach, four different modeling approaches are described in the following: analytical (M_A), surrogate (M_S), CFD simulation (M_C) and 3D printing (M_P).

4.1 The analytical model (M_A)

A broad variety of analytical models intended to predict cyclone separation performance exists in the literature (Löffler 1988; Overcamp and Mantha 1998; Hoffmann and Stein 2007). The analytical approach developed by Barth (1956) and Muschelknautz (1972) can be considered as standard. It will be referred to as the *Barth-Muschelknautz method of modeling* and denoted as (M_{AB}). This method is based on the assumption that a particle carried by the vortex is influenced by two forces: a centrifugal force and a flow resistance. They are expressed at the outlet pipe radius r_i

Table 2 Particle size distribution table

Particle size $x[\mu m]$	Δx	Mean $\tilde{x}$ $[\mu m]$	$\Delta Q_e(x)$	Cumulative
0-1	1	0.5	0.1	0.1
1-2.7	1.7	1.85	0.1	0.2
2.7-5.5	2.8	4.1	0.1	0.3
5.5-8.7	3.2	7.1	0.1	0.4
8.7-12.7	4	10.7	0.1	0.5
12.7-16.9	4.2	14.8	0.1	0.6
16.9-21.2	4.2	19	0.1	0.7
21.2-25.4	4.2	23.25	0.1	0.8
25.4-30.8	5.4	28.1	0.1	0.9
30.8-63	31.2	46.9	0.1	1.0

Values correspond to the dust used in the 3D printing experiments

where the highest tangential velocity occurs. The (M_{AB}) model is based on the balance of forces acting on a particle that is rotating in the cyclone. Small particles leave the cyclone through the vortex finder, whereas large particles move to the cyclone wall. The cut size, x_{50}, plays a central role in these calculations. For cyclones, particles of size x_{50} have a 50-50 chance of being captured. Smaller particles are less likely to be captured and larger particles are more likely to be captured. Two forces act on a particle rotating on the cylindrical control surface, which is assumed to separate the outer region of downward flow from the inner region of upward flow. These forces are (i) the centrifugal force acting outward and (ii) the Stokesian drag acting inward. By equating these forces, Barth (1956) developed an analytical model for the cut size x_{50}. For a given cut size, the fractional efficiency curve, $T_{x_{50}}(x)$ assigns an efficiency to the particle diameter. The *overall collection efficiency* E is predicted according to:

$$E = \int_{x_{\min}}^{x_{\max}} T_{x_{50}}(x) q_e(x) dx \approx \sum_{x_{\min}}^{x_{\max}} T_{x_{50}}(\tilde{x}_i) \Delta Q_e(x_i), \qquad (1)$$

where $x_{\min}$ is the lower bound of the particle size, $x_{\max}$ is the upper bound of the particle size, $\tilde{x}_i$ is the mean particle size in each fraction, $\Delta Q_e(x_i)$ is the change in distribution of particle sizes and $q_e(x) = \Delta Q_e(x_i)/\Delta x_i$. The particle size distribution table, which was used in our studies, is shown in Table 2.

The collection efficiency, which is calculated using the analytical model (M_{AB}), will be referred to as E_B. Results from our collection efficiency calculations for

Table 3 Results from analytical models (M_A)

Type	h_0 [mm]	h_e/h_0	b_e/h_0	D_t/h_0	h_t/h_0	h_z/h_0	D_a/h_0	D_u/h_0	E_B	E_M
Löffler	25,000	0.2400	0.0800	0.1680	0.2600	0.2800	0.5040	0.1680	89.3334	90.52
Muschelknautz E.	9,340	0.1852	0.0621	0.1820	0.3330	0.1852	0.7281	0.2441	89.9122	91.11
Stairmand high eff.	12,650	0.1249	0.0498	0.1249	0.1249	0.3747	0.2498	0.0941	89.0490	87.92
Muschelknautz D.	8,630	0.2167	0.0626	0.1379	0.3685	0.3036	0.4137	0.2260	89.5968	91.33
Storch 4	16,160	0.1609	0.0235	0.0724	0.1089	0.5625	0.1609	0.0563	91.4006	87.04
Storch 3	8,210	0.2034	0.0731	0.1303	0.2436	0.5627	0.2339	0.1121	86.9871	88.29
Storch 2	10,970	0.1714	0.0483	0.0985	0.2179	0.4230	0.2051	0.0766	89.6379	90.37
Storch 1	19,430	0.0515	0.0515	0.0633	0.0726	0.2820	0.1879	0.0329	92.4774	89.49
Tengbergen C	9,300	0.1075	0.1075	0.1204	0.1559	0.2011	0.3624	0.1204	90.4930	92.16
Tengbergen B	6,040	0.2964	0.0927	0.1854	0.3709	0.5364	0.3477	0.1854	83.8606	87.45
Tengbergen A	6,470	0.2087	0.1144	0.1731	0.2427	0.2782	0.4281	0.3122	87.7651	89.80
TSN -11	9,590	0.1919	0.0563	0.1418	0.2523	0.2284	0.3629	0.1606	89.3807	90.58
TSN -15	11,240	0.1477	0.0534	0.1406	0.3114	0.5240	0.2367	0.1059	86.6353	89.50
Stairmand high flow	7,550	0.1868	0.0940	0.1868	0.2185	0.3748	0.2517	0.0940	81.2173	84.33
Van Tongeren AC	12,310	0.1210	0.0544	0.0812	0.2640	0.3542	0.2640	0.1056	92.1290	91.25
Vibco	7,200	0.1542	0.1250	0.1542	0.1722	0.3167	0.3972	0.0917	88.6851	90.47
Lapple GP	11,310	0.1247	0.0628	0.1247	0.1565	0.5004	0.2502	0.0628	88.8657	89.01

Cyclone geometries taken from Hoffmann and Stein (2007). Proportions of the geometries shown in the first three rows (Löffler, Muschelknautz, and Stairmand) were used for the 3D printing experiments. All values were scaled by the height h_0 of the real cyclone. Since the printed cyclones have an absolute height of 160 mm, the relative values from columns three to nine are multiplied with 160. The Barth model, i.e., (1), was used to determine the collection efficiency E_B. Results from the Mothes model, which will be introduced in Section 6, are shown in the E_M column. Both models use the process parameters from our study and the scaled geometry parameters of the original cyclones. The same particle size distribution (silica sand) was used for the calculations to obtain comparable results

models from the literature and for models used in our experiments are shown in Tables 3 and 5, respectively. The corresponding function is available in the R package SPOT as `funCyclone()`.

4.2 CFD Simulations (M_C)

Computational Fluid Dynamics simulations have proven to be useful for the study of the fluid and particle flows in cyclones (Hoekstra et al. 1999). They have clear advantages for the understanding of the details of the flow in cyclones, but also limitations in terms of modeling cyclone separation performance accurately (Hoffmann and Stein 2007). Numerical simulations are performed by solving the unsteady-state, three-dimensional *Reynolds averaged Navier-Stokes* (RANS) equations combined with a closure model for the turbulent stresses and the large eddy simulation approach.

The CFD simulations were carried out with the open source software OpenFOAM, which has been developed to solve numerical problems (Konan and Huckaby 2015). The mesh for these CFD simulations consists of approximately 30,000 to 50,000 hexahedral cells. The transient MPPICFoam solver was chosen to calculate the two-phase flow (Euler-Lagrange). The cyclone simulation from the OpenFOAM cyclone tutorial was used as a basis (OpenFOAM Foundation 2016). The settings for `fvSchemes`, `fvSolution`, `transportProperties`, and `turbulenceProperties` were adapted to obtain the same setup as for the 3D printing experiments. The settings in the `kinematicCloudProperties` file were adjusted to the characteristics of the used particles. The density of the particles was changed to 2,700 kg/m^3 (as in Table 1 above). Using the `generalDistribution` model, the particle distribution from Table 2 can be mapped precisely. In the simulations, 20,000 parcels represent the entirety of the particles, where each parcel has the same mass. This amount was chosen as a compromise between calculation time and mass per parcel. The minimization causes a lower error when a parcel escapes at an outlet. The `heatTransfer`, `surfaceFilm`, `damping`, `stochasticCollision`, and `radiation` submodels were left unchanged at the "off" state. In the experiments, a total of 6 g was spread over 10 thrusts and the waiting time between each thrust was approximately 3 s. The simulation takes only one thrust of 0.6 g instead of performing the 10 repetitions in order to avoid very long simulation times. The particle velocity in the simulation was set to the same value as the determined velocity of the air at the inlet. Overall, a time frame of 3 s is simulated. For this, a total calculation time of approximately 96 h (wall-clock time) using 16 processor cores is required. The simulation was controlled by the time step and relaxation factors and behaved relatively stable.

After each experiment, a certain amount of dust remains in the cyclone. We consider dust as being separated, when it leaves the cyclone through the dust exit (usually at the bottom of the cyclone). The evaluation of the simulation results is shown in column (M_C) in Table 5. If the remaining dust in the cyclone is also considered as being separated, the collection efficiency is increased. Collection efficiencies calculated with CFD models (M_C) will be referred to as E_C.

4.3 Surrogate modeling (M_S)

Computational fluid dynamics simulations are computationally expensive. A well-known approach to handle costly objective functions is to use data-driven XY-models. They are referred to as response surface models or surrogate models (Jin 2003; Kleijnen 2008). That is, data-driven surrogate models may be constructed based on experimental results. Then, an optimization algorithm can search on the cheap surrogate model instead of using the expensive CFD simulations. For the purpose of demonstration, we provide an example by training a simple surrogate model. Here, we have chosen a standard linear regression model, because it is available out-of-the-box. It models the estimated collection efficiency as a function of 17 different geometries, which were specified in Table 3.

The data that were used to fit the surrogate model are based on the the analytical model (M-AB) developed by Barth (1956) and Muschelknautz (1972). The estimated collection efficiencies that were calculated using the surrogate model (linear regression model) will be referred to as E_S.

$$E_S = 95.56 - 0.58 D_t + 0.11 D_a. \quad (2)$$

This linear regression model was reduced from a full model considering h_e, b_e, D_t, h_t, and D_a with variable selection based on stepwise model reduction using the Akaike information criterion (AIC). The parameters D_u and h_z are not present because they are not considered by the analytical model. The parameter h was removed as it is constant (due to the constant height of the 3D printed cyclones). The model reduction via AIC yielded only two remaining parameters, which may mostly be due to the small size of the training data set. The available 17 samples do not seem to allow for concluding much about the removed parameters. Either, the data set explores these values insufficiently, or else, their relation is nonlinear and hence not detected by the model. The parameters that were removed based on the stepwise model reduction are not necessarily without influence. In fact, they are known to influence the behavior of cyclones in practice.

JrnlID 158_ArtID 1934_Proof#1 - 02/02/2018

Table 4 Geometries, i.e., $\mathbf{x}_g$ values, of the 3D printed cyclones shown in Fig. 3

Type	b_e	D_a	D_t	D_u	h	h_e	h_t	h_z
Löffler	12.8	80.64	26.88	26.88	160	38.4	0	44.8
Löffler	12.8	80.64	26.88	26.88	160	38.4	35	44.8
Löffler	12.8	80.64	26.88	26.88	160	38.4	44	44.8
Muschelknautz E.	9.92	116.48	29.12	39.04	160	29.6	0	29.64
Muschelknautz E.	9.92	116.48	29.12	39.04	160	29.6	35	29.64
Muschelknautz E.	9.92	116.48	29.12	39.04	160	29.6	44	29.64
Stairmand high eff.	7.97	39.97	19.98	15.04	160	19.98	0	59.95
Stairmand high eff.	7.97	39.97	19.98	15.04	160	19.98	35	59.95
Stairmand high eff.	7.97	39.97	19.98	15.04	160	19.98	44	59.95

A total height, h of 160 mm was chosen. This table shows absolute values, which were determined using the relative values from Table 3 multiplied by 160

However, based on the available data as well as the chosen linear model, their influence on the measured performance indicator can not be determined with any confidence. The diameter of the vortex finder, D_t, and the cyclone body diameter, D_a, have a significant effect on the collection efficiency. Although this model already shows good accuracy, with an adjusted R-squared of 0.94 (estimated on the same data that was used to fit the model), more sophisticated linear models or Kriging models can be fitted (Turner et al. 2013; Kleijnen 2014). Optimization can be performed on this model, e.g., to improve the collection efficiency for a given pressure drop. Even multi-objective optimization techniques can be applied (Elsayed and Lacor 2012; Zaefferer et al. 2014).

4.4 3D-printing model (M_P)

The experiments with the 3D-printed cyclones used standard laboratory equipment: Erlenmeyer flask, stand, pressure gage, precision scale, and a vacuum cleaner. To ensure comparability and interchangeability of the results, the same design and process parameters as in the other modeling approaches were used for the 3D-printing models. Table 4 shows the parameters of the printed cyclones. The process parameters, $\mathbf{x}_p$, are described in Table 1. The model building step (S-4 in the OMMS approach) consists of the (i) 3D computer model generation and the (ii) printing step. The 3D models, described in the *STereoLithography, Standard Tessellation Language* (STL) are created using a Python script, which uses the FreeCad Python library.[1] The model was exported to an STL file, which can be interpreted by a large number of 3D printers. The printed cyclone models are shown in Fig. 3.

Today, a broad variety of 3D printers as well as different materials are available. The printing technique as well as the materials have to meet certain requirements. The cyclone has to be robust, because it is held in a fixed position and has to withstand the flow of air and dust. Due to the hollow shape of the cyclone, a ProJet CJP 660pro printer was chosen, which uses gypsum powder (Visijet PXL) as printing substrate. The entire cyclone was printed in one step. The gypsum powder has to be hardened after surplus gypsum powder has been removed from the interior. Cyan acrylate ("ColorBon") was used to avoid electrostatic charging. It produces a sufficiently smooth surface and improves the stability. The experiments were performed at room temperature. Significantly higher temperatures may require a different choice of material. Printing a single cyclone model takes about three hours, plus one hour for refinishing, dust removal, and infiltration, and one additional hour for curing.

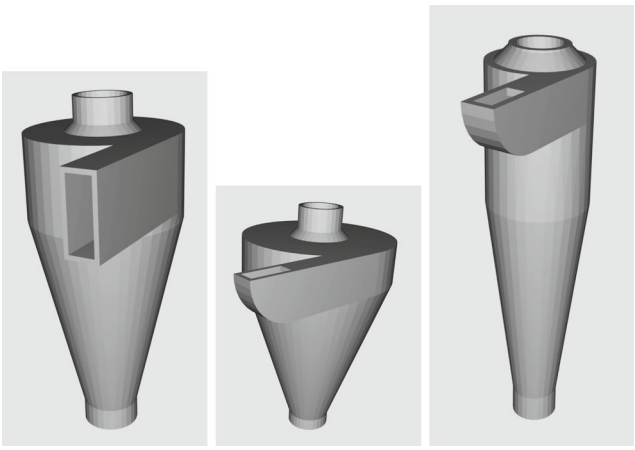

Fig. 3 Three printed cyclone models, which were used with three different h_t values in the (M_P) experiments. From left to right: Löffler, Muschelknautz, Stairmand. See Table 4

[1] http://www.freecadweb.org

Table 5 Estimated collection efficiencies from five different modeling approaches: E_M (Mothes), E_B (Barth-Muschelknautz), E_C (CFD), E_P (3D-printed cyclones), and E_S (surrogate model) for three different cyclone types and three different outlet pipe immersions (h_t)

Type	h_t	E_M	E_B	E_C	E_P	E_S
Löffler	0	–	90.19	88.57	86.80(± 4.39)	89.80
Löffler	35	90.45	89.49	90.91	92.13(± 5.12)	88.95
Löffler	44	90.54	89.27	90.19	90.84(± 3.84)	88.73
Muschelk.	0	–	91.14	82.04	88.50(± 8.24)	92.13
Muschelk.	35	91.02	90.37	82.55	93.67(± 2.17)	91.28
Muschelk.	44	91.07	90.15	82.43	92.87(± 2.94)	91.06
Stairmand	0	–	89.44	91.20	90.53(± 5.68)	88.79
Stairmand	35	89.34	88.70	95.95	94.29(± 1.34)	87.93
Stairmand	44	90.57	88.45	95.60	95.50(± 2.50)	87.71

The 3D-print column (E_P) shows mean (and standard deviation) from five repeats. Obvious outliers were removed. The Mothes model (M_{AM}) requires outlet pipe immersion values $h_t > 0$. The values shown in column $E_S(M_S)$ are based on the simple linear regression model from (2)

Besides the selection of a printer and material, the characteristics of the dust have to be selected. The distribution of particle sizes should not vary to prevent fluctuations in the results. If the particles are too large, they may be too easy to separate from the gas. If there are too many, they may even block the flow inside the cyclone. If the particle size is too small, the particles cannot be separated. The chosen dust is silica sand with a maximal particle size of 63 μm. Its particle size distribution is shown in Table 2. Collection efficiency calculated using 3D-printed cyclones (M_P) will be referred to as E_P.

5 Single models: experimental results

An experimental design was set up to measure the effects of the cyclone shapes and the outlet pipe immersion h_t on the collection efficiency E_P in the 3D printing (M_P) experiments. Three different values for the immersion of the outlet pipe ($h_t = 0, 35, 44$ mm) were chosen for each of the three cyclone geometries (Löffler, Muschelknautz, and Stairmand), see Fig. 3. Since every experiment was repeated five times, altogether 45 experiments were performed. Measurements with unusual collection efficiency values were considered as outliers. For example, an efficiency value larger than 100% was probably a result of insufficient cleaning of the cyclone between tests, because these outliers occurred at the beginning of every series of tests. Results from these experiments are shown in column (M_P) in Table 5.

Figure 4 shows boxplots, which visualize the collection efficiencies of the three different cyclones and outlet pipe immersions. As can be seen in Fig. 4, the median of the collection efficiency values of the Löffler cyclone is lower than 90%, whereas it is larger than 90% for the Muschelknautz and Stairmand cyclones. According to the (M_P) column in Table 5, the Stairmand cyclone has the highest average efficiency. The efficiency at every immersion value is higher than the previous two cyclones with the same outlet pipe immersions. Furthermore, results from the Stairmand cyclones are more robust, i.e., smaller variances, than results from the other two cyclones. A very small outlet pipe immersion value (0 mm) results in a reduced collection efficiency as can be seen in Fig. 4.

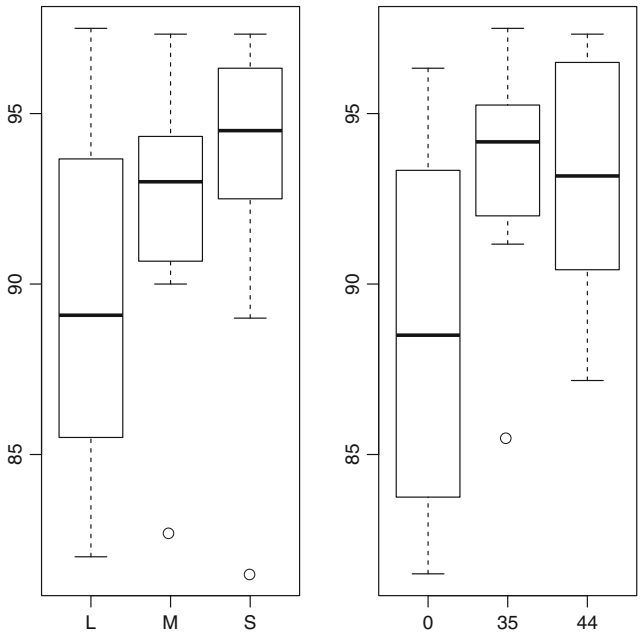

Fig. 4 Collection efficiencies E_P (in %) of the the three different cyclones (left: Löffler, Muschelknautz, Stairmand) and geometries (right: 0, 35, 44 mm) used in the (M_P) experiments. The boxplots indicate that collection efficiency might be improved if Stairmand cyclones and an increased h_t value are used. Thick bars indicate the median values

This result could be observed consistently for all cyclone types. The results from the 3D printing experiments can be summarized as follows: (i) there are high variances in the measured values, and (ii) the experimental results indicate that the collection efficiency, E_P, increases with increasing vortex finder immersion, h_t, values.

Due to the complex nature of the lab experiments, the M_P results were subject to inaccuracies. While 3D printing reduces the cost of experiments significantly compared to full-scale simulations, the experiments themselves are still time-consuming and require material resources. With the cost of experiments, noise of evaluation (due to manufacturing as well as measurement inaccuracies), and the inherent complexity of the search-space due to its combinatorics, the modeling based on 3D-printed cyclones poses a major challenge. Due to these technical difficulties and the complexity of the experimental setup, the (M_P) results presented in this study are not conclusive yet. Problems related to the high variance in the (M_P) model can be reduced by improving the experimental procedure, e.g., by enhancing the cleaning process between the experiments. However, even if the data itself does not enable to draw reliable conclusions for the design of an optimal cyclone geometry, they are suitable to demonstrate the OMMS approach.

In addition to the discussion of the results from the 3D printing experiments, we consider results from the CFD simulations. The experimental CFD results indicate that the collection efficiency is worst if the vortex finder immersion values are zero. This mostly agrees with the findings for the 3D-printed cyclones.

In contrast to the results from the other models, data from the analytical model (M_{AB}) show a negative effect of the immersion length, h_t, on the efficiency: smaller h_t values result in an increased collection efficiency E_B, cf. Table 5. Further calculations revealed, that this is a systematic error in the Barth-Muschelknautz method of modeling. To fix this error, we implemented an additional analytical model, which was developed by Mothes and Loeffler (1984). This model and its collection efficiencies will be referred to as M_{AM} and as E_M, respectively. The Mothes model partitions the cyclone into regions. Differential equations are used to model diffusive interchange between neighboring regions. Details are described in Hoffmann and Stein (2007). The corresponding values can be found in Table 3 (column E_M). If the Mothes model is used, larger h_t values result in an increased collection efficiency. Because the other models (CFD and 3D printing) support this relationship, the Mothes model is preferred to the Barth model in the reminder of this study. Note, the Mothes model is not able to calculate the collection efficiency if h_t equals zero. However, this constraint is of minor practical relevance. Using the Mothes model changes the coefficients of the linear model presented in (2). The modified equation reads $E_M = 90.40 - 0.25 D_t + 0.08 D_a$.

6 Metamodel-based optimization and stacking

The metamodeling is performed in two steps. First, a set of surrogate models has to be chosen in step (S-7) of the OMMS approach. These models will be denoted as level-0 models. They are combined in a second step, which is step (S-8) of the OMMS approach. To combine the level-0 models, stacking is used. Stacking can be seen as a "more sophisticated version of cross-validation" (Wolpert 1992). Cross-validation *selects* the best of a set of models based on their estimated error, whereas stacking *combines* the models. Stacking partitions the available training data set into training and test data subsets. The quality of predictions from models using the training data subsets, the so-called level-0 models, can be evaluated with the test data subsets.

For each of the k hold-out subsets, the level-0 models are trained on the remaining data and are used to predict on the hold-out data. For the purpose of combining the level-0 models, stacking trains a so-called level-1 model, which receives the predicted values of each level-0 model instance, and the corresponding true values from the hold-out data as observations. Thus, the level-1 model combines the level-0 models, at the same time trying to minimize their combined error. The pseudo code in Algorithm 1 explains the employed implementation of the training of the stacking ensemble. Once the stacking ensemble (which includes level-0 and level-1 models) is trained, predictions for new data can be made as follows. For each model type, the trained level-0 models from the k folds are combined by averaging their predictions. The average prediction of each model type is then fed into the level-1 model to produce the ensemble prediction. For further details on this, the reader is recommended to read Wolpert's seminal paper (Wolpert 1992). The source code of the ensemble procedure is also publicly available. During our experiments, stacking was performed with the freely available R package SPOT (*sequential parameter optimization toolbox*).[2] It provides several tools for the analysis and optimization of complex problems (Bartz-Beielstein et al. 2005). The SPO function `buildEnsembleStack` implements the metamodel building step (S-8).

[2]Source code and data for performing experiments from this study are available at http://www.gm.fh-koeln.de/~bartz/bart16e. The open source R software package SPOT can be downloaded from https://cran.r-project.org.

Algorithm 1 Training procedure for stacking as implemented in the `buildEnsembleStack` function of the `SPOT` package. The following variables are used: number of folds: k, number of data samples: n, input data: x, output y, number of ensemble models: p. "modelL0" denotes the level-0 model, e.g., random forest or Kriging, whereas "modelL1" denotes the level-1 model, e.g., a linear regression model

```
begin
  Generate k-fold data split s
  comment: Fit p level-0 models:
  for j := 1 to p step 1 do
    comment: For each of k folds of the data:
    for i := 1 to k step 1 do
      comment: Exclude current fold from training data
      xt := x[s ≠ i, :];
      yt := y[s ≠ i];
      comment: Fit j-th modelL0
      fit0[j][i] := modelL0[j](xt, yt);
      comment: Predict excluded data
      yhat [s = i, j] := predict(fit0[j][i], x[s = i, :]);
    od
  od
  comment: Fit modelL1 on results from all models & folds:
  fit1 := modelL1(ybar, y)
end
```

As level-0 models, a simple regression model (lm), a random forest (rf), and a Kriging (kr) model were used in this study. For each of the k folds, these models are trained as usual, e.g., via ordinary least squares (lm), or maximum likelihood estimation (kr). A linear regression model (lm) trained by ordinary least squares is also used as a level-1 model. This level-1 model, which combines results from the level-0 models, uses the following coefficients:

$$\text{Ensemble:} -51.38 + 0.40\ \text{lm} + 0.92\ \text{rf} - 0.90\ \text{kr}, \qquad (3)$$

i.e., the stacked model uses mainly the information from the rf and kr surrogate (level-0) models, and includes information from the lm surrogate model to a lesser amount as well. This result was obtained by training all three level-0 models repeatedly on 10 folds of the 80 sample data set. Hence, each level-0 model instance was trained with 70 training data samples. The training data stems from the earlier described sources: analytical models, CFD models and 3D-printing experiments. Note, that further data-sources can easily be integrated into the data-set. Also, the ensemble model could easily be extended by a larger set of data-driven surrogates.

After generating an objective function from the fit, an optimizer can be applied (S-9). In our example, differential evolution was used, but any other optimizer is applicable. Results from the optimization are as follows: $h_e = 0.0938$, $b_e = 0.0938$, $D_t = 0.1808$, $h_t = 0.2068$, $h_z = 0.2799$, $D_a = 0.2436$, $D_u = 0.0916$. All values are relative to the cyclone height $h_0 = 160$ mm. This geometry results in an estimated efficiency of 93.69%.

Notably, the estimated efficiency of the proposed solution ($E_S = 93.69$) is lower than some of the single-model results for the cyclone variants shown in Table 3. This inconsistency is probably caused by noise in the real-world experiments and impreciseness of the models. The optimization results are compared to the search bounds in Fig. 5. This figure illustrates the recommendations from the optimization on the metamodel. For example, the proposed solution requires intermediate diameters for the dust outlet, D_u, and the immersion depth, h_t. The diameter of the vortex finder, D_t, is close to the maximum possible value. On the other hand, the body diameter, D_a, is rather low. If the optimization budget is not exhausted, these recommendations can be used to print a new cyclone or to perform a CFD simulation and start the next iteration of the OMMS loop.

The validation of the accuracy of the stacked model is an inherent bootstrap problem, because the true collection efficiency is unknown. This also applies to the quantification of the improvement. As a possible solution, results obtained from the metamodel-based optimization could be compared to randomly generated parameter configurations. Since there is no data about the real-world collection efficiency of the proposed cyclone geometries available, existing results could be interpolated or long-running CFD simulations could be considered as the "ground truth".

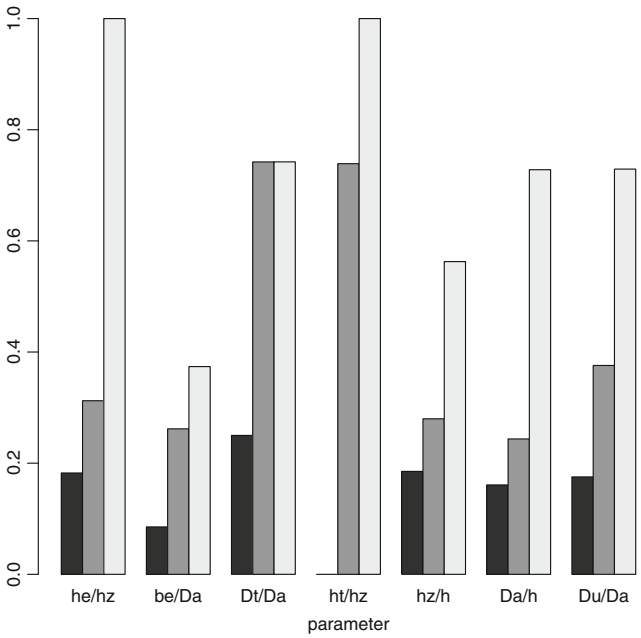

Fig. 5 Optimization on the metamodel (S-9). Barplot of the optimized geometry parameters. Comparing lower and upper bound of decision space with optimum. All values are presented as ratios, relative to the indicated parameters, because the bounds have to be specified as ratios

However, there are two reasons why the CFD simulation can be a problematic choice for validation purposes. First, CFD simulations can only approximate the behavior of real cyclones and may miss important effects. Secondly, due to the high computational cost, the corresponding CFD simulation data is very sparse and hence problematic to be used for validation purposes.

Reasonable artificial test functions that approximate the behavior of problems with different data sources (here: CFD, analytical, 3D-printing) are also not easily available.

While it would be possible to simply estimate quality on some hold-out data (from all combined data sources), this does not provide an estimate of the model with respect to the actual real-world performance.

7 Conclusion and outlook

7.1 Summary

This paper can be seen as a proof-of-concept for the OMMS approach, which combines results from several different modeling approaches. One important research question that we attempted to answer was: *Are there any benefits in combining different modeling approaches and can the weakness of one approach be compensated by other approaches?* While the OMMS approach has theoretical merit, this question cannot be conclusively answered based on the numerical experiments described in this study.

This is owed to the fact, that problems occurred during the data collection with the 3D-printing model (M_P) and the analytical model (M_{AB}). Potential solutions to these problems were discussed in Section 5. These problems occur during the data generation process (experiments, S-1) and not during the OMMS core procedure (stacking, S-7 and S-8). Such problems affect most methods that are driven by experimental procedures.

The proposed OMMS methodology improves the robustness of the optimization, accelerates the optimization-via-simulation process, and provides a unified approach. These benefits will be outlined in the following.

Robustness First results, i.e. the interpretation of the improved values from the OMMS as shown in Fig. 5, indicate that the weakness of one approach can be compensated by other approaches. The stacking procedure also enables OMMS to avoid deterioration of model quality due to selection of models with poor quality (Wolpert 1992). Our observations are in line with results from Goel et al. (2007), who use the best or a weighted average surrogate model instead of individual surrogate models. They demonstrate that "using an ensemble of surrogate models can improve robustness of the predictions by reducing the impact of a poor surrogate model (which may be an artifact of choice of design of experiment or the inherent unsuitability of the surrogate to the problem)". An experimental study performed by Bartz-Beielstein (2016) indicates that these results also apply to stacking, which is used in the OMMS approach. Obviously, if the model information does not reflect the behavior of the true system at all, stacking cannot fix this problem. But this is a general problem, which affects every modeling approach.

Acceleration The combination of several modeling approaches in a single loop can be time consuming. So, one might ask: "What is the advantage of doing so?" In practice, the main costs will be caused by complex simulations (CFD) as well as real-world or lab experiments (here: 3D printing). These results have to be integrated, as they are the only trustworthy way to gather information. Because the running times of the OMMS evaluations are negligible compared to the run time of the CFD simulator, OMMS provides an effective and efficient approach to find improved cyclone geometries. A single CFD simulation took approximately three days, whereas the OMMS runs require a few seconds. If costs are measured in run time, OMMS significantly accelerates the optimization.

Unified approach In practice, there may be situations where one model is preferred to the other. But often, it may be unclear which model is closer to the real system: CFD simulations may fail to account for certain complexities of the real system (e.g., potential chemical reactions) while lab experiments will depend on some down-scaled system. Both may hence have different inaccuracies with respect to different aspects of the system. Therefore, we argue that they should complement each other. That is the reason why they are ultimately combined with an ensemble approach. The stacking approach is also conceptually attractive, because it is a single, unifying approach. The failure of one level-0 model is not critical for creating good predictions. And, ensembles do not require much parameter tuning.

Stacking allows the transfer of knowledge from the ensemble back to a simpler model. In the context of machine learning, Bucila et al. (2006) demonstrate how to compress the knowledge in an ensemble into a single model (Hinton et al. 2015).

7.2 Future research

Investigations about the capability of a combination of surrogate models to deal with noise are of great interest. We employed three individual level-0 model types (random forest, Kriging, and a linear model) and a linear model as a level-1 model. Random forest uses bagging, which is a model averaging approach and relatively robust against

noisy data and outliers. So are linear models, which nonetheless may be deteriorated if strong outliers are present. In such cases, robust linear regression might be more appropriate. Finally, Kriging is able to incorporate noise via the so-called nugget effect, which allows Kriging to smoothen rather than interpolate the data. Regarding the level-1 model, the situation is not so clear, because interactions between these diverse sets of models with different features are very complex. Furthermore, there is not only one type of noise. Therefore, concluding remarks about the effect of noise are beyond the scope of this study and will be subject of further research.

Another issue that may be subject to further research is the level-1 model of the stacked regression ensemble. As shown in (3), the coefficients of the Kriging and the random forest model have opposite signs. The interpretation of this observation is difficult. Previous research showed that restricting the coefficients of the linear level-1 model to non-negative values can produce better results (Breiman 1996; LeBlanc and Tibshirani 1996). Further improvements are possible by restricting coefficients to sum up to one. This will also improve the interpretability, as it may more easily provide a link between coefficients and the importance of a single level-0 model for the ensemble.

Because of the technical difficulties of the 3D-printing models, alternative model combinations, which include several analytical models, can be recommended. Instead of using the five models from our study, we will combine one CFD model with several analytical and surrogate models.

On the basis of the observations in this article, we believe that the novel OMMS approach may be successfully used for many applications in the field of optimization-via-simulation. Last but not least, it should be noted that OMMS cannot be applied without any modifications to every optimization problem. It has to be adapted to the particular application, e.g., cyclone geometry optimization. We do not claim that OMMS outperforms all other algorithms on every problem class. This is a consequence of the no free lunch theorem (Wolpert and Macready 1997; Haftka 2016).

Acknowledgements This work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 692286. We would like to thank Horst Stenzel, Beate Breiderhoff, Dimitri Gusew, Aylin Mengi, Baris Kabacali, Jerome Tünte, Lukas Büscher, Sascha Wüstlich, and Thomas Friesen for their support.

References

Barth W (1956) Berechnung und Auslegung von Zyklonabscheidern aufgrund neuerer Untersuchungen. Brennstoff-Wärme-Kraft 8(1):1–9

Bartz-Beielstein T (2016) Stacked generalization of surrogate models - a practical approach. Technical Report 5/2016, TH Köln, Köln. https://cos.bibl.th-koeln.de/frontdoor/index/index/docId/375

Bartz-Beielstein T, Zaefferer M (2017) Model-based methods for continuous and discrete global optimization. Appl Soft Comput 55:154–167

Bartz-Beielstein T, Lasarczyk C, Preuss M (2005) Sequential parameter optimization. In: McKay B et al (eds) Proceedings 2005 congress on evolutionary computation (CEC'05), Edinburgh, Scotland. IEEE Press, Piscataway, pp 773–780

Barzier MK, Perry CJ (1991) An approach to the construction and usage of simulation modeling in the shipbuilding industry. In: Nelson BL, Kelton WD, Clark GM (eds) 1991 winter simulation conference proceedings. IEEE, pp 455–464

Breiman L (1996) Stacked regression. Mach Learn 24:49–64

Bucila C, Caruana R, Niculescu-Mizil A (2006) Model compression: making big, slow models practical. In: Proceedings of the 12th international conference on knowledge discovery and data Mining (KDD'06)

Chaudhuri A, Haftka RT, Ifju P, Chang K, Tyler C, Schmitz T (2015) Experimental flapping wing optimization and uncertainty quantification using limited samples. Struct Multidiscip Optim 51(4):957–970

Dempster AP (1968) A generalization of bayesian inference. J R Stat Soc Ser B Methodol 30(2):205–247

Elsayed K, Lacor C (2010) Optimization of the cyclone separator geometry for minimum pressure drop using mathematical models and CFD simulations. Chem Eng Sci 65(22):6048–6058

Elsayed K, Lacor C (2012) CFD modeling and multi-objective optimization of cyclone geometry using desirability function, artificial neural networks and genetic algorithms. Appl Math Model 37(8):5680–5704

Fishwick PA, Zeigler BP (1992) A multimodel methodology for qualitative model engineering. ACM Trans Model Comput Simul 2(1):52–81

Forrester A, Sóbester A, Keane A (2007) Multi-fidelity optimization via surrogate modelling. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science 463(2088):3251–3269

Fu MC (1994) Optimization via simulation: a review. Ann Oper Res 53(1):199–247

Goel T, Haftka RT, Shyy W, Queipo NV (2007) Ensemble of surrogates. Struct Multidiscip Optim 33(3):199–216

Haftka RT (2016) Requirements for papers focusing on new or improved global optimization algorithms. Struct Multidiscip Optim 54(1):1–1

Haftka RT, Villanueva D, Chaudhuri A (2016) Parallel surrogate-assisted global optimization with expensive functions—a survey. Struct Multidiscip Optim 54(1):3–13

Hinton G, Vinyals O, Dean J (2015) Distilling the knowledge in a neural Network. arXiv:1503.02531

Hoekstra AJ, Derksen JJ, Van Den Akker HEA (1999) An experimental and numerical study of turbulent swirling flow in gas cyclones. Chem Eng Sci 54(13-14):2055–2065

Hoffmann AC, Stein LE (2007) Gas cyclones and swirl tubes. Springer, Berlin

Jin Y (2003) A comprehensive survey of fitness approximation in evolutionary computation. Soft Comput 9(1):3–12

Jin R, Chen W, Simpson TW (2001) Comparative studies of metamodelling techniques under multiple modelling criteria. Struct Multidiscip Optim 23(1):1–13

Kazemi P, Khalid MH, Szlek J, Mirtič A, Reynolds GK, Jachowicz R, Mendyk A (2016) Computational intelligence modeling of granule size distribution for oscillating milling. Powder Technol 301(Supplement C):1252–1258

JrnlID 158_ArtID 1934_Proof#1 - 02/02/2018

Kleijnen JPC (2008) Design and analysis of simulation experiments. Springer, New York

Kleijnen JPC (2014) Simulation-optimization via Kriging and bootstrapping: a survey. Journal of Simulation 8(4):241–250

Konan A, Huckaby D (2015) Modeling and simulation of a gas-solid cyclone during an upset event (presentation). OpenFOAM Workshop

Law AM (2007) Simulation modeling and analysis, 4th edn. McGraw-Hill, New York

LeBlanc M, Tibshirani R (1996) Combining estiamates in regression and classification. J Am Stat Assoc 91(436):1641

Löffler F (1988) Staubabscheiden. Thieme, Stuttgart

Meerschaert MM (2013) Mathematical modeling (fourth edition), 4th edn. Elsevier, Amsterdam

Mothes H, Loeffler F (1984) Bewegung und Abscheidung der Partikel im Zyklon. Chem-Tech-Ing 56:714–715

Müller J, Shoemaker CA (2014) Influence of ensemble surrogate models and sampling strategy on the solution quality of algorithms for computationally expensive black-box global optimization problems. J Glob Optim 60(2):123–144

Murphy KP (2012) Machine learning: a probabilistic perspective. MIT Press, Cambridge

Muschelknautz E (1972) Die Berechnung von Zyklonabscheidern für Gase. Chemie Ingenieur Technik 44(1-2):63–71

Nelson BL (1995) Stochastic modeling: analysis and simulation. Dover, New York

OpenFOAM Foundation (2016) OpenFOAM tutorials lagrangian MPPICFoam cyclone. Official OpenFOAM repository. https://github.com/OpenFOAM/

Overcamp TJ, Mantha SV (1998) A simple method for estimating cyclone efficiency. Environ Prog 17(2):77–79

Preen R, Bull L (2014) Towards the coevolution of novel vertical-axis wind turbines. IEEE Trans Evol Comput PP(99):284–294

Santner TJ, Williams BJ, Notz WI (2003) The design and analysis of computer experiments. Springer, Berlin

Simpson T, Toropov V, Balabanov V, Viana F (2012) Design and analysis of computer experiments in multidisciplinary design optimization: a review of how far we have come - or not. In: 12th AIAA/ISSMO multidisciplinary analysis and optimization conference. American Institute of Aeronautics and Astronautics, Reston, pp 1–22

Turner AJ, Balestrini-Robinson S, Mavris D (2013) Heuristics for the regression of stochastic simulations. Journal of Simulation 7(4):229–239

Wolpert DH (1992) Stacked generalization. Neural Netw 5(2):241–259

Wolpert DH, Macready WG (1997) No free lunch theorems for optimization. IEEE Trans Evol Comput 1(1):67–82

Yang Y (2003) Regression with multiple candidate models: selecting or mixing? Stat Sin 783–809

Zaefferer M, Breiderhoff B, Naujoks B, Friese M, Stork J, Fischbach A, Flasch O, Bartz-Beielstein T (2014) Tuning multi-objective optimization algorithms for cyclone dust separators. In: Proceedings of the 2014 conference on genetic and evolutionary computation, GECCO '14. ACM, New York, pp 1223–1230

Zeigler BP, Oren TI (1986) Multifaceted, multiparadigm modeling perspectives: tools for the 90's. In: Proceedings of the 18th conference on winter simulation. ACM, New York, pp 708–712

Zerpa LE, Queipo NV, Pintos S, Salager J-L (2005) An optimization methodology of alkaline–surfactant–polymer flooding processes using field scale numerical simulation and multiple surrogates. J Pet Sci Eng 47(3):197–208

Q5