Proceedings World Geothermal Congress 2020+1
Reykjavik, Iceland, April - October 2021

Addressing Geophysical Inverse Problems for Basin-Scale Heat Flow Models by Using a Physics-Based Machine Learning Approach

Denise Degen[1,2], Karen Veroy[2,3], Magdalena Scheck-Wenderoth[4,5] and Florian Wellmann[1,6]

[1]RWTH Aachen University, CGRE, Wüllnerstraße 2, 50262 Aachen

[2]RWTH Aachen University, AICES Graduate School, Schinkelstr. 2, 52062 Aachen, Germany

[3]RWTH Aachen University, High-Performance Computation for Engineered Systems, Faculty of Civil Engineering, Aachen, Germany

[4]GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany

[5]RWTH Aachen University, Faculty of Georesources and Materials Engineering, Aachen, Germany

[6]JARA-HPC, Wilhelm-Johnen-Straße, 52425 Jülich, Germany

denise.degen@cgre.rwth-aachen.de

Keywords: Geothermal, Model Order Reduction, Reduced Basis Method, Uncertainty Quantification

ABSTRACT

In order to determine suitable locations for geothermal exploration, reliable predictions of the earth's subsurface temperature field are essential. For these predictions, it is necessary to consider the uncertainties of the involved parameters. However, standard uncertainty quantification methods, such as Markov Chain Monte Carlo, are computationally intractable for high-resolution simulations of basin-scale models. We thus require numerical methods that considerably accelerate the forward simulations to enable the use of uncertainty quantification approaches that can easily require up to a million forward simulations. For this purpose, we introduce the reduced basis method, a physics-based machine learning approach. Our previous studies show that we obtain speed-ups of four to six orders of magnitude in comparison to standard finite element simulations. One main advantage of the reduced basis method in contrast to other surrogate models is that we obtain temperature values at every point in the model and not only at the observation points. Consequently, we can generate uncertainty maps of the temperatures at the target depth of the geothermal wells for the entire extent of a basin of interest.

We use the Brandenburg (Germany) model to illustrate the application and benefits of the reduced basis method for large-scale geological models. The numerical simulations are realized within the DwarfElephant package, an open-source, high-performance application based on the Multiphysics Object-Oriented Simulation Environment (MOOSE) developed by the Idaho National Laboratory. The DwarfElephant package offers a physics-independent and user-friendly access to the reduced basis method within a high-performance finite element library, allowing computations of spatially high dimensional models. In addition, we present how the method can be used for other inverse processes, such as automated model calibrations. Inverse problems are becoming rapidly extremely expensive computationally even without including all major sources of uncertainty. In that regard, the reduced basis method is very promising because it allows a significant reduction in computation time without introducing additional physical uncertainties.

1. INTRODUCTION

Many methods in the field of geosciences aim at gaining an accurate understanding of the spatial distribution of the earth's subsurface properties and the involved physical processes, such as fluid and heat transport, chemical species transport and mechanical processes. Understanding the subsurface has a significant impact on our society because it is essential for providing meaningful predictions, which can be used to promote sustainable use of valuable subsurface resources.

A meaningful prediction, however, presents us with two main challenges: First, complex coupled processes must be analyzed over a large spatial and temporal domain. Additionally, the highly heterogeneous nature of the subsurface numerically leads to a high dimensional parameter space yielding the so-called "curse of dimensionality". Hence, the encountered tasks are computationally intensive and expensive, thus requiring high-performance computing (HPC) infrastructures. Second, retrieving information from the subsurface is a non-trivial process due to, for instance, limited access to the areas of interest. Consequently, the measurements are associated with high uncertainties.

As a step towards addressing these challenges, we use a model order reduction (MOR) technique, specifically the reduced basis (RB) method, to reduce the dimensionality of the problem. The RB method constructs low-order approximations (i.e. with a small number of degrees of freedom) to, for instance, high-dimensional finite element models. One requirement for constructing these low-order RB approximations is that the problem is affine parameter-dependent; this requirement can, however, be readily relaxed. Taking advantage of this decomposition, the RB method results in "offline" and "online" phases. In the offline phase, the expensive pre-computations (that involve solving multiple high-dimensional finite element simulation problems), which are required to build low-order models, can be performed on HPC infrastructures. In the online stage, the RB low-order model is then used as a computationally cheap surrogate for the computationally expensive high-fidelity finite element solution. The faster RB simulations allow in turn to perform inversions, uncertainty analyses, to perform more complex analyses, or to obtain results in real-time. Contrary, to other MOR techniques, the RB method has the advantage of providing an error bound for elliptic and parabolic partial differential equations (PDEs), allowing the assessment of the quality of the approximation as well as the construction of efficient reduced-order models (Hesthaven et al. 2016, Prud'homme et al. 2002).

However, constructing parallel computer code for HPC infrastructures is often not trivial. We, therefore, present a software implementation within the Multiphysics Object-Orientated Simulation Environment (MOOSE) primarily developed by Idaho National Laboratory (Alger et al. 2019), offering a built-in parallelization on top of libMesh (Kirk et al. 2006) and PETSc (Balay et al. 2019). The presented RB implementation is mainly based on the rbOOmit package provided by libMesh (Knezevic and Peterson 2011).

We demonstrate the benefits of the RB method for basin-scale heat flow applications by using the case study of Brandenburg. We will show that for these kinds of applications we obtain a speed-up of several orders of magnitude for highly accurate approximations. In contrast to many other surrogate models, we can retrieve the entire temperature field, allowing the generation of uncertainty quantification maps as shown in this paper.

2. MATERIALS AND METHODS

In the following section, we briefly introduce the geological model, the governing equations and the numerical methods used throughout this paper.

2.1 The Brandenburg Model

We use a combination of the Brandenburg model (Fig. 1) presented in Noack et al. (2012, 2013) with a steady-state geothermal heat flow problem. At the top and bottom of the model, we apply Dirichlet boundary conditions of 8˚C and 1300˚C, respectively. In order to account for errors in the geometrical parameterization of the lithosphere-asthenosphere boundary (LAB), we allow the temperature of the lower boundary condition to vary by 10%. For the thermal conductivities, we allow a variation of 50%.

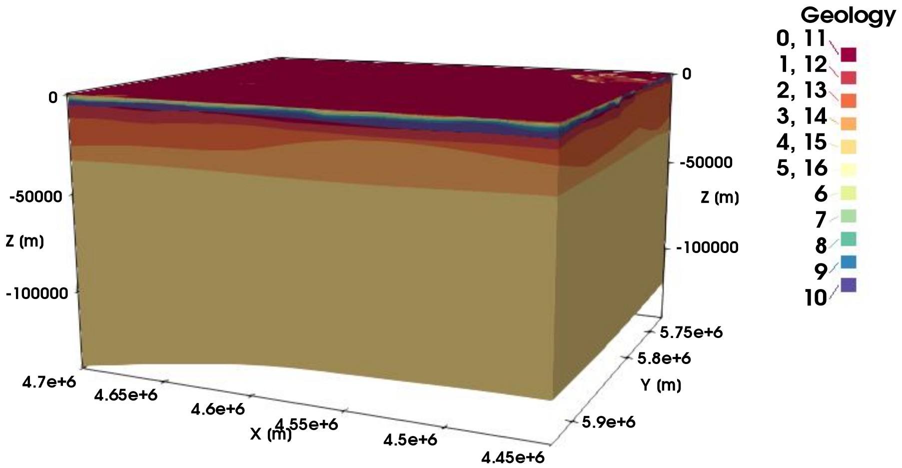

Figure 1: Geology of the Brandenburg model. 0: Quaternary, 1: Tertiary-post-Rupelian, 2: Tertiary Rupelian-clay, 3: Tertiary-pre-Rupelian-clay, 4: Upper Cretaceous, 5: Lower Cretaceous, 6: Jurassic, 7: Keuper, 8: Muschelkalk, 9: Buntsandstein, 10: Zechstein, 11: Sedimentary Rotliegend, 12: Permo-Carboniferous Volcanics, 13: Pre-permian, 14: Upper Crust, 15: Lower Crust, 16: Lithospheric mantle.

2.2 The Physical Model

For the forward simulations, we take a geothermal conduction problem with the radiogenic heat production *S* as the source term (Bayer et al. (1997)):

$$-\lambda \nabla^2 T + S = 0. \tag{1}$$

Here, λ is the thermal conductivity, and *T* is temperature. For efficiency, we non-dimensionalize the equation, which leads to the following equation:

$$-\frac{\lambda}{\lambda_{ref}\, S_{ref}} \nabla^2 \left(\frac{T-T_{ref}}{T_{ref}}\right) + \frac{S\, l_{ref}^2}{S_{ref}\, T_{ref}\, \lambda_{ref}} = 0. \tag{2}$$

Here, λ_{ref} is the reference thermal conductivity, which is set equal to the maximum thermal conductivity of the Brandenburg model of 3.95 $Wm^{-1}K^{-1}$. The reference temperature T_{ref} is set to 1300˚C (the temperature at the LAB), and the reference radiogenic heat production S_{ref} is set to the maximum radiogenic heat production of the Brandenburg model of 2.5 μWm^3.

2.3 The Reduced Basis Method

The high-dimensionality of the presented problems is tackled by relying on the reduced basis (RB) method, which allows obtaining simulations results in real-time, due to a drastically reduced model dimensionality and complexity. We decided to use the RB method because, in contrast to other surrogate models, it has the advantage that it allows to obtain temperature values everywhere within the model and not only at predefined points. Additionally, it provides error bounds for elliptic and parabolic PDEs, allowing the assessment of the quality of the approximation as well as the construction of efficient reduced order models. The RB method is decomposed into an offline and online stage. Where, the offline-stage compromises all expensive pre-computations (cost depends primarily on the high-dimensional finite element problem), and the offline stage only has to be carried out once. The obtained reduced basis can afterwards be used as a surrogate model when carrying out inverse modeling.

For implementing the offline-online procedure the problem has to be reformulated into its integral form. First, we decompose the stiffness matrix into:

$$a(w, v; \lambda) = \sum_{q=0}^{13} \lambda_q \int_\Omega \nabla w \, \nabla v \, d\Omega \quad \forall v, w \in X, \ \forall \in \lambda \, D, \tag{3}$$

where, w is the trial function, v the test function, λ_q is the thermal conductivity corresponding to the different geological layers, X is the function space ($H_0^1(\Omega) \subset X \subset H_1(\Omega)$), Ω denotes the spatial domain in $\mathbb{R}^3$, and D is the parameter domain in $\mathbb{R}^{13}$. Afterwards, we decompose the load vector into:

$$f(v; \lambda, s) = \sum_{q=0}^{13} \lambda_q s \int_\Gamma \nabla v \, g(x, y, z) \, d\Gamma + s \int_\Gamma \nabla v \, S \, d\Gamma \quad \forall v, w \in X, \ \forall \in \lambda \, D, \tag{4}$$

$$g(x, y, z) = T_{top} \frac{h(x,y,z) - z_{bottom}(x,y)}{d(x,y)}. \tag{5}$$

Here, Γ denotes the boundary in $\mathbb{R}^2$, s is the scaling parameter for the lower boundary condition, $g(x,y,z)$ is the lifting function, T_{top} is the temperature at the top of the model, T_{bottom} is the temperature at the bottom of the model, $h(x,y,z)$ a location in the model, $z_{bottom}(x,y)$ is the depth of the model, and $d(x,y)$ is the distance between the bottom and top surfaces.

3. RESULTS

We use Markov Chain Monte Carlo (MCMC) as our uncertainty quantification method. For the analysis, we rely on the MCMC implementation of PyMC with a Metropolis sampling. Prior, to the uncertainty quantification we reduce the number of parameters that need to be considered as uncertain. First, we combine layers of equal thermal conductivity into one. Furthermore, we fix the radiogenic heat production since the influence of the radiogenic heat production is minor in comparison to the thermal conductivity. This reduces the number of parameters to 14. Afterwards, we perform a global sensitivity analysis to identify the model parameters that notably influence the model response. Through the global sensitivity analysis, we are able to reduce the number of parameters from 14 to six. For all thermal conductivities in the MCMC algorithm we allow a variation of ±50%. The number of iterations for the MCMC run is set to 1,000,000 with a thinning of 1,000 and 10,000 burn-in-simulations. We do not have any prior information about the uncertainties of the parameters; therefore, we use uniform priors. The proposal standard deviation is set to 10 for all parameters. The temperature data used for the validation consists of 81 temperature measurements from 44 wells in the area of Brandenburg. The formation temperatures have been measured at various depth levels and can be found in Noack (2012, 2013). Since the correction might not fully capture the perturbation of the temperature field, we apply a standard deviation of 0.02 for the observation data. The highest posterior standard deviations are about 67˚C at a depth of 30 to 35 km. We observe very low uncertainties at the top and bottom boundary of our model domain and increasing uncertainties towards the center of the model (Fig. 2).

We observe for the posterior mean temperatures (Fig. 3) the influence of the salt structures in the northern part of Brandenburg. South of Potsdam, we have an interface between the high posterior mean temperatures and significantly lower posterior mean temperature values. The highest posterior standard deviation is at the target depth of 5 km and southeast of the region majorly influenced by the salt structures (Fig. 4). Note that at our target depth the uncertainties range from 10 to 20˚C, which is extremely low.

4. DISCUSSION

We observe decreasing uncertainties towards the boundaries and increasing uncertainties towards the center part of the model. The uncertainties are lower at the boundaries because we defined Dirichlet boundary conditions, and hence the temperature is fixed at the boundaries. The highest uncertainties are observed at a depth of 30 km to 35 km. Note that the LAB, although at a depth of 100 to 140 km influences the model up to a depth of 60 km. This is not critical since the target depth is at 5 km. However, this shows how important it is to place the boundaries as far away from the area of interest as possible. Furthermore, the upper boundary influences the model down to a depth of 10 km. Consequently, the target depth is significantly influenced by the upper boundary condition, which is manifested in the low standard deviations (Fig. 4). The high uncertainties adjacent to the salt area is caused by the high contrast in thermal conductivity between the Zechstein and its neighboring layers.

For the generation of the uncertainty quantification (UQ) maps, we performed, as an initial step, a sensitivity analysis with 300,000 function evaluations yielding an execution time of less than 48 minutes and an MCMC algorithm with 1,000,000 function evaluations requiring 2.3 hours. Postprocessing of the results for the UQ maps took 4 minutes.

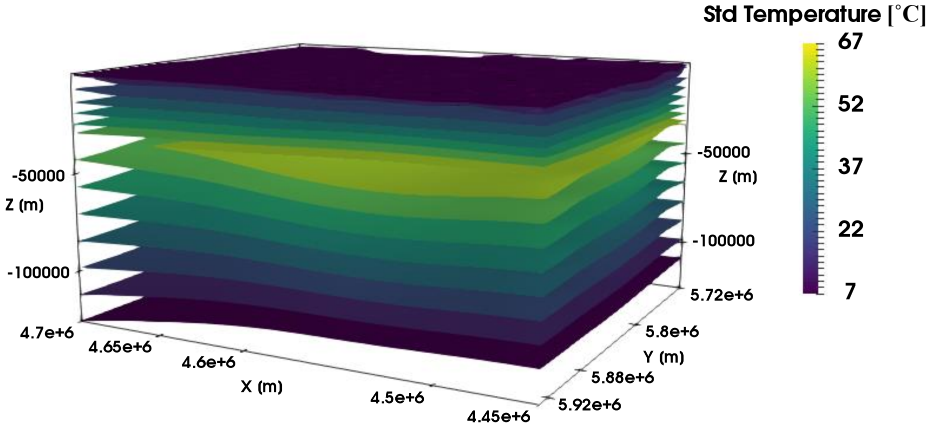

Figure 2: Posterior standard deviations of the temperatures within the entire Brandenburg model. The parameter distributions used for the generation of this map were generated by performing an MCMC analysis. The image is meant to provide a general impression of the uncertainties within the entire model. A detailed image of the uncertainties in the area of interest is provided in Figure 3. A discussion of the uncertainties at a depth of 1 km to 4 km is beyond the scope of this paper.

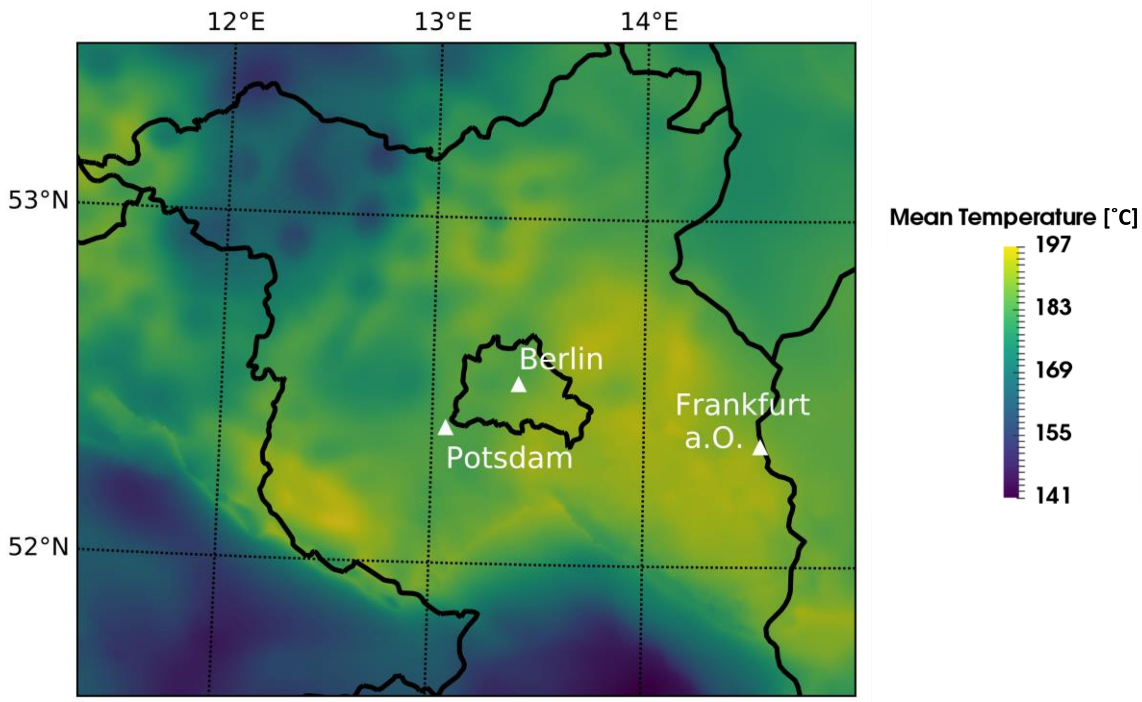

Figure 3: Posterior mean temperatures of the Brandenburg model at a depth of 5 km. The parameter distributions used for the generation of this map were generated by performing an MCMC analysis.

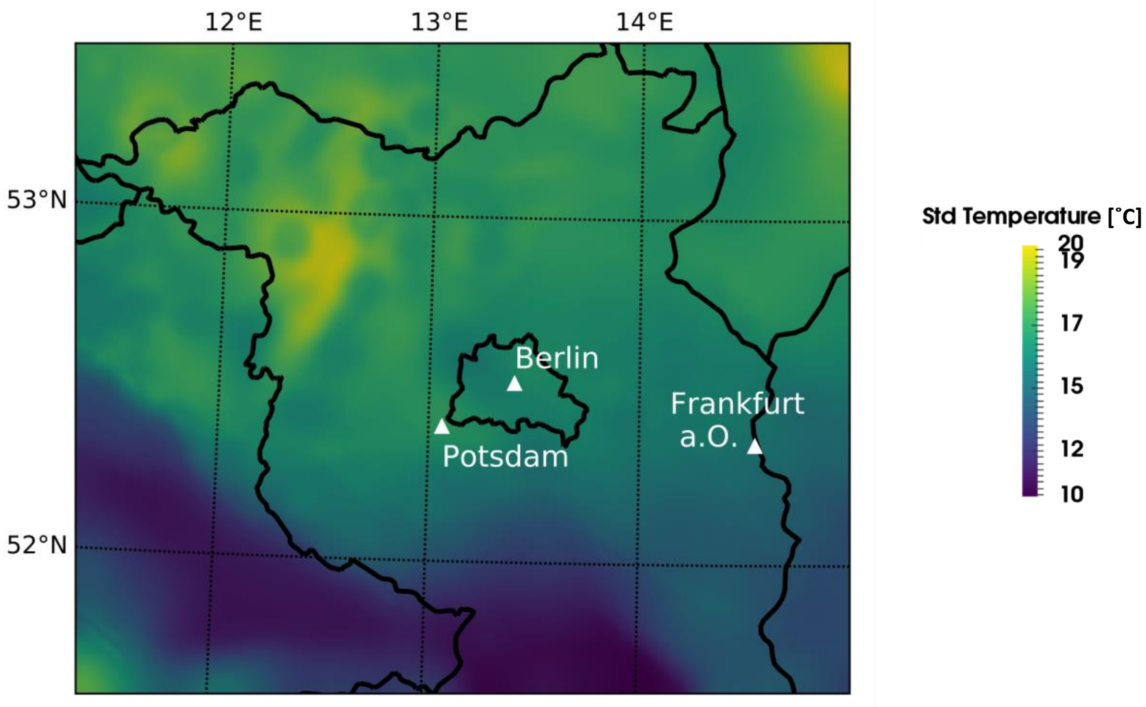

Figure 4: Posterior standard deviation of the temperatures of the Brandenburg model at a depth of 5 km. The parameter distributions used for the generation of this map were generated by performing an MCMC analysis.

5. CONCLUSION

We presented the generation of UQ maps for basin-scale heat flow models, involving a Sobol sensitivity analysis and an MCMC algorithm. The process required less than 2.5 core-hours for the whole analyses, whereas the analyses, using a state-of-the-art finite element solver, would have taken 16 core-years.

It would be interesting to investigate further the influence of the upper boundary condition on the model, by incorporating, for instance, climate interactions since we demonstrated that the upper boundary condition is significantly influencing the temperatures at the target depth.

6. ACKNOWLEDGEMENTS

We like to acknowledge Dr. Vera Noack (Federal Institute for Geosciences and Natural Resources (BGR)), for providing and generating the Brandenburg model. Furthermore, we would like to acknowledge the funding provided by the DFG through DFG Project GSC111. The temperature data used in this presentation is available in Noack (2012, 2013).

REFERENCES

Alger, B., Andrš, D., Carlsen, R. W., Gaston, D. R., Kong, F., Lindsay, A. D., Miller, J. M., Permann, C. J., Peterson, J. W., Slaughter, A. E., & Stogner, R. (2019). MOOSE Web page. https://mooseframework.org.

Balay, S., Abhyankar, S., Adams, M. F., Brown, J., Brune, P., Buschelman, K., Dalcin, L., Dener, A., Eijkhout, V., Gropp, W. D., Karpeyev, D., Kaushik, D., Knepley, M. G., May, D. A., McInnes, L. C., Mills, R. T., Munson, T., Rupp, K., Sanan, P., Smith, B. F., Zampini, S., Zhang, H., & Zhang, H. (2019). PETSc Webpage, http://www.mcs.anl.gov/petsc.

Bayer, U., Scheck, M., & Koehler, M. (1997). Modeling of the 3D thermal field in the northeast German basin. Geologische Rundschau, 86(2), 241–251.

Degen, D., Veroy, K., & Wellmann, F. (2019). Certified Reduced Basis Method in Geosciences - Addressing the challenge of high dimensional problems. EarthArXiv. doi:10.31223/osf.io/neh9j.

Hesthaven, J. S., Rozza, G., & Stamm, B. (2016). Certified reduced basis methods for parametrized partial differential equations. SpringerBriefs in Mathematics, Springer.

Kirk, B. S., Peterson, J. W., Stogner, R. H., & Carey, G. F. (2006). libMesh: a C++ library for parallel adaptive mesh refinement/coarsening simulations. Engineering with Computers, 22(3-4), 237-254.

Knezevic, D. J., & Peterson, J. W. (2011). A high-performance parallel implementation of the certified reduced basis method. Computer Methods in Applied Mechanics and Engineering, 200(13-16), 1455-1466.

Noack, V., Scheck-Wenderoth, M., & Cacace, M. (2012). Sensitivity of 3D thermal models to the choice of boundary conditions and thermal properties: a case study for the area of Brandenburg (NE German Basin). Environmental Earth Sciences, 67(6), 1695-1711.

Noack, V., Scheck-Wenderoth, M., Cacace, M., & Schneider, M. (2013). Influence of fluid flow on the regional thermal field: results from 3D numerical modelling for the area of Brandenburg (North German Basin). Environmental earth sciences, 70(8), 3523-3544.

Prud'homme, C., Rovas, D. V., Veroy, K., Machiels, L., Maday, Y., Patera, A. T., & Turinici, G. (2002). Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. Journal of Fluids Engineering 124(1), 70-80.