Dear Mr. James Kelly,

This is a supporting letter to the first round of review to our manuscript entitled ""Perspectives of Physics-Based Machine Learning for Geoscientific Applications Governed by Partial Differential Equations" by Denise Degen, Daniel Caviedes Voullième, Susanne Buiter, Harrie-Jan Hendricks Franssen, Harry Vereecken, Ana González-Nicolás, and Florian Wellmann to be submitted to the GMD journal.

We have received the comments from two reviewers on our original manuscript. Despite their positive feedback on the quality, novelty and scientific merit of our study, they highlight some points which we have carefully revised in the other attachments to this letter. The main concerns were on only presenting two physics-based machine learning methods and on using only one method for the numerical examples. This said, we originally included only these two methods since they demonstrated two end members on how to integrate physics into machine learning approaches. We want to highlight how the differences impact the potential of the individual methods in context of applications dominated by partial differential equations. We clarified this aspect in the manuscript and added additional physics-based machine learning methods, explaining how they relate with respect to the presented end-members. Furthermore, we added reference to other overview papers comparing both data-driven and physics-based machine learning methods, which in contrast to our manuscript focus on data-dominated applications, having a significant impact on the perspectives of the presented methods. To address, the second point, we extended the numerical examples by constructing the surrogate models additionally via a neural network providing a comparison between physics-based and data-driven machine learning methods.

All other points raised have been carefully answered and we are grateful to the two reviewers for their criticisms which indeed help to raise the quality standard of our publication.

The file system uploaded consists of:

1. Reply to Reviewers.pdf --> document that provides our answers to each point raised by the two reviewers.
2. PhysicsBasedMachineLearning_rev.pdf --> revised manuscript with tracked list of changes made during the round of review.

We hope that you would find our answered to their comments and the revised version of our manuscript sound and improved to grant its submission to the journal, and we are ready to answer to any other questions they should arise.

On behalf of all authors,
Denise Degen

Dear reviewers,

Thank you for taking the time to read our manuscript and for providing feedback. We have revised the manuscript based on your suggestions. For your convenience, changes from the previous version are marked in red. The answers to your individual remarks are inline below.

Reply to Reviewer 1:

This paper provides an overview of physics-based ML method, mainly non-intrusive reduced basis method, for geoscientific problems and presents a workflow to implement physics-based methods. Authors provide geothermal example, geodynamic example, and hydrological example as three benchmarks to validate that this type of method has potential to solve general challenges in geoscience.
The paper is supposed to be a "perspective" paper, but the paper doesn't provide sufficient overview of the field and existing methods. The numerical examples are also only limited to one method.

- Thank you very much for your comments regarding our manuscript.

  The idea behind the paper is to present the potential of physics-based machine learning for geoscientific applications governed by partial differential equations. Narrowing down the possible application field (as presented through the three examples) is necessary since the evaluation of the different physics-based machine learning methods strongly depends on the amount of available data and physical knowledge.
  We on purpose restrict our paper to physics-based machine learning since for data-driven machine learning several review papers exist. Within the field of physics-based machine learning, we chose to present physics-informed neural networks (PINNs) and the non-intrusive reduced basis (NIRB) method since they represent two end members of the spectrum found in physics-based machine learning. PINNs originated from the field of machine learning, treating physics only as a constraint, whereas NIRB originates from the field of applied mathematics aiming at building the model with only the physics in mind. PINNs are therefore ideal in a situation where some data is available and some physical knowledge. In our case, we have a situation where data is extremely sparse, and a lot of physical knowledge is available. Hence, methods such as NIRB are preferential and thus also the choice to perform the benchmarks for the NIRB method.
  We realize that these points might not have become clear in the manuscript. Therefore, we highlighted these aspects in the revised manuscript in line 213-219.

  "These various approaches are already compared in several papers (e.g., Faroughi et al., 2022; Swischuk et al., 2019; Willard et al., 2020). However, they focus on applications were substantially more data is available than in most subsurface applications. Therefore, we shift the focus to application with very sparse data sets. This impacts the potential of physics-based machine learning methods differently, depending on the paradigm used to combine physics- and data-based methods. So, instead of presenting numerous of these techniques in detail, as already done in the

afore mentioned papers, we want to discuss the different paradigms that exist. Here, we identify two end-member cases: i) physics-guided loss functions, and ii) hybrid models."

Furthermore, we list other physics-based machine learning methods and provide further references to them in line 234-239 and explain how they compare to the two presented endmembers in Section 2.3.3, to highlight that many different physics-based machine learning methods are existing.

"Note that both methods act as examples to better illustrate the different concepts used for combining physics- and data-based approaches. Numerous other methods such as the Physics-Encoded Neural Network (PeNN) (Chen et al., 2018a; Li et al.,2020; Rao et al., 2021), Physics-Encoded Recurrent-Convolutional Neural Network (PeRCNN) (Rao et al., 2021), the Fourier Neural Operator (FNO) (Li et al., 2020), and DeepONets (Lu et al., 2021). We will not discuss in detail but shortly explain their relations to the two paradigms, represented by PINNs and the non-intrusive RB method."

We also provide references to review papers on data-driven methodologies in line 200-201.

"For an overview of data-driven machine learning techniques, we refer to Jordan and Mitchell (2015); Kotsiantis et al. (2007); Mahesh (2020)."

Additionally, we slightly adapted the title including the world "Strategies" to better reflect the intentions of the manuscript.
"*Perspectives of Physics-Based Machine* ***Strategies*** *for Geoscientific Applications Governed by Partial Differential Equations*"

Major comments:
- Section 4 "Challenges" does not provide enough information, but rather repeat those already mentioned in introduction. It is also not clear how these challenges will be solved by new methods.
- There is a similar issue in conclusion section.

- In the revised manuscript, we clarify how the described methods can be used to solve the presented challenges at the end of the corresponding challenge sections (Section 4.1 – 4.3).
- Challenge 1: Sensitivity Analysis
  "Consequently, we can address the challenge of performing computationally expensive global sensitivity analysis by using surrogate models constructed by the non-intrusive RB method. This reduces the computational cost by several orders of magnitude making it a suitable technique to ensure the feasibility of global sensitivity studies also for large-scale models (Degen et al., 2021b; Degen and Cacace, 2021; Degen et al., 2021a, 2022b)."
- Challenge 2: Uncertainty
  "To conclude, as for the sensitivity analyses the computational challenge is addressed by using surrogate models. Note that as before methods that map from model parameters to state information are better suitable within an uncertainty quantification framework."

- Challenge 3: Real-Time
  "Consequently, the challenge of producing real-time predictions is addressable by the use of surrogate models. We generally see greater potential in methods that follow the conceptual ideas of the non-intrusive RB method since they allow for both parameter and state estimation. Nonetheless, in the context of state estimation and in situations where some physical knowledge is available, methods that map from state-to-state information (such as PINNs) are advantageous since they allow also a direct incorporation of the measurement data. However, these are not the target applications of this manuscript."

For benchmark examples, the objective of learning and test metric are not clearly pointed out. For example, the input and output of the non-intrusive RB method and the target to learn should be emphasized in the text.

- We clarified the objectives of the method on lines 326-332.
  In general, the NIRB method consists of two steps.
  For the proper orthogonal decomposition, we have a set of snapshots as inputs (meaning full dimensional simulations for specific material properties and time steps) and obtain eigenvectors, which are our basis functions.
  For the machine learning method (in our case either a neural network or a Gaussian process regression), which follows after the proper orthogonal decomposition, we have as our input the material properties. As the training data, we have the matrix product of the snapshots and the basis functions. Hence, we obtain the weighting factors for our above determined basis functions.

- It's better to list computational costs for traditional methods and new methods to have a clear comparison as this paper focuses on speed-up of classical methods.
  - So far, we only presented the computational cost of the classical methods and the new methods in the text of the respective examples. Additionally, we provide a Table 5 in Section 3.5 listing and comparing the computational cost for a better illustration.

- Some paragraphs and sentences are hard to read and need to be revised.
  - We revised the text parts.

Minor comments:

- Line 767-769 repetitive sentence.
  - The sentence has been removed.
- Format of equation should be consistent, such as all centered (equation 11 is on the left).
  - The formatting has been adjusted.

Reply to Reviewer 2:

In this article, the authors discuss the potential of combining physics-based and data-driven methods for geophysical applications. One physics-based machine learning method, namely the non-intrusive reduced basis method, is highlighted and tested in three case studies.

The manuscript is overall rather well written. However, for a perspective article I am a bit surprised that so few methods are presented and only one is tested, especially since in my opinion the manuscript is very long.

- Thank you very much for your comments regarding our manuscript and taking the time to review the manuscript. We presented only two physics-based machine learning methods since they represent two end member cases. However, we see that this might not have been apparent in the manuscript. Therefore, we extended the explanation (see lines 213-219) and added an additional Section 2.3.3. describing other physics-based machine learning methods and explaining how they behave with respect to these end members.
- Additionally, we slightly adapted the title including the world "Strategies" to better reflect the intentions of the manuscript.
  "*Perspectives of Physics-Based Machine* ***Strategies*** *for Geoscientific Applications Governed by Partial Differential Equations*"

General comments

1) In general, I think that clear statements of the objectives are missing in section 2. From what I understand, we are looking for a surrogate model, but for what kind of applications: forecast? inversions? parameter estimation? other? In my opinion this is very important as what works for an application might not work for another. The requirements for the surrogate model may also differ from one application to the other.

- We focus mainly on parameter estimation, uncertainty quantification, and global sensitivity analysis. This has been specified in lines 99-101 (Section 2) as its influences the requirements for the surrogate models.

2) I am under the impression that I have a different definition of machine learning (or data-driven) methods than the one used in the present manuscript. From what I know, machine learning methods are methods where a problem is solved without being explicitly programmed, using a large dataset. By contrast, with physics-based methods the same problem is solved using a model derived from physical laws. Consequently, methods such as POD or RB, which only require data to work and which are independent of the physics, are typically machine learning methods, and I do not think they can be classified as physics-based. This needs to be clarified.

- Both the POD and the intrusive RB method, described in Section 2.1., use a Galerkin projection which requires access to the discretized version of the PDE in form of the stiffness matrix and load vector and are not independent of the physics. However, the POD without the combination of a Galerkin projection can also be used independent of the physics, where it would be right to no longer classify the method as physics-based. To clarify this, additional explanations have been added in line 134-139 and 150-151.

3) In the test series, I understand that the data comes from simulations. However, in figure 1 we see that in some cases the data comes from measurements. Can we have more details about these measurements? Furthermore in this figure, physics-based methods are used with measurements only. This is weird as in the description in section 2.1, "physics-based" techniques such as POD or RB are built using simulated data.

- We added a clarification of what is meant by measured data in the caption of Figure 1. For the explanation of POD and RB, we refer to the previous comment.

4) Here and there, for example L253, it is mentioned that data-driven models do not preserve the input-output relationships. This statement seems a bit bold since the input and output of the problems are not properly defined. I would highly recommend to explicitly define them.

- The statement originates from the Model Order Reduction community and in the original manuscript, we did not realize that input and output are naturally differently defined depending on the community. We clarified, the definition of input and output throughout the manuscript.

5) In the numerical illustrations, only one method is tested, in such a way that we have no idea of how good the performances are. I would really advise to add a baseline method for comparison.

- To offer the comparison to a base line method, we constructed the surrogate models also via a neural-network and detailed the implication in the newly introduced Section 3.5.

6) In the numerical experiments, only one initial condition is used. Readers from communities such as meteorology might be surprised by this choice. Perhaps it would be interesting to discuss the sensitivity to initial conditions and to compare it to the sensitivity to other parameters.

- In Section 3.5 (line 669-673), we described how different initial and boundary conditions can be considered.

Specific remarks

L80: "physically consistent" What is meant exactly here? Physically consistent can have many different meanings.

- We specified the meaning of physically consistent (see lines 80-82).

L 89/90: "It is critical that the solutions produced by the surrogate models do not change the general behavior if, for instance, the accuracy of the models is increased." I do not understand this sentence.

- The sentence has been revised.

L98: "Physics-driven approaches preserve the governing equations," What does it mean to "preserve an equation"?

- We clarified the meaning of this statement.

“Physics-driven approaches preserve the governing equations, meaning that they, for instance, conserve mass, momentum, and energy in the same way as the original discretized version does. They also maintain the characteristic that for a given set of model parameters (e.g., material properties), they produce information about the state variables (e.g., temperature, pressure) at, for instance, every node of the discretized model.”

L 131: "The RB method is fairly unknown in the field of Geosciences." Please tune down this subjective statement.

- We removed the subjective statement.

L 134/135: "Alternatively, also a POD can be used for the sampling instead of the greedy algorithm." In that case, RB would be the same as POD, right?

- Unfortunately, the POD method with a Galerkin projection is in literature sometimes called RB with POD sampling, POD or POD + Galerkin. A note of caution is found in line 150 -151. We removed the statement from line 134/135 because it is an inconsistency in the definitions available in literature but does not impact the results presented in this manuscript. By removing the statement and only leaving the note of caution, we aim to avoid further confusions.