In blue are our answers to the reviewer reports based on the revised manuscript (first revision).

Report #1
Submitted on 09 May 2025
Anonymous referee #1
accepted subject to technical corrections

Overall, I find the response to my comments acceptable, and I value the provided clarification. I have some suggestions that might enhance the manuscript's readability:

Thanks for taking the time to assess the revised manuscript and for your positive and constructive comments.

Table 1. Rounding the numbers to 2 decimals looks better in this case.

We agree that rounding the values improves the readability of the table. We have therefore updated Table 1 so that all values are now rounded to two decimal places.

Line 100: Could you elaborate on the necessity of having duplicated edges in this situation?

As mentioned in line 102, we chose to work with a directed graph in order to capture the predominant direction of the flow. Without this directionality, paths in the graph could propagate information backward, which would not be physically meaningful in our context. Duplicated edges (i.e., edges in both directions) allow the model to learn distinct behaviors depending on the flow direction.

Line 133: Is this meant to be Re or we from Eq 4?

Thank you for pointing this out. You are right, this should be we. We have corrected the manuscript accordingly.

Line 157: Consider moving the reference to the figures (Fig. 3 (a) and (b)) between "mass map _ over the 80 cases"

Thank you for the suggestion. We have followed your advice and moved the reference to Fig. 3 (a) and (b) accordingly in the revised manuscript.

Figures 3a and 3b do not include units of dimension for the quantities being plotted. For clarity, consider including the coordinate x of the plane being presented, as it is the first time you introduce the points at which the approach is measuring the similarity index.

Thank you for your comment. The axes in Figures 3(a) and 3(b) are given in $y/\Delta y$ and $z/\Delta z$ which are dimensionless coordinates corresponding to the number of discretization steps in the y and z-directions. As stated in lines 79 and 135, these maps are plotted at the outlet of the domain (the last layer in the x-direction), located at x=7000 m. We have clarified this more explicitly in the manuscript and updated the figure caption accordingly to avoid any ambiguity.

Figure 7: Specify that this is comparing the 'cumulative mass' of scenario 65 with 'distances' from all scenarios in S5

Yes, we agree that clarifying the caption would help the reader better understand the content of Figure 7. We have revised it accordingly.

Discussion:
Line 355: Do the authors have any insights into what is considered a strong difference in the conductivity of the medium? As part of acknowledging the limitations of the proposed approach, is there any sensitivity regarding when the model begins to lean more towards the heterogeneous-no-faults case?

> Here, the strong hydraulic conductivity contrasts of two orders of magnitude (or factor 100) is controlled by the definition of the faults. We did not perform a sensitivity analysis of the dividing/multiplying factor on the performance of the proxy.

Figures:
Figure captions need to be more descriptive to align with the author's interpretation. Some figures have detailed captions, while others lack sufficient information. Although there's a description in the main text, the figures should be self-explanatory.

> Thank you for this valuable comment. We have revised and expanded the captions of several figures to make them more descriptive and self-contained, better reflecting the interpretation provided in the main text. In particular, we updated the captions of Figures 1, 2, and 4 to ensure that they are informative and understandable independently from the main manuscript.

Report #2
Submitted on 13 May 2025
Anonymous referee #3
reconsidered after major revisions

GENERAL COMMENTS
This manuscript presents an extension of the work by Rizzo and de Barros (2017). The developed surrogate model is tested on an example, where the goal is to determine the permeability character of three faults in a domain for which the conductivity field, the position of the sources and the position of the faults is known.

I was not one of the reviewers of the original submission, and, therefore, I will add some comments to the remarks of the colleagues who reviewed the original submission.

The manuscript is well organized and written with a good language.

However, in my opinion the innovative content of the work is quite limited and the test used to validate the surrogate model is excessively simple.

I think that the work suffers from some scientific weaknesses, that are listed in the specific comments below.

Overall, I think that the manuscript cannot be considered for publication in its present form, but it requires a major revision.

Thanks for providing a review on this revised manuscript. All comments have been carefully addressed point by point.

SPECIFIC COMMENTS

1) Lines 2 & 3. I found this sentence rather obscure and, if I understood it properly, I do not fully agree with the concept it conveys. More generally, at the end of the abstract, I had the idea that the work proposes a surrogate model, based on graph theory, rather than a physically-based model, in order to exclude some hydrostratigraphic or hydrogeological scenarios. However, I am not sure that I got the right point.

The beginning of the sentence line 2 has been reformulated to clarify its meaning. You are correct in your understanding of our objectives.

2) Lines 120 to 124. This remark is intriguing, but I do not understand the physical motivation. Moreover, if I understood correctly, equation (4) implies that the local flow direction is always the same as the main direction of flow. In the case of a highly heterogeneous medium, the geometry of low permeability structures could yield a local flow with a direction opposite to the average flow. Therefore, I am afraid that this formula could not be optimal for some conditions.

As explained on line 100, the graph is similar to a non-oriented graph. Depending on weights, a shortest path might have directional components opposite to the main flow direction, which is unrealistic given the steady state boundary conditions. Equation (4) implies that the local flow direction (in the graph) can drift with an angle bounded by the]-90,90[interval. An angle of 0 would mean a local flow direction aligned perfectly with the main flow direction.

3) Section 2.3 Observation time. I think that this discussion is not well developed. In particular, the surrogate model proposed here gives a steady-state picture of flow, which is then used to estimate transient solute transport, without solving the ADE. So, the problem of the selection of the observation time should be analysed very carefully.

This is what we acknowledge in section 2.3. We already explain and justify our choices. A sensitivity of the retained observation time is not the scope of this work.

4) Section 2.4 Metrics. The proper selection of metrics for this comparison is a well known problem, for which, to my knowledge, there is no definite solution. the authors provide an interesting discussion, but I would appreciate if they can support their remarks with some numerical tests.

The use of pixels ranked by distances / cumulative mass is supported by the quantitative correlation analysis (Fig. 3). The arguments advanced to build the proposed metric are supported by the literature cited in the section, so we do not deem necessary to add a numerical example.

5) Lines 224 to 229. The authors assume that the conductivity field and the position of the faults (this is not explicitly stated in the text, but I deduced it from the rest of the description here) are known. Then the problem is restricted to the determination of the permeability o the faults. This problem could be solved quite simply with few runs of a numerical, physically-based model. Indeed, the problem is often that scarce data do not allow to map preferential flow paths (permeable faults, fractures, permeable sediments, etc. for different types of aquifers) and their connectivity.

Indeed, we try to identify if faults act as barrier or pathway. The challenge with physically based models is their high computing requirements (to solve the transport - ADE). This is why we are interested to see how good a (faster) proxy can do, as it provides the opportunity to test easily more scenarios.

6) In several steps of the work, thresholds are introduced. As it is common in such approaches, the selection of thresholds is often a subjective operation. A thorough sensitivity analysis with respect to the applied thresholds, extending what is found in Appendix B, would be necessary.

Section 3.3 already provides a sensitivity analysis of the success rate (for scenario selection) as a function of the threshold lambda (applied to the similarity metric).

7) Throughout the whole work, MODFLOW is mentioned as the model applied to generate synthetic data. However, to my knowledge, MODFLOW solves the flow equations, whereas MT3D or other models are used in cascade to model transport. Is this the case? I think this should be specified in a better way.

On their webpage, https://www.usgs.gov/software/modflow-6-usgs-modular-hydrologic-model the USGS presents MODFLOW 6 as a container for Groundwater Flow (GWF) and Groundwater Transport (GWT) models. We have update the term MODFLOW to MODFLOW 6 throughout the manuscript.

TECHNICAL COMMENTS

1) Line 2. What is "it"? the subject of the previous sentence is "groundwater contaminant transport problems".

The pronoun "it" in line 2 refers to the computational cost associated with groundwater contaminant transport problems, which limits the exploration of conceptual uncertainty. To clarify this, we have rephrased the sentence in the revised manuscript.

2) Lines 14 to 16. Sentence "The study of... heterogeneous environments" is quite tautological. Sentence "Contamination... mitigation strategies" is quite generic and it ends with a couple of citations that do not seem to be the most relevant to support this statement.

We acknowledge that the original sentences were somewhat tautological and generic. We have revised this passage to improve clarity and precision, and we have removed the first reference while keeping the second one. In addition, we have added a relevant reference (Bear and Cheng, 2010) that specifically addresses the modeling of contaminant transport in groundwater systems to better support this context.

3) Line 40. "CO2" should be corrected.

We have corrected "CO2" as requested.

4) Line 48. Expression "multi-heterogeneous-layer" should be rephrased.

We are keeping the original formulation. It is a bit tedious, but it states what it represents: a faulted medium composed of multiple layers, each of them being heterogeneous.

5) Figure 1. The three geological units mentioned at line 67 are not represented. It would be useful to provide the length of the domain along the x and y directions, otherwise using the scale length remains quite imprecise. These values are given at line 79; they could be given earlier.

The three geological units are synthetic and their role in the study is entirely defined by the porosity and conductivity values described in lines 78–79. Their purpose is to introduce heterogeneity relevant for contaminant transport modeling, without referring to real-world lithologies. To improve clarity, we have now added the spatial dimensions of the domain in the figure 1 caption, as suggested. However, we chose to present all the analytical data regarding soil properties together in lines 74–83, in order to keep the description consistent and easily accessible in one block.

6) Line 70. Sentence "Flow... equation" is rather imprecise. ADE (advection-diffusion equation) is used to model transport and it is based on Darcy's law and on Fick's law for diffusion and dispersion. Moreover, Darcy's law is used to model flow, together 3 with the continuity equation (i.e., the mathematical formulation of the mass conservation principle).

We agree that the original sentence was imprecise and could lead to confusion between flow and transport modeling. We have revised the sentence to clearly distinguish between the equations used for flow (Darcy's law with the continuity equation) and those used for transport (the advection–diffusion equation based on Darcy's law and Fick's law). Thank you for pointing this out.

7) Line 71. The authors mentioned MODFLOW, which is a finite difference code, but here they underline the use of finite elements. OK, this is not a big problem for the specific work presented here, but I'd prefer to avoid confusion. Moreover, these sentences are not well related with the following sentences. And at lines 79 & 80, cell size is mentioned: this seems to be related to a finite difference approximation, rather than to a finite-element simulation.

Thank you for pointing this out. You are absolutely right — we mistakenly referred to a finite element solver, while in fact, the numerical simulations were conducted using a finite difference approach, consistent with the use of MODFLOW and the reference to cell size later in the text. We have corrected this in the revised manuscript.

8) Line 72. Increase or decrease K with respect to what?

Thank you for this remark. We agree that the sentence was ambiguous. The increase or decrease in conductivity due to the faults is defined with respect to the value that would be assigned by the underlying multi-Gaussian conductivity field in the absence of faults. We have clarified this point in the revised manuscript.

9) Lines 76 & 77, 80 & 81. Measurement units are missing for the coordinates of the reference point (lines 76 & 77) and of the border planes (lines 80 & 81).

This has been corrected in the revised manuscript.

10) Table 1. The position of wells is given with a precision of 10-8 m, i.e., 10 nm. This is not physically significant!

This has been corrected in the revised version of the manuscript.

11) Line 81. The average hydraulic gradient is 1/70, which is slightly greater than the typical value of hydraulic gradients, whose magnitude is of the order of 1/1000.

It is true, and any choice here would be subjective. A slightly greater gradient requires less time steps for transport simulations and thus lowers computing requirements as well.

12) Line 85. I do not like the “e” format used here to denote values: 3.5×10-5 ms-1 is much better, in my opinion. Measurement units should be attached to each value of a list, as required by rules of the SI system.

> We thank the reviewer for this suggestion. We agree that the scientific notation with “$\times 10^{-5}$” and proper unit formatting is clearer and more consistent with SI conventions. We have corrected this notation in line 85 and throughout the manuscript to follow this format

13) Lines 104 & 105. I do not understand the link between the two sentences “The conductivity field… discrete fields” and” here… 3D space”.

> We agree that the link between the two sentences could be clearer. While physics-based solvers like MODFLOW can handle discrete conductivity fields defined on non-regular grids (e.g., grids with variable cell sizes), in our work, we focus exclusively on regular grids within a bounded 3D space. We have revised the manuscript accordingly to clarify this point.

14) Lines 196 & 205. Using i for both the point index and the coordinates might be confusing.

> We thank the reviewer for pointing this out. In the original version, the index i always referred to the point index, not to a coordinate component. The expressions xi represented the 2D coordinates of the i-th pixel. However, to avoid any possible confusion, we have revised this section by renaming the variable x by the variable z= {x,y} and we now explicitly state that for each i, zi={xi,yi} denotes the coordinates of the i-th pixel in the image. This clarification should remove any ambiguity regarding the use of indices and coordinate notation.