Subject: Final author comments (ACs)

Title: Integrated approach for characterizing aquifer heterogeneity in alluvial plains

Author(s): Igor Karlović et al.

MS No.: egusphere-2025-327

MS type: Research article

RC2 (https://doi.org/10.5194/egusphere-2025-327-RC2)

General comments

Thank you for submitting this manuscript. Personally, I am curious about applying stochastic modelling for the purposes of electrical resistivity response of the subsurface, so I found it an interesting and engaging read. The reasoning, or wider context, behind the research is well formed (in that the aquifer which is the subject matter of the research is critical to the local water supply). The methodology in this manuscript seems competent, though I'm not familiar with the T-PROGS software that underpins the stochastic modelling efforts. I find the results insightful, however, I do have concerns about the statistical significance the stochastic outputs with and without the benefit of geoelectrical information.

Regardless I wish the authors well, and hope my comments are useful to them. Furthermore, I would like to recommend that this manuscript be accepted on the basis of moderate revisions.

Specific comments

Lines 93 – 95: Is it known how K is determined in these boreholes? Are they determined in situ (e.g. falling/rising head tests) or via laboratory investigations on samples.

The reliability classes presented in Figure 1 were developed to reflect the range and quality of available hydrogeological data in the boreholes, where parts related to information for K determinations refer to all possible estimation methods, including grain-size distribution analyses, permeameter tests from borehole samples, field investigations such as slug tests, pumping tests etc.

Lines 198 – 202: As the ERT is sensitive to 20 m below the ground level, I agree that the synthetic modelling should also be limited to 20 m. Although, I'm unsure if the inclusion of synthetic modelling improves the sensitivity of the electrical measurements to various hydrofacies below 20 m depth. If one proposes a tomography model (be it through a smoothness regularised gauss newton approach, as with R2, or via stochastic approaches) the raw data is still fundamentally only sensitive to the upper 20 m in this case. It's a limitation of where the electrical current flows, as the authors point out (Lines 198-199). Nevertheless, I think the authors could argue here that the synthetic modelling benefits ground model development by incorporating prior knowledge of the geology.

We agree with the reviewer that both field data-based and synthetic models are reliable down to a depth of approximately 20 meters. More precisely, this sensitivity limit corresponds to the depth where the conductive CSs/Sgcs layer is encountered, e.g. in borehole SPV-5 these two materials form thin conductive layer ranging from 20.7 to 23 m depth. We fully agree with the reviewer's observation that ERT measurements are fundamentally sensitive to the upper 20 meters, regardless of the approach used to estimate a possible geological model. This is the primary reason why Model Area 2 was limited to the upper 20 meters, allowing us to assess how ERT-derived lens lengths and varying grid resolutions influenced model prediction accuracy, while avoiding potential uncertainties arising from synthetic modeling at deeper levels. The purpose of employing synthetic modeling was not to improve sensitivity of ERT measurements, but rather to assess or test different possible geometries, improving the ERT imaging interpretation. The synthetic modeling was beneficial because it provided reliable and realistic hydrofacies dimension estimates that were used as critical input parameters for the T-PROGS model, which in turn improved the procedure for constructing the hydrofacies model in the subsurface. We acknowledge that our original text may have created some confusion about ERT sensitivity limitation, and we have carefully corrected this in the revised manuscript in Section 2.2.

Figure 4: How did the authors build the lens shapes in the mesh? The results are quite exciting. What is the data misfit of the traditional tomography section and synthetic model versus the real data?

In the revised manuscript, we include a step-by-step description of our methodology for estimating lens dimensions in Sections 2.2 and 3.1. The synthetic model development in ResIPy was informed by both ERT field data interpretations and borehole records. Particularly valuable were the borehole data, which provided information about depth intervals and thicknesses of individual hydrofacies units. Some boreholes did not intersect the CSs and Sgcs layer, indicating its discontinuous, lens-shaped nature. In the revised manuscript, we included the final RMS misfit values for both the synthetic model and real data inversions in the caption of Fig. 4.

Figure 5 b: Referring to Figure 2 (which shows a 3D visualisation of the boreholes) the "Clay to silt, sandy" (CSs) hydrofacies appears to dominate the near surface, furthermore, previous studies (Karlović et al., 2021) suggest that the geology is layered. Yet in this figure (5 b) the CSs hydrofacies distribution has little lateral continuity. Can the authors comment on why that might be?

The main purpose of developing the hydrofacies model was to more accurately characterize subsurface heterogeneity. Previous aquifer characterizations relied on simplified, layer-based conceptual models that often overlooked the covering layer. As discussed in the text, this layer is typically thin or absent, suggesting high infiltration potential and increased groundwater vulnerability to surface contamination. Consequently, the layer likely exhibits limited lateral continuity near the surface. Fig. 5b was constructed using 80% of the boreholes from Fig. 2 as hard constraints, ensuring the presence of the CSs hydrofacies at borehole locations. Another important aspect is the model scale, the entire Model area 1 domain (27,984 m × 16,142 m × 100 m) was discretized into 1,000,000 cells (100 × 100 × 100 in the x, y, and z directions), resulting in a horizontal resolution of approximately 280 m × 161 m per cell.

Line 254: What is not clear to me is which information from the electrical resistivity tomography is included or how it is included into the T-PROGS software. Is it just the lens lengths as stated on line 167?

This observation aligns with Reviewer 1's comments (RC1). In response, we have revised the text to more clearly define the role of ERT data as soft constraints within our methodological framework, specifically as ERT-derived hydrofacies lens lengths.

Figure 7 b: See above comment about Figure 5 b.

Similar to our response regarding Fig. 5b, the lateral continuity of the CSs hydrofacies is limited at the surface of Model Area 2 (Fig. 7b), but remains well-constrained at borehole locations. Note that the hydrofacies model in this figure has a horizontal cell resolution of 20 m × 20 m. Moreover, Fig. 3 effectively illustrates this pattern, as high-resistivity materials, consistent with the gravel and sand observed during field ERT measurements, are frequently present at the surface of the Vinokovščak wellfield.

Lines 308 – 309: Is an improvement of 0.3 to 5.0 % statistically significant enough to argue that the inclusion of ERT has improved the outcome of stochastic modelling? My experience with Markov chain Monte Carlo methods is that the answer can differ for repeated runs by a few percentage points. If the T-PROGS software is utilising Markov chains as the authors state (Line 136) then the question is whether the results are repeatable.

Both approaches (combining borehole data with either ERT-derived lens lengths or default lens lengths) produce a range of prediction accuracies. The mean values for all tested horizontal grid resolutions (10 × 10 m, 20 × 20 m, 40 × 40 m, 60 × 60 m, 80 × 80 m, and 100 × 100 m) consistently demonstrate better performance with ERT-derived lengths across 1,200 realizations (Fig. 8). While these results demonstrate repeatability, the observed improvement remains slight, as noted in the Abstract, Sections 3.2.2 and 4 (4). Other reviewers (RC1 and CC1) similarly observed the modest improvement in the prediction of hydrofacies distribution when ERT is added. In our discussion in Section 3.2.2, we propose a possible explanation to comparable results, related to similarities between ERT-derived lens lengths and T-PROGS default values (set at 10 × hydrofacies thickness). Additional details are presented in our response to CC1.