<u>Response to Reviewer Comments 2</u> (Responses in blue)

The manuscript presents novel heat transport/heat transfer experiments that address a long-standing research gap on the relevance of LTNE effects in aquifer systems. The manuscript is a highly valuable contribution to the research topic and provides great insights into the microscopic temperature distribution in large grain size sediments, building on previous experiments by some of the authors.

Thank you for your positive comments pointing out the value of our research. All of your comments will be considered carefully to improve our manuscript. The detailed replies are provided below each of your question.

Abstract: It would be great if the abstract could include some more quantitative values and clarifications. Examples:

- "some temperature differences were negative" -> the reader doesn't know what positive/negative temperature differences mean at this point.

We will correct the sentence by clarifying the results of inverse pulse ΔT, which are likely caused by delayed thermal arrival in fluid phase by the local advective flow.

- "relatively good agreement", "became too significant" -> can these statements be quantified?

The quantitative evaluation of the model fit will be included in the revised manuscript by providing the values of root mean squared error (RMSE) for each model result.

- I believe that the term "magnitude of LTNE" is not universally known, given the wide readership of HESS. Therefore, I propose to define the term if it is used in the abstract.

Thank you for your suggestion. The term of "Magnitude of LTNE" was used to describe the peak of $\Delta T(t)$ curve. Considering your comment, the term will be improved to better reflect its meaning in the revised manuscript.

Materials and methods

** The authors state at the beginning that "specialised experimental instrumentation" was used, but it remains unclear in the rest of the text what this "specialised instrumentation" is and how the current procedure was "adapted" from Gossler et al. (2019). The lessons learned from Gossler et al. (2019) are nicely highlighted in the following paragraphs.*

In this study, only one refrigerated bath circulator was used to prepare warm water as a heat source. To establish initial temperature, water circulation with outlet water was performed in this study as it is stated in line 126-127, while Gossler et al. (2019) provided cold water from a refrigerated bath. In addition, LTNE probes with Pt100 type A sensors are newly introduced in this study. We will improve this in the revised manuscript.

** Why were all samples measured in one experimental setup and stacked on top of each other? Why not test all grain sizes separately? Did the authors reverse the order of the spheres to check for a possible influence?*

Previous research has explored the relationship between LTNE effects and hydrogeological properties such as flow velocities and grain sizes. This study aims to investigate how varying grain sizes influence heat transport under different flow conditions. To do this, we drew on deep insights from previous studies (Rau et al., 2012a; Gossler et al., 2019; Bandai et al., 2023), when designing our experiment to effectively measure the temperature difference between fluid and solid phases. We decided that our experimental setup is particularly suited for testing heat transport for the following reasons:

- Placing all six different grain sizes into one experiment allows experimentation with identical flow velocities. This is difficult to do in separate experiments.
- We arranged the spheres in the column based on the spatial gradient of the thermal front. As heat propagates downward through the column, the spatial derivative (or steepness) of the thermal front decreases. Consequently, the smallest grain size, expected to have the least LTNE effects (Gossler et al., 2019), was placed at the beginning of the column to respond to the steepest gradient. Larger grain sizes were then arranged progressively deeper in the column.
- Experimentation is labour and time intensive. Overall, this approach provided us with the ability to efficiently experiment with two degrees of freedom (grain size and velocity) in one experimental setup.

We hope this clarifies your questions and welcome any further discussion or specific questions for additional clarification

** When filling the layers with the small glass beads, was it necessary to keep the outer PT100 free from contact with the smaller beads or did the authors rely on a (rapid) LTE between the small glass beads and water?*

PT100 sensors for fluid phase are surrounded by small glass beads without additional structure to keep them away from the beads. Although it might cause contact between the sensors and the glass beads, we assume that fluid phase and solid phase of small glass beads (d_p = 1 mm) establish an instant thermal equilibrium (LTE) resulting in the same temperature. We relied on a rapid thermal equilibrium between the glass beads and water. This is justified given the findings by Gossler et al. (2019) as well as our own results.

Analysis/Results

** What effect did the "correction" steps (pp. 225-229) have on the results? What was the difference/spread between the sensors? Should this be shown in Fig. 3? The curves in Fig. 3 are almost indistinguishable visually. In particular, which would be of interest, it is not possible to see whether the difference between Ts and Tf is consistent but shifted across the probes. Was there a systematic relationship between the spread and grain size or flow velocity? I would encourage the authors to try a different way of visualising Figure 3,* *perhaps adjusting the range of the x-axis or visualising differences rather than absolute values.*

The data was corrected to make the temperature difference between fluid and solid phases comparable. This allowed to determine streamlined LTNE effects as ΔT curves with a peak. Fig. 3 is presented to show comparison between temperature measurements from all sensors at the same depth. Considering your comments, the relationship between the spread and grain size or flow velocities will be analysed, and the further explanation will be included in the revised manuscript. Also, we will consider an improved presentation for Fig. 3 as suggested.

** The wording in the text and figures does not seem to be consistent throughout the text. For example, Figure 2 refers to "calibrated temperature", which is not used in the text. Also, "inner" and "outer" Tf measurements are not explained in the text (Fig. 3) - could this be done in Fig. 1?*

We will remove these inconsistencies in our revised version of the manuscript.

** The relevance of the negative temperature pulses is difficult to assess. l. 267f states: "was observed in some pairs of (...) measurements over all (...) flow velocities". What does this mean? Some replicates showed negative temperature pulses while others showed positive pulses? Because this is associated with the smallest grain sizes: Could this be an experimental problem, as the Pt100 sensor is the same size for all grain sizes, so its placement within the glass beads is more sensitive in smaller beads? Also, l. 270: "both sides of the grain" suggests that there is a discrepancy between the Pt100 sensors to the left and right of the glass beads. This would be interesting to see rather than Fig 5a,c,e which shows pretty much two lines on top of each other. Does this grain size discrepancy only occur for the small grains or is it only of less significance for the larger glass beads? Is this what you want to show in Fig. 3?*

Thank you for your comments with detailed questions. Although we have fixed the sensor position using the PVC frame and placed at the specific depth inside column as shown in Fig. 1., the possibility of sensor misplacement cannot be eliminated. We will further analyse the data investigating the relationship between the sensor positions and inverse curve results in our revised manuscript. Additionally, the sensitivity of small grain sizes to the Pt100 will be investigated taking your comment into consideration.

Discussion

** Can the "agreement" between numerical models and experiments be quantified in some way to quantitatively demonstrate the "better fit"?*

As it is stated earlier, we will add the root mean squared error (RMSE) for each model to the revised manuscript.

** The authors measure the temperature in the middle of the glass beads and refer to it as "solid temperature". However, within the glass beads there will be a radial thermal gradient. The "mean solid temperature" across the glass beads might be higher than the temperature* *measured observed by the authors. Hence, the heat transport within the larger grains takes longer time due to the larger distance from the outside of the glass bead to the temperature sensor. At the contact surface of the grain and the water, there still might be thermal equilibrium. While the subsequent theoretical considerations in terms of the influence of flow velocity go beyond the scope of this work, a short discussion on the scale effect could enrich the Discussion section.*

Temperature for solid phase was measured at the centre of glass spheres to represent solid temperature. The surface of glass sphere was technically difficult to measure due to the thickness of the sensor and the small contact area of each glass sphere. The challenges and inconsistency of the surface solid phase temperature measurement have been reported in the study of Bandai et al. (2023), which conducted heat transport experiments measuring solid phase temperature. We will add a short discussion of the implications to our discussion in the revised manuscript.

References

Bandai, T., Hamamoto, S., Rau, G. C., Komatsu, T., & Nishimura, T. (2023). Effects of thermal properties of porous media on local thermal (non-)equilibrium heat transport. Journal of Groundwater Hydrology, 65 (2), 125-139. doi: https://doi.org/10.5917/jagh.65.125

Gossler, M. A., Bayer, P., & Zosseder, K. (2019). Experimental investigation of thermal retardation and local thermal non-equilibrium effects on heat transport in highly permeable, porous aquifers. Journal of Hydrology, 578, 124097. doi: https://doi.org/10.1016/j.jhydrol.2019.124097

Rau, G. C., Andersen, M. S., & Acworth, R. I. (2012a). Experimental investigation of the thermal dispersivity term and its significance in the heat transport equation for flow in sediments. Water Resources Research, 48 (3). doi: https://doi.org/10.1029/2011WR011038