

## Review of the manuscript

*“Characterizing groundwater heat-transport in a complex lowland aquifer using paleo-temperature reconstruction, satellite data, temperature-depth profiles, and numerical models”*

Manuscript ID: hess-2021-586.

### **General comments:**

The manuscript introduces a numerical groundwater flow and heat transport model for the region of northeast Flanders, Belgium. It is a very detailed case study that is very useful for other hydrogeologist studying this region (and potentially other students working on setting up comparable models for other regions), but it is sometimes hard to follow for someone not familiar with the site. As I am not an expert in numerical modeling I cannot speak to this aspect of the study, however at this stage there are some shortcomings in the structure of the paper and in the selection of surface temperature data as an upper boundary condition.

I am particularly worried about the lack of separation between satellite derived land surface temperature, air temperature, and (ground) surface temperature/soil temperature (which would be relevant for the subsurface thermal regime). While these temperatures correlate, they are not the same.

I am furthermore not clear of the benefit of running the model based on temperature time series starting in 8500 BC – we do not have accurate data for this far back and more importantly, the model reaches down only 400 m at most, at this depth surface temperatures from maybe a couple hundred years ago are dominating the thermal regime. Hence, I must recommend focusing rather on high-frequency and more accurate data that's a bit more recent than on long time-series such as paleo-temperatures (e.g., climate reanalysis data such as ERA 5 or MERRA-2 or GISS Model E2)

On a smaller note – I had a hard time really understanding study area. The authors could consider renaming their boreholes and faults to somethings more intuitive. Right now, the boreholes are called R-51 which doesn't give a ton of info. I'd suggest something like borehole1\_2019 s in first borehole from the left, measured in 2019. They also often refer to parameter names or other sites in the text that are not mentioned anywhere else and thus pretty irrelevant and non-gettable for a reader not already familiar with either the study site or the model (e.g. Fig 11 or Line 401).

### **Detailed comments:**

1. Line 55 “However, in many cases [...], when there are multiple unknowns on flux exchanges and thermal and hydraulic properties, it is likely to reproduce temperature correctly ...” What is ‘it’?
2. Line 65: “Several studies where temperature profiles are measured are typically used for qualitative interpretations of the effects of anthropogenic stressors.” Do you mean “... are measured, they are typically used...”?
3. Line 110 – discussion of the ON-MOL-2 site seems highly irrelevant. It is not shown in any of the figures or discussed in the results and discussion sections.
4. Chapter 3.1.1.
  - reference to Figure A1.
  - I’d also appreciate a clearer indication of the year (and maybe even month for seasonality impacts) the measurements were taken in Fig A1 and Fig 2
  - Give some info here about the depth of these TD-profiles
  - Where any of the measurements repeated or are the 35 TD profiles taken in 35 different locations?
5. Chapter 3.1.2. what is the depth of the soil temperature measurements?
6. Chapter 3.2.3.
  - While not really relevant to the study the suggested missions (Landsat, ASTER, MODIS and Copernicus) and the discussion of those are worded in a way that could be misleading. Specifically, Landsat LST are available globally, just not daily, and with Copernicus I guess you are referring to the second gen Meteosat?
  - Please give info what kind of MODIS data you are using. Day or night? Terra or Aqua (i.e., MOD or MYD)?
  - Line 169 “From this sampling, it is clear that the raster LST images overestimate the peak temperature values, mostly during the summer months” you are comparing LST to air and soil temperature. They are not overestimating air and soil temperature; they simply measure a different (but correlated) value! Hence you are not correcting LST, you are estimating (soil? air? I’m not clear) temperatures from LST.
  - Line 175 LST is fitted to soil or air temperature?
7. Chapter 3.1.4 – are these temperatures air, soil or LST? I assume air but should be made clear. Also, as mentioned in the main comment I’m doubtful about the need for such a long (and inaccurate) time series. I’d suggest keeping the 90 time steps but for a shorter time series.
8. Line 212 – what depth is the bottom boundary? It is shown in the figure, but I’d appreciate a short sentence here
9. Line 228 – bracket is never closed (‘) missing).
10. Line 245 “(Databank Ondergrond Vlaanderen; DOV).. The model” one dot too much.

11. Line 317 – how where LST downscaled?
12. Chapter 4 – generally I'm not sure for what date your model gives output. Dec 2019? How are you then comparing this to the TD-profiles from previous years?
13. Line 356 – why are you comparing temperature gradients instead of absolute temperatures when comparing these models? And how did you (statistically) calculate the gradients?
14. Line 360 – Why did you split the model at -5m.a.s.l.? In temperature modeling meters below surface are of course much more relevant hence I believe a split at maybe 20 m below ground would be more useful, above those seasonal impacts may be way more relevant than anything else.
15. Line 387 – Space missing between “Figure 12b.Figure 12a”
16. Line 401 – I'm not really seeing any blue dashed lines in Fig 12 a (or b) – also seems highly irrelevant as “HFB\_094” etc are never mentioned again.
17. Line 426 – are measured and observed shown for the same time?
18. Line 461 – how does this strongly support the statement that heat transport is driven by advection? You are comparing two models, you have not shown me yet which one is more accurate in these locations, have you?
19. Chapter 5
  - Could you give some sort of accuracy assessment of your models? This info is in part scattered throughout the manuscript, but as a reader I would benefit from an extra chapter just giving me all the info in one location which models are how accurate in which locations (near surface, near faults, etc.)
  - I'd like to see a (qualitative) cost-benefit analysis of your model. Is the extra time and computing power needed to run the long time series worth the effort or is an analytical solution basically as good and much faster, less costly in computing power.
20. Fig 1 – really hard to read. It was rather low res in my pdf and the colors. I cannot read any of the writing on the map where LST is shown. The legend is missing the yellow marks showing boreholes, are temperatures shown as discrete values (as shown on the legend) or as a stretched values along a color ramp?
21. Some of the methodology figures could be moved to the Appendix (Fig 1-9)
22. Fig 8 – is meters here m.a.s.l. or meters below ground?
23. Fig 10a – make it square so it's easier to compare. Unit for  $z \leq -5$  and  $z > -5$  missing. Also use  $\geq$  instead of  $>=$
24. Fig 11 – the parameter names along the x-axis mean nothing to me, they are not mentioned in the manuscript anywhere else.
25. Fig 12 – the cyan and green arrows are barely distinguishable in my pdf
26. Fig 13 – pink circles not in legend, include B-B' line, in b the names of the boreholes are different than anywhere else (i.e. the small case letters at the end)

27. Fig 14 – give the names of the wells so I can compare.  
28.