Effectiveness of distributed temperature measurements for early detection of piping in river embankments
Authors: Sylvia Bersan et al

The manuscript discusses both the analysis of the results of a large-scale (so-called 'IJkdijk') experiment in the Netherlands to test the effectiveness of DTS in the early detection of backward piping erosion under dikes and a theoretical numerical study to predict thermal response below an embankment under seepage conditions. The study is of high relevance for the dike health monitoring community and extracts the relevant results of the IJkdijk experiments which at the moment have been described mainly in grey literature. It present novel data and the analysis to try to use DTS for piping erosion (internal erosion) is scientifically challenging. The paper has a strong geotechnical focus and terminology and somewhat less hydrological / earth scientific. Consequently the authors have not benefitted from the impressive amount of published research on DTS in hydrological systems. Moreover, I found the paper's structure confusing to say the least and the (modelling) assumptions not well explained and discussed. Maybe because it has been written like a work flow or learning curve of the authors. To tackle this, the paper needs major restructuring (I will elaborate on that below). In my opinion, the writing style and some figures also need some more work. Lastly, unfortunately, the paper seems to consist of two part (experiment and model) but the two parts do not get together. In conclusion, to me it is a potentially interesting paper but in current state it is not ready yet for publication in HESS.

In short I suggest restructuring as follows (details follow):
1 Intro
2 Theory and methods (including IJkdijk set up, maybe partly in supplement)
3 Perception and conceptual model of how DTS could detect piping
4 IJkdijk experiment results
5 Numerical results
6 Analysis/discussion of the both results and compare DTS and model
7 Conclude about possibility to detect piping and suggest optimal conditions and location

First, the authors start from DTS applications in geotechnics like dam monitoring but do not take advantage of the large amount of literature that is available in the hydrological community using high resolution DTS measurements to interpret the hydrological system, especially groundwater-surface water interaction (see work of e.g. Selker, Krause, Mamer, Fleckenstein, Cassidy, Briggs and many others to start a literature review). The here-described experiment has many similarities with published ground- surface water interaction studies and awkward statements can then be avoided like P3L27: "heat transfer in soils mostly occurs by conduction". So I suggest to write a generic intro on heat and water transport and how it can and will be used in your study. The paper should add comparison with published groundwater–surface water studies here and as such way stress the novelty and show the relevance for the hydrological community.

Second, the paper needs a clear methodology section in which the theory coupled heat and water flow modelling and the large-scale experiment are given. Describe both methods in the section 2 (bring 5.1 to section 2). How is the DTS calibrated? How did you make sure to distinct temperature from strain induced changes as there is quite some strain on the FO cable which influences your measurements. The authors should clearly explain why the numerical model is needed/used (to what purpose do you model), why this formulation has been used and what assumption they made. Also add scientific references here. I think the main aspects of Bersan model (PhD thesis) should be included in the paper (maybe as supplement) as it is very relevant for this research. Furthermore, I found parts of this description unclear and confusing. For example, why you use specific discharge rate and not Darcian velocity, or P10L30 "neglects effect of conduction" followed by P10L31 "describes the effect of heat conduction". P11L8 the authors start writing annual temperature fluctuation under the assumption that "u=0" (P11L10). But this does not hold for your condition so elaborate on this assumption and why it can or in hindsight cannot be used. This effects strongly the interpretation of your results. Can the authors elaborate on this?

Third, the authors give some confusing if not conflicting description on what they expect from DTS under an embankment. Moreover, they do so in different parts of the paper. I suggest to include a structured section on the perceptional and conceptual model the authors have of using thermal measurements at the toe of the embankment (fig 13). This in terms of expected advective and conductive heat transport (see also first lines of section 4). A first statement on the expected flow travel time underneath a dike which you will elaborate later. Maybe also interesting to say something on expected temperature difference in river water during a year and during a flood (do you know if river water temperature changes a lot during typical flood events?). Basically, under which seepage conditions could DTS be informative for detecting piping (initiation) and when not. Obviously, if all water has identical temperature, nothing will be detected using DTS. Second, if the river water is warmer/colder than the GW, it is not expected to see piping developing at the toe of the embankment. In contrast to what the authors write (P7, L22-23), if you detect T differences due to river (reservoir) water the preferential flow path (possible pipe) already exists from river to toe of embankment and does not need to develop anymore. And your objective is detection the initiation and backward development of the pipe. Backward eroding cavities do not connect yet to the river and will logically first see inflow of the surrounding water (no T-difference). However, in case of warmer/colder water at the toe of an embankment compared to water underneath the embankment, piping initiation starting from the toe and eroding backwards, can be detected (as fortunately the case in this experiment). Note that strictly speaking one does not detect piping with DTS but rather preferential water flow. In my opinion such structured concepts will help describing your field data analysis and theoretical modelling results.

Fourth, the field results (current 3.3). This can be reduced to only data as concepts are given already. I suggest to the authors to stop describing the experiments after t=90hours as hereafter it is not a piping experiment using DTS anymore but a failure experiment which is not part of the paper's focus.

Lastly, the results of the dimensional and numerical analysis can be presented and the influence of the model and boundary assumptions be discussed. I, as a small example, do not understand why Figure 11 presents data using a constant hydraulic load of 3 m was used while the auhors have time series of hydraulic load during the experiment. The dimensional analysis presented in the manuscript allows to quantify the relative importance of conduction versus advection. This conclusion is important for the propagation of the front assumption which requires advection to prevail. The results of such analysis are based on the assumptions of steady-state boundary conditions as acknowledged by the authors. Then the question arises if the model used for determining the thermal front propagation is representative of the actual conditions of the experimental set up?

This brings me to my final point, I was somewhat disappointed to see that the comparison of field and model results is quite minimal to almost absent. This is a missed opportunity and should be included in the paper.

Some minor points:

- Comparison with use of DTS in concrete dams seems not so relevant in your study
- Change terminology "Downstream and upstream of water under embankment" into "landward and river side"
- P4L10: "easily installed also in existing embankments" That is quite a statement. This only holds if the FO is installed in the toe (landward side) of the dike. But it is unclear if that is an optimal location for the FO in case of piping detection. I believe this is far from optimal location. Authors should explain this more.
- The derivation of the thermal front velocity starting from the heat transport equation in porous media is used to determine the "Retardation" coefficient. This coefficient is presented in equation 3 but its derivation is not yet clear.
- Time scale of figures could be homogenized to facilitate interpretation for reader
- Figure 6: It is written that the toe location temperature (F1) is highly influence by the air temperature whereas in the plot it is observed that it seems to have only influence between t= -7 days and t= -2 days air temperature is constant between t=0 days and t=5 days but the temperature at F1 continuously drops. Please explain.
- Why F1 reaches a temperature lower than any measurement in subsurface? Did you take into account that the FO cable could be strained due to displacement and therefore shows deviating temperature measurements (this links to the calibration question of the DTS.).
- Figure 7: plot this figure in color gradient for improving the readability.
- Figure 9 shows negative gradients in X=18. Why is this?
- Figure 10 could be explained based on the possible variability of the hydraulic conductivity.
- Figure 11 seems like is not considering the same initial conditions as the experiment. Also, the figure 8 at t=60h (approximately 2.5 days) shows a boundary temperature of around 14 degrees but the model results remain in 12 degrees. This may be interpreted as a large error taken into account that the steady state is achieved by a total increase of only 4 degrees.
- Figure 12 is not explained neither the variables or the system. This figure is a copy (very minor modification) of the original presented in Van der Kamp and Bachu 1989.
- Figure 15 and 16 show the temperature gradient along the x axis but it is not clear if they are located in y=0 or further. Please explain and compare with the experimental measurements.
- In addition, include the information of the moments when sand boils are observed as in figure 5 in other relevant figures.
- The grammar should be checked as there is an indiscriminate use of commas and frequent use of very long sentences.