Review to: nhess-2020-391 paper, "Fault distance-based approach in thermal anomaly detection before strong Earthquakes", by Arash Karimi Zarchi and Mohammad Reza Saradjian Maralan

The paper deals with a novel method of using thermal anomalies in LST records to investigate their correlation with strong Earthquakes and to assess their possible use for EQ forecasting. The paper is well written and organized but I cannot recommend publication in its present form, according to my comments here in the following.

General comments

Many papers claiming any relationships between physical/environmental parameters (as possible precursors) and earthquakes typically deals with the occurrence of anomalies before (or after) the quake. Only a few also look at the possible existence of such anomalous patters even in unperturbed periods (i.e. when no EQ occurred). The present paper suffers from the same strong limitation: it investigates 6 intense EQs occurred in the study area by means of land surface temperature (LST) anomaly analysis, demonstrating the occurrence of LST anomalies before or after the EQ, but it doesn't verify if similar thermal patterns occurred even when no EQ occurred in the area. This is, in my opinion, the main weakness of the paper that needs to be improved, including, at least, a full "falsification/confutation" analysis to be considered for publication.

Other general comments refer to some hypotheses, at the basis of the rationale of the study, that are poorly scientifically based and should be carefully handled.

1) Fault identification: often, a "fault systems" rather than a single fault, is activated during strong EQs. How the method can handle this? Not clear. In addition, the identification and selection of the fault is, even in this work, carried out aftermath, when the epicentre position is well known: how can manage this in terms of EQ forecast?

2) LST anomaly identification: to remove natural/observational noise, authors consider "…a previous year with no strong seismic activities…" and subtract this "linear function" from the data. In my opinion, 1 year is not enough to be considered as representative of the actual "normal" conditions of the area. A single year, in fact, can be affected by other (e.g. meteorological) forcing factors, limiting its representativeness for a "normal" year and significantly impacting in thermal anomaly identification.

3) coefficients k: to detect a thermal anomaly, authors apply two different methods, both requiring a threshold test based on a "k" coefficient that seems determined in a totally arbitrary way. K values determination, as well as their possible variability and dependence on different environmental/observational/geographical conditions should be better justified and assessed. Additionally, it is not completely clear how the average and sigma are computed.

4) Buffer radius: authors consider LST mean computed in buffer radius from 1 to 20 km. Why limiting R < 20km? It is well known that the "preparation zone" of a strong EQ could be as large as the Dobrovolsky theory (Dobrovolsky et al. 1979), i.e $R = 10^{0,43M}$ (with M= magnitude). In particular, for M=6 a radius of about 400 km can be expected for the EQ preparation zone. Moreover, if thermal anomalies are related to fault degassing, gases (e.g. CO2) might be spread in large areas, also depending on meteorological conditions (e.g. wind intensity and direction) and/or local topography. Therefore, the limitation at 20 km should be better explained and scientifically justified.

5) How ANN works to estimate EQ intensity? Have the authors trained the network using the same dataset they used for results or an independent one?

Specific comments

- Figure 3: what do the black dots mean?
- It is not clear how the fault distance map is used in the process. Please clarify.
- Not clear how the LST mean is computed in buffer radius. In particular, please better explain the following sentence: "*It should be noted that width of each buffer is only 1 km and R is the buffer radius (distance) from the related fault*."
- Lines 204-207: authors assert that anomaly far from the epicentre are not used for ANN. Thus, authors are using epicentre information for filtering data, but epicentre position is only available after EQ occurred, so this study appears as a retrospective analysis as well, and cannot be used for forecasting.
- Lines 230-231: "*Although these anomalies are not as strong as the anomalies detected near the time of the earthquake, they seem to be related to some seismic activities rather than being a false alarm*": how authors can assert this? What is the scientific basis for asserting this?
- Lines 233-234: …" *Many anomalies detected by this method are related to the earthquake*…". Again, how authors can affirm this?
- Lines 244-245: "*ANN results also show high correlations between thermal anomaly data and 245 the earthquakes intensity*." What exactly does this sentence mean?

Cited reference

Dobrovolsky I P, Zubkov S I and Myachkin V I, 1979 Estimation of the size of the earthquake preparation zones Pure Appl. Geophys. 117 1025–44.