The authors present a data-driven model to predict flood susceptibility at ~10 km resolution, using 11 geographic datasets trained against point observations of flooding from DFO and EM-DAT. It is regrettable that I must judge this work to be of poor quality in many respects, and can therefore only recommend it be rejected for publication in NHESS.

Flood observations of this provenance over 20 years are not good enough to be skillful in the characterisation of flood hazard at this scale. Just because a flood hasn't been recorded since the year 2000, doesn't mean a flood did not happen (within the time period, or indeed before or after it). All your model has therefore done is replicate where DFO/EM-DAT has recorded floods in the past 20 years, having been trained on where DFO/EM-DAT has recorded floods in the past 20 years. The skill score (AUC) is extremely high, but it is skillfully replicating something that is neither useful nor interesting: GIGO. It is therefore not clear how any of the data produced in this study advance our understanding of flood hazard, or present novel methods capable of doing so.

The writing is generally of poor quality – for which I am sympathetic to the authors – but it does make the manuscript difficult and/or unpleasant to read in some places. The paper is too long, containing much superfluous information, much of which is incorrect anyway (for instance, claiming the GDP ($2.3bn) of a region spanning 3 continents being roughly equivalent to that of a small English market town). It is overloaded with references that are not needed, which are often unrelated to the point being made. I personally have never heard of the "Belt and Road region", and so would refer to the study area as something else (or make it less arbitrary). Indeed, there's no reason not to deploy this method globally, given its simplicity.

I am sorry to not be able to give a more positive review of this work.