I have to say I am a bit unsatisfied with the paper by Sozdinler and co-authors. The subject is very interesting and relevant, and the title claims for high expectations. But in my opinion several improvements at least in a couple of key aspects must be made for the paper to be considered for publication.

First of all, I need clarifications regarding the general philosophy adopted to build up the database. The authors clearly state that a scenario approach is adopted. What kind of scenarios are we dealing with? Just credible? Worst-case scenarios? This aspect remains obscure throughout the paper. What is the database conceived for? In the "Conclusions" the authors state that the database serves for tsunami hazard assess-
ment related to earthquake sources, which is partially in contrast with the second of the two main objectives declared in the Introduction, i.e. to provide a suitable instrument for the RETMC. The way in which the term "scenario" applies to the two different scopes changes dramatically. If you want to carry out an hazard assessment, very likely you will need to adopt a credible worst-case scenario approach; in the other case, you will rather need a database of Green’s functions which you can play with in real-time to iteratively improve your forecast. The approach adopted by the authors remains unclear to me and appears to be an hybrid between the two previous extremes. With reference to Table 1 and 2, I ask the authors to clarify what is the sequence of the steps. If I got it correct, the authors first adopt geometries for different fault segments in the Marmara Sea, taken from previous studies and from the MARSite database. In other words, length and width, and hence the fault area, are fixed, as well as the focal mechanism. Then, with the area the authors use two different regressions (Leonard 2010 and Wells and Coppersmith 1994) to retrieve magnitude and on-fault slip. Is this correct? Then, how is seismic moment computed? Through the Hanks and Kanamori (1979) relationship or as the product of area, slip and rigidity? Which of the columns are used to build up the 30 scenarios? The one relative to the regression by Leonard (2010) or the other one? This is a key step in the study and is poorly explained. It deserves a more detailed and in-depth explanation. The authors also state that "Slip values have been assigned using the same logic but in an arbitrary manner without any prior assumption, so that heterogeneous earthquake rupture scenarios can be represented". But I understand that just one realization of slip heterogeneity is adopted among the several possible. The way this heterogeneity is introduced dramatically influences the results in terms of tsunami maximum coastal elevations, so the choice cannot be single and arbitrary.

A second key point that deserves further clarifications is the need to introduce landslides as tsunami sources in order to justify the historically observed maximum tsunami heights/run-up/inundation. I can agree on that, but I need to see at least a map with the distribution of the historically observed maximum tsunami effects. By the way, the first
of the two motivations for the study indicated in the Introduction reads as "investigating the nature of historical tsunamis in Marmara Sea, namely whether they are generated solely due to those significant earthquakes or not". I believe this point has not been discussed at all in the paper. The reader has no mean to judge about it.

Other aspects should be further discussed. Regarding the tsunami simulations, maximum water elevations are computed at 1333 coastal points placed at some depth. This means that, despite the fact that NLSW equations are solved, no inundation is computed. Probably, linear equations would have served well enough for the same scope. Moreover, not accounting for inundation probably brings an underestimation of the maximum water heights, at least in some places. I ask the authors to comment on this. The argument that a probabilistic approach would result in larger maximum water heights is misleading. If this was true, then why not adopting a probabilistic approach from the start? Still regarding the tsunami simulations, I believe that only 2 hours of simulation is too a short time. This is confirmed by the largest part of the computed time histories shown in the supplement. I strongly encourage the authors to extend the simulations at least to 4 hours.

The presentation style is not as "well-finished" as a publishable paper would require. Apart from the English, which is not up to standards, too few explanations are given for the key points of the approach and for the applications of the database. Despite this, due to the relevance of the subject and of the geographical area, I think the paper could be considered for publication after major revision. I am also attaching an annotated pdf version of the manuscript.

Please also note the supplement to this comment: https://www.nat-hazards-earth-syst-sci-discuss.net/nhess-2019-186/nhess-2019-186-RC4-supplement.pdf