Reply on CC2

I have read this preprint with great interest and think that it is an important study for under-standing the influence of heavy rainfall events on damage to rail transport and infrastructure. My colleagues and I developed CatRaRE and we highly appreciate the use of the dataset in this study. However, there are a few issues in the description of CatRaRE and the results that I would ask the authors to address. In the case the authors have any questions or would like to discuss some of the issues mentioned below, please feel free to contact me or my colleague Ewelina Walawender.

- CatRaRE catalogue: In the abbreviation CatRaRE the word "catalogue" is already included, therefore CatRaRE catalogue would mean Catalogue of Radar-based heavy Rainfall Events catalogue. I suggest to use either just CatRaRE or catalogue of radar-based heavy rainfall events

Thank you for the comment. We will use the term „CatRaRE" in the revised version.

- P.3, L.72: CatRaRE W3 and T5 are DOI referenced datasets. Please use the appro-priate reference for the catalogue used in this study, which I guess is the Version 2022.01:

Lengfeld, K., Walawender, E., Winterrath, T., Weigl, E., Becker, A., 2022, Heavy pre-cipitation events version 2022.01 exceeding DWD warning level 3 for severe weather based on RADKLIM-RW version 2017.002, DOI:10.5676/DWD/CatRaRE_W3_Eta_v2022.01.
In case another version is used, please check
https://www.dwd.de/DE/leistungen/catrare/catrare_daten.html?nn=16102&lsbId=751876

We will ad the reference for the catalogue in the revised version. The verion 2022.01 is correct.

- P.3, L.73-74: In CatRaRE events with 11 different durations between 1 and 72 hours are listed. In the catalogue W3 we use the lower boundary of warning level 3 as a threshold. Not only the warning levels for 1 hour (25 mm) and 6 hours (35 mm) are used, but also the ones for 12 (40 mm), 24 (50 mm), 48 (60 mm) and 72 hours (90 mm). There are no official warning levels for rainfall events with durations of 2, 3, 4, 9 and 18 hours. Therefore, for these durations we linearly interpolate the official warning levels and get thresholds of 27 mm in 2 hours, 29 mm in 3 hours, 31 mm in 4 hours, 37.5 mm in 9 hours and 45 mm in 18 hours.

Thanks for the clarification. We will add the description in the revised version.

- P.4, L.98 and Table 1: The SRI does not describe the speed at which rainfall accumulates within a specific duration of time. The SRI is based on the return period of the rainfall amount for indices 1-7, where 7 corresponds to a return period of 100 years. Indices 8-12 are based on the rainfall amount compared a precipitation with a return period of 100 years. Please clarify the description and see Schmitt (2017) and Schmitt et al. (2018) for more information.

We will add the correct definition and add the references Schmitt (2017) and Schmitt et al. (2018) in the revised version.

- P.6, L.136: What is the reason for choosing the HYRAS dataset over the climatological radar dataset RADKLIM? RADKLIM would correspond to CatRaRE and has a higher spatial resolution comparable to the soil moisture dataset.

The soil moisture and HYRAS dataset (contrary to the stated description that will be corrected) have the same resolution. It was chosen in connection with the soil moisture dataset, as they are both based on station based observations of precipitation.

- P.10-11, L.252-256 + Section 3.3.: The CatRaRE event variables the authors have chosen (Tab. 1) are calculated for the whole event area (e.g. as an average over the event zone), that can in extreme cases cover several thousand km². However, the damage data used in this study are available for a given route segment (point location), so the cross-analysis makes sense only if a given rainfall event is undifferentiated within its zone in terms of precipitation characteristics (RR, SRI, V3) and occurs over an area with similar landscape pattern (TPI, VSGL, STRM). A pixel-based analy-sis would be more appropriate in this case. Also using the ETA as a measure of extremity is not proper in case of point-analysis, as it is calculated exactly on the basis of the event area.
The accuracy of the geolocalisation of the damage events is fairly low. A pixel based evaluation would simulate a fake accuracy. Hence, we decided to use the area-averaged characteristics of the heavy rainfall event.

- P.11, L.270: The fact that only 23% of the flooding events are linked to heavy rainfall events seems surprising to me. Did the authors also check for rainfall events in a certain radius around the flooding since rainfall does not necessarily cause flooding in the region of its occurrence but in the region where the water flows to.
See response to comment from RC2, line 79.

- P.13, L.312-314: It is not clear to me why a rain event should cause a flood one or two days after it's occurrence. If there is no more rain in the area there shouldn't occur a flood unless the water comes from another region, e.g. from upstream a river. But then the flood is probably triggered by another rainfall event that occurred upstream and not by the one that occurred in the area with the flooded railway section There-fore, not only a temporal but also a spatial buffer should be taken into consideration. In case the flooding occurred one or two days after the rainfall event I would also suggest checking the HYRAS dataset if there was more rainfall in the damaged area or the surroundings that could have caused the damage but wasn't classified as an event in CatRaRE. I understand that a detailed analysis of flow paths is beyond the scope of this paper, but the issue as well as the difference between damages caused by a heavy rainfall event and by a flood event should at least be mentioned in the discussion.
We added the HYRAS dataset in our analysis with the intention to detect additonal rainfall events which are not included in CatRaRE. In our opinion, the approaches of the recording methods and the spatial reference from HYRAS and RADOLAN complement each other well. A detailed analysis of the flow path is not the focus of the paper, but we can add a few sentences on this and the differences in the recoring method in the discussion.

- Section 3.3: I am not sure if increasing each parameter by one unit is appropriate. 1 mm increase in mean precipitation is not comparable to increasing the duration by 1 hour or the SRI by 1. Let's e.g. assume a precipitation sum of 50 mm in 1 hour has a return period of 100 years, which corresponds to SRI = 7. Increasing the precipitation sum by 1 mm leads to 51 mm in 1 hours which most probably still has a SRI of 7 because the return period won't increase that much. Increasing the SRI by one to SRI = 8 would mean according to Schmitt (2017) that the precipitation would be 1.2 to 1.4 times the precipitation sum for return period of 100 years (which is 50 mm in our case). Therefore, increasing SRI from 7 to 8 would increase the precipitation sum from 50 mm to a value between 60 and 70 mm, which is 10 to 20 times more than the increase of 1 mm that was assumed for investigating the influence of increasing the precipitation by 1 unit. Therefore, the influence of

increasing the SRI by one unit is by definition larger than the influence of increasing the mean precipitation by one unit. Also, I don't quite understand how the duration of precipitation is increased. In my example I had a duration of 1 hour and 50 mm. Does increasing the duration by one unit mean that it will rain 50 mm in 2 hours instead or 2*50 mm = 100 mm in 2 hours?

Our analysis represents an initial evaluation of the quantitative effects of the control variables. Without further analysis, the values between the control variables are only comparable to a limited extent.