Reply on RC3

Major comments

1. For the reader it is not easy to gain an impression of what the DB Netz AG database looks like. Because some of the results are somewhat puzzling and surprising (see point 2), we request the authors to provide more information on how this dataset looks. A few concrete suggestions:
    1. Include a few example records from the database showing the raw data (i.e. the raw text record that was identified), preferably with examples showing all three types of natural hazards; and examples that did match with rainfall events and examples that did not match.
    2. Please give an example(s) of how a well-known flood event (example the July 2021 floods) was mapped in the data, ideally with some pictures of how the damage looked in the real world. Earlier NHESS articles (e.g. https://doi.org/10.5194/nhess-22-3831-2022) and the author's https://doi.org/10.3390/atmos13071118 contain many relevant details that can be used to link the records to.
    3. Please give examples of that were tagged as a flood by DB, but for which no extreme rainfall was reported.

    The authors are not part of the DB Group, and the DB Netz AG database is an internal documentation that is not publicly available. Unfortunately, we have no rights to publish raw data from the database and only receive extracts from the database and have no access to the database itself. We are therefore unfortunately unable to comply with point 1 for data protection reasons. Point 2: The studies mentioned use different data sources than we use in our manuscript. Depending on the size, severity and impact of an event, there is an obligation to document it in varying degrees of detail, meaning that significantly more (and different) data is available for the July 21 event than for smaller events. The database used in our paper contains damage reports that are recorded during ongoing operations. This means that if traffic is completely suspended in the event of a major incident, the effects on the routes are only recorded at a later date and in a different form. In order to ensure comparability between the individual events and the years, we have decided for this manuscript to use only one data source for damage events and not to compile information from different sources.
    In the revised version, we can add a more detailed explanation of the dataset and give examples of some selected events with pictures how the damage look like.
2. We find the finding that only a quarter of the flood events could be linked to extreme rainfall events (line 414) very puzzling. In our view, the potential causes of this are insufficiently discussed in the draft version. Please critically reflect on the following possible causes:
    1. Is it possible that most of the reported flood events relate to river (fluvial) flooding? On the one hand, one might expect that this is the case, because it would present a natural explanation for the fact that the damage is observed in a different location than the extreme rainfall event. On the other hand, we would be surprised if the ratio between fluvial and pluvial flood events would be 3:1. Also, if 75% of the flood cases would concern river flooding, one may wonder if the methodology of the paper is sound. The authors mention this aspect in line 463, but what is missing is a reflection on whether this explains the low correlation in their results. *As mentioned above, the information in the database on each event are limited, so that it is not*

possible to evaluate which proportion of the events relate to river flooding. Another limitation is that the spatial intersection does not take into account the whole catchment area and flow direction (see also comment from RC2, line 79). However, the infrastructure operators (DB) are currently changing their documentation system for such events, so we hope, in future, it will also be possible to compare events with images and unique keywords. Unfortunately, this is not relevant for the manuscript, as the changeover and generated data sets will only be available in a few years' time.

2. Another take would be that most of the floods are caused by rainfall events that do not qualify as heavy rainfall, which is reasonable given that there is no standardized guideline for defining heavy rainfall (line 72). If this is the case, it would be best to clarify it in the text and highlight it as relevant further research. *As mentioned in the response to comment 2.1 above, the data quality of the data base from DB is limited and the whole catchment area is not taken into account. Precipitation events not classified as heavy rainfall events in the CatRaRE-data set but occurring in the catchment area could be detected by a detailed analysis oft he HYRAS data set used in our study. However, that would be a very complex analysis (high resolution hydrological modelling over whole Germany), which goes far beyond the scope oft he study.*
3. Is it possible that due to other reasons, there is a mismatch between what the DB understands as a 'flood' is very different from what the radar data shows? *This may also be, due to the fact that the database is not filled with events by experts, but by the staff on site along the route. It is therefore possible, that technical terms are not always used correctly and flood events are not differentiated according to their cause. An event is only recorded if it disrupts the railway operations, so there may occur heavy rainfall events or river floods that are not captured in the dataset.*
4. Are there other possible explanations? *As already mentioned in the response to major comment 1, events are recorded in varying degrees of detail depending on their size and extent. In the case of line closures, there may be delays between the time of the event and the documentation of the event in the database. It must also be taken into account that only events that have caused disruption or damage to operations or infrastructure are recorded in the database.*

We mainly raise the above points because for modelling studies, the outcomes of the present study may have large implications. Most modelling studies assume a deterministic relationship between rainfall, inundation and damage. The present study seems to suggest that such relationships would only explain 25% of reported flood damage events. We invite the authors to further reflect on this. Are the author's aware of any other empirical studies that looked into this relation? Did they find similar results? *As mentioned in lines 487-489, there are only few scientific studies available on analysing natural hazard-related disruptions in the transport sector based on event databases from infrastructure operators. To our knowledge, none of them used a similar approach as presented in this manuscript.*

Minor comments

- In the abstract, please list the 'three associated natural hazards' upfront. Now it takes a while before the reader knows which hazards you examined, namely: floods, gravitational mass movements / landslides?, tree fall. *We decided not to mention the three associated natural hazards namely in line 7 due to the limitation of 200 words in the abstract. The current*

abstract consists of 199 words. If the editors wish to list the hazards, we have to rewrite the abstract and shorten it in other parts.

- Line 47: and smaller tolerance of risk compared to road transport? *Disturbances on railway lines have often more impact on the traffic than on roads because of fewer alternative routes and usually longer times for restorations. A lower risk tolerance therefore makes sense from the operator's perspective.*
- Line 50: Can the authors provide any additional information on what type of damage is reported in the DB database? Do railtrack characteristics play a role? What part of the track is damaged? Could it also be damage to a pier or abutment of a bridge that supports the track? *DB does not currently have guidelines on the specific information that must be recorded when recording incidents. The level of detail is therefore dependent on the person making the report. This means that for some events there is also information on which part of the rail system is damaged (e.g. overhead line, switch), but in most cases this remains unclear. Railtrack characteristics play a role particularly in the effects, e.g. the duration of the disruption. If, for example, the overhead line is damaged by a falling tree, it usually takes much longer to restore access to the line than if such an event occurs on a non-electrified line.*
- Line 59: do you mean bias or correlation? *We mean correlation.*
- Line 64: One or two figures to illustrate the data sources could be useful for a reader that is unfamiliar with them. *See response to major comment 1.*
- Line 88: Figure 1 description indicates Monthly and yearly distribution of heavy rainfall, stating it as Yearly and monthly would better match the content of the figure. Figures c-h are also more likely to be consulted by the reader when reading the results, making the figure placement inconvenient (though understandable). Could changing the position or splitting the figure into two make it more readable? *Figure caption can be changed as proposed. The position of the figure can also be changed if wished by the editors. We do not find the division into two figures helpful, as in our opinion it is clearer if the monthly and annual distribution can be viewed directly next to each other.*
- Line 147, can you describe in a few lines how the polygon data looks like? Is it a polygon indicating a uniform amount of rainfall within that polygon? *The polygon area describes the entire area that has been issued as a heavy rainfall event by the DWD. The attributes of the polygon are standardised to the entire polygon.*
- Line 149: Punctuation can be improved for readability. *We can split this long sentence in two shorter ones for better readability.*
- Line 163: The selection of 2 days as a time window could be better explained; how often does it take longer than 2 days to record the damage events? During periods of heavy rainfall, it may not be safe to collect the data, for example? *Unfortunately, there are no official or general statements about a possible delay. However, the documented events are all events that hinder and disrupt operations or infrastructure. This also means that in case of doubt not all events are documented, but only those that directly affect the rail in operation. Because of this, the time delay can be viewed as not very large (usually a maximum of a few hours) and negligible.*
- Line 185-200: The definition of route segments could be further clarified and justified, specifically: (1) what is an "operating point"? (2) What are the implications of such wide range of lengths (140 m to 12.7 km) – is the starting point in a long stretch equally representative as that of a short one? (3) Taking 5-meter segments may be unmanageable, but why not, for example, use 1 km segments as a standard? *(1) An operating point is a*

railway system defined according to the Railway Construction and Operating Regulations (EBO). Most operating points fall into the categories of stopping point, block point or switch. Operating locations are an important measure and category in the railway industry and were therefore selected. (2) As operating points are not evenly distributed over the whole railway network, the lengths differ. This may lead to mis-representations, which are unavoidable, especially in combination with the different accuracies available for the geolocation of the natural hazard events. (3) The reported natural hazard events refer to route segments and a corresponding operation point, additional details vary. A geolocalisation on geometries or points that differ from this reference frame are difficult and prone to error.

- Line 209 (and other location): I find the term natural hazards a bit ambivalent in this context, because it could be used to indicate either the extreme rainfall event, or the flood/gravitation mass movement/ tree fall. *The term is derived from its usage in the operational context, where the damage events due to natural hazards are separated from damage due to i.e. infrastructure failure. We therefore suggest to further specify the events as „natural hazard damage event" and can adjust it accordingly in the whole manuscript.*
- Line 393: what is meant with: can be spatially overlaid. *We meant „spatially intersected" as shown in Figure 2. We can replace „overlaid" by „intersected".*
- Line 397: does this conclusion follow from the data, or from other literature, or from common understanding? *This conclusion is from common understanding, therefore no reference is added here. It is also supported from the damage data from DB, because events occur mainly on routes following valley courses or crossing low mountain ranges.*
- Line 401-402: Is flooding triggered by heavy rainfall events (Line 402) an example of the statement in the previous sentence? This can be clarified. The use of the word connections makes it sound like it is a separate idea or concept. *This is meant as an example for the statement in the previous sentence. We can replace the word „connections" by „Zusammenhänge" for clarification.*
- Line 410: This idea is not very clear. An additional sentence with a specific example may help. Something like "[…] but to establish connections between the processes through X or Y, for example, by looking at mass land movements triggered by flooding", if this was indeed the idea. *We can formulate an additonal sentence with a specific example in the style the reviewer proposed.*
- Line 451: What is meant by "the different background of the data collections"? Is it their intended or original purpose rather than their background? *„Intended or original purpose" is correct; we can replace „background" by this terminology.*