Answers to Reviewer#1 comments

We thank Reveiwer#1 for his/her comments. The reviewer comments appear in black and the answers appear in blue.

1/ This manuscript presents an innovative model regarding the risk of pluvial local flooding and the possible impacts on a traffic infrastructure, in this case a railroad line. The manuscript mainly presents the model concept and an application to the Rouen – Le Havre railway in NW France.
The contents of the paper is innovative, it is written in a well-structured and mostly clear manner and the research and the results are of high relevance for NHESS. The results of the paper are convincing! I suggest the publication with a few minor amendments and some additional explanations:
We thank Reviewer#1 for his/her positive appraisal of the paper.

2/ Page 2, line 1-5: I think that the options of using a physically based model are written in a rather pessimistic manner. I think that such a model could be applied with a similar data base and with similar results. However, the big difference would be the required time for setting-up such a model and the required computational time. But it could be run with a rather high temporal resolution instead. Maybe you could elaborate a bit more on those differences.
We agree with Reviewer#1 that the comparison of the IRIP model with other types of models is quite short. We will improve the presentation as suggested by Reviewer#1. We also underline that the positioning of the IRIP model already appears in previous publications (Lagadec et al., 2016, 2018). As the focus of the present paper is more on the evaluation methodology of such a model, we had chosen not to discuss too much the interest of the IRIP model, but we agree that it is necessary to elaborate a bit more for a reader not familiar with the IRIP model. One of the major difference between the IRIP model and hydrological models is that the model is a static model and there is no quantitative simulation of discharge (see also answer to comment 3/).

We propose to modify the sentences as follows:
"As runoff can occur everywhere on a territory, there is a need to provide maps of susceptibility to surface runoff at the scale of a whole territory or an entire transport network. Physically based distributed models may be deployed (e.g. Dabney et al., 2011; Le Bissonnais et al., 2002; Schmocker-Fackel et al., 2007, Smith et al., 1995) have the ability to provide the spatial and temporal evolution of runoff dynamics (water depth and sometimes velocity). However, they require many input data for their set up and calibration that may not be available everywhere. Thus this kind of model may be difficult to deploy on large territories. An alternative solution, called IRIP….."

3.1/ Page 3, introducing the IRIP model: Can you talk a bit on the temporal resolution. I understood it is a quasi-static model, i.e. no temporal resolution. This should be mentioned.
Reviewer#1 is right. There is no temporal resolution in the IRIP model as it is a static model. It will be explicitly stated in the revised version of the manuscript.
"An alternative solution, called IRIP for "Indicator of Intense Pluvial Runoff" was proposed by Dehotin and Breil (2011) for mapping the susceptibility to surface runoff. The model is a score method that provides 3 maps of runoff susceptibility to runoff generation, transfer and accumulation, with 6 susceptibility levels from 0 to 5. The maps are static, therefore, the IRIP model does not have any temporal resolution and the maps do not provide quantitative information about runoff dynamics. The IRIP model allows the creation of three maps representing three different phases of the surface runoff phenomenon: generation, transfer, and accumulation."

3.2/ Furthermore, please explain if (and possibly how) variability of rainfall in space is considered and how one may approach/ guess the critical rainfall intensity thresholds.
This point is partially discussed in the Discussion section 4.3 but can be further elaborated in the discussion section (see also answer to point 9/). In the methodology section, the use of a rainfall intensity threshold to define the evaluation area is presented as an alternative example to the definition of the study area from a linear transportation network.

3.3/ I also think that one should mention the question of an appropriate / meaningful spatial resolution here. You discuss this well in the discussion chapter, but you may refer here already to this discussion, because it is essential for model application.
In the revised version of the manuscript, some elements about the appropriate spatial resolution of input data will be added. Note that the IRIP model can be applied with data (in particular DTM) at various resolutions, depending on the objectives of the study. Coarse resolutions can be used when the objective is to get a broad view of sensitive-non sensitive areas over a territory. If higher resolution data are included (eg Lidar DTM data), it is possible to get more precise information and to have explicit representation of linear features such as ditches or roads. In this case, it is not necessary to provide exogenous information about road networks that are not seen by coarse resolution DTM but are detected with high resolution ones. In this case, adaptation of the model may be required.
We propose to modify the end of section 2.1 as follows, taking also into account comment 9/ of Reviewer#1:
"The resolution of the susceptibility maps retains the resolution of the Digital Elevation Model (rasterized topography map) used as input data. To determine the thresholds separating the topographic indicator values (slope and topographic index respectively) into values favorable or not to runoff, an automatic classification, the "K-mean clustering method for grids" provided in SAGA GIS was used. The third option that combines two methods: the iterative minimum distance (Forgy, 1965) and the hill-climbing method (Rubin, 1967) to divide the grid values into two classes was retained. The principle of the method is to maximize the inter-class variance, while minimizing the intra-class variance. As the classification is performed using all the grid points located in the study area, the threshold value, separating the two classes (favorable or not to runoff), depends on the study area. The IRIP model can therefore be applied to various territories without a priori local knowledge on the area, as the thresholds can be automatically computed. If local knowledge about threshold values is available, these threshold values can be specified by the user. Note that the choice of the two thresholds for the slope and topographic index has an impact on four indicators out of the 15 presented in Figure 1. Note also that If higher resolution data are included (e.g. Lidar DTM data), it is possible to get information that is more precise and to have explicit representation of linear features such as ditches or roads. In this case, it is not necessary to provide exogenous information about road networks that are not seen by coarse resolution DTM but are detected with high resolution ones. In this case, adaptation of the model may be required."

4/ Page 3, line39: Typo ("reduced" instead "reduces")
Taken into account

5/ Page 4, line 13: how to guess the "rainfall large than a specified intensity"
See discussion for more details. See also answer to comment 3/.

6/ Page 4, step 2.2.2: Can you explain why the runoff susceptibility map is not important here
We do not say that the runoff susceptibility map to runoff generation is not important, but that it is not used for characterizing the hazard level, when data of localized impacts are used for the model

evaluation. Indeed, those data are not directly related to runoff generation process but generally to transfer (erosion and water arrival) and accumulation processes (water stagnation and flooding). This highlights that there are generally no direct observation of runoff and in particular of the generation process. So the map cannot be used for the comparison with localized impact data as the latter do not characterize the generation process. It does not mean that the map has no interest, on the contrary: by pointing out where are the runoff generation areas, the map can suggest mitigation measures to retain runoff where it is produced in order to limit its further transfer and accumulation (see discussion in section 4.4).

We propose to modify step 2.2.2 as follows:

"As the IRIP model provides three maps, it also means choosing the maps that will be considered in the evaluation. The susceptibility map to runoff generation is not used for characterizing the hazard level, when data of localized impacts are used for the model evaluation. Previous experience (Lagadec et al., 2016b; 2018) showed that, when compared to localized impact data, susceptibility maps to transfer and accumulation were relevant, with the susceptibility map to transfer generally associated with erosion, and the susceptibility map to accumulation associated to sediment deposition and flooding. Note that the same area can have at the same time a high susceptibility to runoff transfer and/or accumulation and/or generation. On the other hand, impact data are not directly related to runoff generation process. Therefore, a composite of the susceptibility maps to runoff transfer and accumulation was built for defining the hazard. This map is the union of the susceptibility maps to transfer and accumulation, i.e. each pixel retains the maximum level of both maps."

7/ Page 5, line 36: I cannot see easily the thalweg structure in figure 4, can you improve this visibility?

The figure was improved in the revised version (see answer to comment 14/.

8/ Page 6, line 31: Figure 5 needs more explanations.

Reviewer#1 is right: the figure caption will be enhanced to better explain the various parts of the figure. On p.6, the reference to the figure was only to illustrate the various types of railway profiles that can be encountered and that are illustrated in the second column of the figure).

The new caption reads as follows

"Figure 5: Vulnerability tree of the railway based on expert judgment. Each column corresponds to one criteria considered when computing the vulnerability: the exposure (non-exposed sections are long tunnels or viaducts) (column 1)", the type of railway profile (column 2), the length of the section (column 3) and the existence of a singularity (either level crossings, road bridges or tunnel inlets or outlets) (column 4). The +1 in the red circles indicate that 1 is added to the vulnerability score of the section to provide the final score that appears in the last column of the figure."

9/ Page 11, line 38: Can you elaborate a bit more, if (and if yes, how) this model could account for individual rain storm events. This would be rather interesting.

Since the submission of the paper, further work has been done on analyzing past events and some elements of methodology can be added to the discussion. The approach that has been proposed – in an operational use of the model, -not in the perspective of assessing the relevance of the evaluation method like in this paper is the following.

A rain threshold triggering localized impacts can be assessed if a time series of spatialized rain and geo-referenced localized impacts are available for the same storm event. The principle consists in searching for the maximum rainfall for the different impacts, over different durations from 5 minutes to 1 hour. The durations of interest are based on the assumption that the higher the average hazard level that is located in the vicinity of the impact, the lower the amount of rain required to trigger the

impact. Initial results points to a relevant duration of 15 minutes to 1 hour. Once the duration has been selected, the minimum rainfall intensity over this duration is selected and considered as the rainfall threshold necessary to trigger all observed incidents. This assumption allows restricting the model's evaluation area to the areas where it has rained enough.
We propose to add this paragraph to the discussion.

10/ Discussion section: Did you gain any information about possible blockage of culverts / drainage pipes under the railroad track during heavy rainstorms? If yes, it would be very interesting to read about this. I have been informed about such incidences in Germany during after flash floods. Those created a big problem, when the street or railroad dams were impounded, overflown, eroded and partly broken. I think tis risk is under-estimated if not neglected at all.
Reviewer#1 is right: blockage of culverts/drainage pipes is a common problem in the railway context. In addition to blockage related to a particular intense event, progressive filling of the infrastructure by diffuse sediment transport is also a difficulty, since there is a large number of small hydraulic works that are difficult to maintain. Unfortunately, the information is rarely documented in the reports about the impacts, found in the archives. Thus, it was not possible to consider this information in the evaluation methodology. On the other hand, IRIP map of susceptibility to transfer, by highlighting areas prone to sediment transport can allow management and warning to be concentrated on these areas.
We propose to add these elements to the discussion.

11/ Discussion (or conclusions): can you add a paragraph summing the limitations of the model?
Some elements will be added in the discussion about this point. We can mention limitations related to the methodology itself such as the fact that the model does not provide quantitative estimates of runoff. The other limitation is that the produced susceptibility maps are relative to the study area, as the thresholds that divide the factors maps into areas sensitive or not sensitive to runoff are computed on the study area. Therefore, it is not possible to compare maps from two areas, and if the study area changes a little, the map will also change. There are also limitations related to the application of the model. The IRIP map of transfer of accumulation strongly depend on the DEM quality, since 3.2 and 4.2 indicators over 5 are derived from the topography. The required computing time is large when large study areas are considered or if the DTM resolution is high. Finally, there are limitations in the evaluation itself as the maps of susceptibility to runoff generation were poorly evaluated, due to the lack of appropriate data.

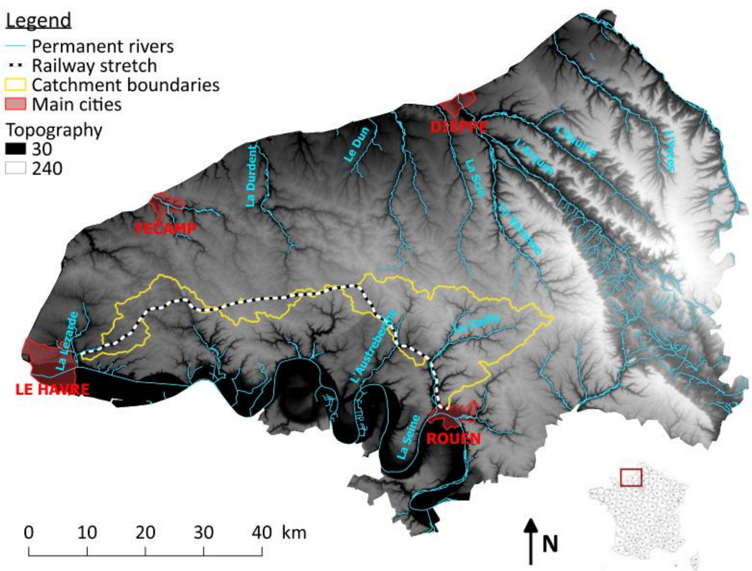
12/ Table 3: Where can one get these estimates for model parameters from? Are these the default estimates? Can you give some reasonable parameter ranges?
The parameters presented in Table 3 are either obtained using the automatic classification, either from expertise of the IRIP model application in various contexts.
IRIP can be applied everywhere without prior knowledge of runoff over the area, however, the relevance of the maps improves significantly when some parameters are adjusted after a study of the area. Some example of such adjustments are provided below:
- the drained area depending on the level of detail expected at the head of the basin, but it remains between 0.5 and 5 ha. Above 5 ha too much information is lost. In addition, localized impacts of runoff were recorded for catchments of about 1 ha.
- the break of slope: the threshold depends on the calculation method used to compute the break of slope factor, as the number of pixels over which the factor is computed can be chosen by the user. The number is always an odd number and 3 pixels is the minimum number that must be used. If the number of pixels increases, information about microtopography becomes less accurate. The number

of pixels must be adapted according to the resolution of the DTM: for instance 3 to 7 pixels are recommanded for a DTM of 25m resolution, and between 9 and 25 for a DTM of 5m resolution.
- The types of land use that are considered favorable to runoff can also be modified, for example according to the agriculture cycle on a same plot. The modification of these parameters and the evaluation of their relevance depends on the expert conducting the study, his knowledge of the area, but also on his objective (precise study, or large mesh for larger territories).
We propose to add these elements to the description of step 1 in section 2.4.

13/ Figure 2: somehow difficult to understand
The figure caption will be modified as follows:
"Figure 2: Scheme of the evaluation methodology to assess the relevance of susceptibility maps to runoff, using localized runoff-related impacts proxy data. The grey boxes indicate the information that is used in the various steps of the methodology. Yellow circles presents the various steps of the evaluation methodology leading to the final quantitative evaluation (orange box)"
14/ Figure 4: Please improve / extend this figure a bit:

- Include an inlet, where you show the location of this region within France
- Names of the river are hardly readable.
- Dry thalwegs difficult to guess.
- Rouen and Le Havre urban areas could be shown?

This figure was improved following the Reviewer's recommendations and the caption will be modified as follows:
"Figure 4: Map of the study area in Normandy (northern France). The yellow contour is the boundary of the catchments intercepted by the Rouen- Le Havre railway (line in black and white). Blue lines are the permanent river courses. One can note the dense network of dry talwegs (darker on the DEM) upstream de rivers that can be activated during a rainfall event."



References
Dabney, S.M., Yoder, D.C., Vieira, D.A.N., Bingner, R.L., 2011. Enhancing RUSLE to include runoff-driven phenomena. Hydrol. Process. 25, 1373–1390. http://dx.doi.org/10.1002/hyp.7897.
Forgy, E., 1965. Cluster Analysis of multivariate data: efficiency vs. interpretability of classifications, Biometrics, 21, 768-780.

Lagadec, L.-R., Moulin, L., Braud, I., Chazelle, B., and Breil, P.: A surface runoff mapping method for optimizing risk assessment on railways, Safety Science, 110, 253-267, 2018.
Lagadec, L.-R., Patrice, P., Braud, I., Chazelle, B., Moulin, L., Dehotin, J., Hauchard, E., and Breil, P.: Description and evaluation of a surface runoff susceptibility mapping method, Journal of Hydrology, 541, Part A, 495-509, 2016.
Le Bissonnais, Y., Montier, C., Jamagne, M., Daroussin, J., and King, D.: Mapping erosion risk for cultivated soil in France, CATENA, 46, 207-220, https://doi.org/10.1016/S0341-8162(01)00167-9, 2002.
Rubin, J., 1967. Optimal Classification into Groups: An Approach for Solving the Taxonomy Problem, J. Theoretical Biology, 15,103-144.
Schmocker-Fackel, P., Naef, F., and Scherrer, S.: Identifying runoff processes on the plot and catchment scale, Hydrology and Earth System Sciences, 11, 891-906, 2007.
Smith, R.E., Goodrich, D.C., Woolhiser, D.A., Unkrich, C.L., 1995. KINEROS – a kinematic runoff and erosion model. Comput. Models Watershed Hydrol. 20, 627–668.