The authors would like to thank the referee for his positive and constructive comments as well as for his careful lecture and detailed understanding of our manuscript. Hereafter is a point by point discussion of the referee's comments. The referee's comments are in red and our responses in black.

Minor changes will be implemented directly in the text at the position referred to in the comments. The referee also asks for a further discussion of several different points. The discussion section will hence be partly reworded/adapted to add a discussion of all the points mentioned by the referee.

General remarks:
This paper is an example of intelligent data harvesting exercise. Of course such analyses are only possible if adequate data-sets are available as in the case of the manuscript by Froidevaux et al. This makes this manuscript a very interesting and novel piece of work that surely deserves to be communicated by HESS. I miss a link to the "Large sample hydrology" literature and a sound discussion on applicability of the methodology in other areas and using alternate (e.g. European or global) datasets.
The style of the manuscript is in some parts quite unusual. It seems that the authors seek the dialog with the reader by asking many questions and stressing several times which plot and which line and which feature of a graphs is meant to answer a specific question. I know that science is driven by questions, but I have the feeling that this manuscript make way too much use of the "question style".

Issues to be addressed:

In general. This study addresses many questions. I would be interested to have an assessment by the authors on how their findings now deviate from their a-priori perceptions. E.g. Figures 5 and 6 are no real surprise with respect to my gut feeling.
Yes, figure 5 and 6 were no surprises for us neither, figure 8 (especially 8b) was more unexpected. A discussion of the results including our a priori expectations will be added in the discussion section

Abstract: The authors make large use of the "we" formulation. I think this should be avoided in an abstract.
This is a fair comment. The abstract will be reworded consequently.

Introduction: I find the introduction well formulated but I miss links to the "Large sample hydrology" literature, which is one very active community that contributed to PUB and is now contributing to the IAHS Decade "Panta Rhei". See Gupta et al. (2014) and references therein.
Thank you for mentioning this opinion paper which nicely expresses the need for "large samples" studies and hence constitutes an interesting motivation for our study. We will mention it in the introduction. We will replace the sentence in 3250-7 by:
"Instead, following the idea of large sample hydrology (e.g. Gupta, et al. 2014), we make use of two extensive networks of rain gauges and river discharge stations to derive robust statistics from an important number of catchments and events."

Concerning your goals and changes in floods and streamflow you could integrate a discussion including the works of Schmocker-Fackel and Näf (2010) and Birsan et al. (2005).
Concerning future floods there is a very recent paper by Alfieri et a l. (2015) with an analysis at the European scale.
We are aware of these three papers which focus on flood and streamflow frequency changes (observed and projected). We nevertheless wish to keep the discussion of flood frequency changes very short since it is not a focus of the study, but rather a simple motivation for the analysis. With this regard, we wish not to include these papers in the introduction and only mention the separated effects of short- and long-term precipitation on future flood frequency (as currently stated in 3249-19—26). We wish thereby to remain close to the scope of the paper.

I would also welcome a more detailed elaboration of the different response time between small and large rivers. Do you think it would be possible to include these effects in a dynamic formulation of PAP and API?
A dynamical formulation of PAPs and API would require a significant effort of calibration considering the amount of catchments considered. This is an interesting idea for a following study. The interpretation of the results will however be considerably more complex if PAPs are not equally defined for all catchments. Our opinion is rather that maintaining a simple experiment is primordial while investigating very large samples. We will discuss this point in the discussion section.

Discussion: do you think it would be possible to make such evaluation at the European-Alp scale using the data set by Isotta et al. (2014)?
Yes indeed, this is a very pertinent suggestion. The dataset of Isotta et al. offers a sufficient space and time homogeneity, a sufficient station density, as well as time coverage to serve for such an analysis. Further potential for the application of our method elsewhere (using global satellite-retrieved precipitation for e.g.) will be discussed at the end of the discussion.

3249 – 13,15: Not in Switzerland, and in other countries?
There probably exist several studies worldwide which offer "A comprehensive and systematic assessment of the relationship between precipitation and flooding". We discuss some of them in our paper (e.g. Marchi et al. 2010, Pui et al. 2011). Of course the flood-precipitation topic is widely studied. The sentence at 3249-13,15 is thus too vague to extend it to the whole world. The important message is that we did not find any study which addresses strictly and statistically the significance of the relationship between precipitation and floods involving 100 catchments, several thousands of discharge maxima and several antecedent precipitation indices. This is an important motivation for our study. The novelty of our work is however already stated explicitly at the last sentence of the introduction so that an additional sentence at 3249-13,15 is not necessary. We will hence remove that sentence.

3250 – 7,8: Which studies, please provide references.
Here we refer to the studies previously discussed like Ranzi et al. 2007, Norbiato et al. 2009, Marchi et al. 2010 and also to e.g. Merz and Boeschel (2009). These studies will be cited now at the end of the sentence.

3251-12 : Do you use the HQ5 values and HQ20 values published by the data provider or have you completed own estimations of return periods? In any case shortly describe how the return periods re computed.
We computed empirical return periods for each catchment (e.g. for a 50 years time series the highest HQ has a RP of 50 years, the second of 25 years, the third 50/3=16.6 years, etc.). We choose this computation because of its simplicity, its objectivity, and because this simple method is of sufficient precision for the purpose of our study (separate thousands of HQs in three intensity classes). Information about the calculation of the return periods will now be added at the 3251-12.

3255-13: I think you do not define "n" of Equation 1.
Indeed, this has been forgotten, thank you. The following definition will be added at 3255-15:
"K is the decay factor and n+1 is the number of days since measurements beginning"

3259-9 (an Figure 3): I think you do not clearly declare which API (API2 or API4) is now shown in Figure 3. Please update this in the text and in the caption.
API2 and API4 both result from the calculation in equ. 1. The equation is simply applied for different days *i* (2 and 4 days prior to the floods, respectively). The climatology of API2 and API4 is therefore the same and the API panel in Fig. 3 is valid for both APIs. This fact will now be stressed explicitly at line 3259-16: "API2 and API4 result from the calculation (see equ. 1). Their climatology (shown in Fig. 3c,f ) is thus the same."

3259 – 24,25: Maybe you could also include the PRE-AP terminology in Table 1.
This is a clever suggestion, thank you. We will add PRE-AP to Table 1.

3260-1,5: In Figure 4 & 5 you display the return periods of rainfall in days. This is not usual. Would it be possible to put on the left side of the RP scale also the tics for return periods of let's say 2, 5 and 10 years?
Yes, it is possible and makes sense. The graphic has been updated.
I would also remove the word "Extreme" at the beginning of the sentence "Extreme or very intense precipitation (return period > 100 days) is frequently
associated with floods".
Yes, this will be changed.
If you would use the same scales in Figures 4 & 5 would you decrease the readability of one of the two plots?
No, in fact. Thank you for this suggestion. Now both figures have the same scale.
You could merge the two Figures to have a left and right part of them. I suggest you to publish them so that Figure 4 is just left from Figure 5.
This is an interesting suggestion. We tried to bring figures 4 and 5 together but this leads to a very busy plot and to very small features. Combining the two figures does not lead to a satisfactory result, even when reducing margins and bringing the colorbar to the bottom. Consequently, and because the colorbar scale (which is now the same for fig. 4 and 5) allows for a direct comparison of the two figures, we wish to keep figures 4 and 5 separated.

3260 – 8,9: "There are more floods without intense D0-1 in Nival and Glacial regimes as compared to then Pluvial regime." Do you think this can be attributed to rain-on-snow

occurrence (see Wever et al., 2014) or this is just because high altitude areas have more shallow soils and thus the triggering event has more weight since less rain is needed to saturate the contributing areas?
Yes, we had a similar discussion internally and we think that both effects probably play a role. Our feeling is that the presence of rain-on-snow events may be the main explanation but only hypotheses can be formulated based on our results.
Both hypotheses will be shortly discussed in the discussion section.

3261 – 1,2: "We will now move on to further quantify these qualitative observations." Unnecessary sentence.
The sentence will be removed.

3261 – 23-25: Nice, please discuss this. Sometimes floods at lake exits can be also triggered (or damped) by lake regulation. Maybe you should also look at the "antecedent lake level".
Yes, we agree that the special behaviour of lake outlets deserves a specific discussion. A corresponding paragraph will be included in the discussion section.

3281: Please consider to either reduce the range from day -20 to day +20 or to make steps by 2 days. I think this would increase the readability of this figure. Or after day 15 you could make jumps in 5 days steps. Just a suggestion.
We updated the Figure to a range of +/- 20 days as suggested and this indeed improves the readability. We would continue to display each day separately so that the precipitation enhancement around day 0 can be compared to the "random" day-to-day variability.

3286: Suggestion for a reduced caption: "Figure 7. The relevance of the different precipitation periods for the occurrence of annual floods is tested using logistic regression for each precipitation period and each catchment (a) D0-1, (b) D2-3, (c) D4-6, (d) D4-14, (e) D4-30, (f) D0-30, (g) API2, and (h) API4. Several thresholds are tested (P50, P75, P90, P95, P99) and the most significant P value is displayed symbolically (squares, dots and triangles indicate a non-, weakly-, and strongly-significant influence, respectively). The colors of the symbols refer to the hydrological regimes of the catchments. Circles denote a negatively significant correlation, i.e. the exceedance of a given precipitation threshold significantly reduces flood probability." "Negative correlations are almost exclusively found in Glacial catchments." Belongs in the main text.
Thank you for this suggestion, the caption will be adapted as suggested

How do I see that different thresholds (P50, P75, P90, P95, P99) are tested? AS far as I understand here it can be that for one station the symbol is allocated by P95 in in another station by P75?
Yes, you understand it right. The most significant P-value out of 5 different tests is displayed here.
Can I ask you to make this Figure only with one threshold and avoid mixing?
Or does it help you to have a mixed-plot?
Yes it helps to have a mixed plot.
We agree that the choice of only showing the most significant result is somehow not usual. The fact that the most significant P-value is extracted out of 5 tests is however described in detail in the method section, it is stated explicitly in the figure caption, it is mentioned in the results section, and finally it is discussed in the discussion section. We therefore think that the reader has several possibilities to get reminded about this fact.
We also considered showing 5 different equivalents of figure 7 (one for each threshold) or choosing one threshold as mentioned by the reviewer. We preferred (and still prefer) the "mixed plot" option because a mixed plot is most in-line with the statements of the paper. Namely, one of the main finding of the paper is the demonstration of the absence of significance between PRE-AP and floods at most catchments. It is thus important to insure that the absence of significance is not linked with the choice of the wrong threshold (we made the special effort to test different thresholds in order to be on the safe side when stating the absence of significance). The mixed plot is for us the best solution for summarizing the complex logistic regression analysis while focussing on the most relevant information: whether there is at all at significant relationship between precipitation and floods at a particular catchment and for a particular PAP. The exact nature of this relation (whether the tests are most significant with P99 or P50 for e.g.) is never addressed in our study and is out of its scope.

3265: Section 4.5 and Figure 8 are really nice and interesting. Again, how much deviate these findings from your expectations?
Also for us, figure 8 (especially panel b) is one of the most unexpected and interesting finding of our study. We agree that this deserves a discussion including our a priori expectations. A corresponding paragraph will be added in the discussion section.

3267: Your open question “Regarding flood forecasting, it would be interesting to define a minimum threshold: what amount of event precipitation is required to trigger a HQ5 given that PRE-AP is known?” reminds me the definition of the flash flood guidance. I think it is appropriate to elaborate on this here (http://en.wikipedia.org/wiki/Flash_flood_guidance_system).
Yes, this idea is similar to the flash flood guidance (FFG) system. We propose to refer to FFG by rewording the sentence to:
“Regarding flood forecasting, it would be interesting to define a minimum threshold of precipitation following the flash flood guidance idea of e.g. Norbiato (2009): what amount of event precipitation is required to trigger a HQ5 given that PRE-AP is known?”

We would also add a sentence to emphasize that our results do not imply that FFG would not work for Switzerland (FFG would just be inefficient if based on daily precipitation sums only):

1. our analysis is based on daily rainfall sums and yearly flood peaks, but data for a flash flood guidance system (FFGS) is usually much more comprehensive (e.g. antecedent moisture for FFGS is usually based on (deterministic) soil moisture accounting models, which consider physical catchment properties and evapotranspiration much more in detail.)

2. We analysed a large number of catchments and within this large dataset (and classified subsets) we did not find evidence for a simple relationship between rainfall threshold, antecedent precipitation (very simplified model) and occurrence of 5 or 20 year floods. However, this does not mean that an FFGS approach will fail for single catchments.

Best regards
Massimiliano Zappa, WSL
13.05.2015
References:

Alfieri, L., Burek, P., Feyen, L., and Forzieri, G.: Global warming increases the frequency of river floods in Europe, Hydrol. Earth Syst. Sci., 19, 2247-2260, doi:10.5194/hess-19-2247-2015, 2015.
Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G. and Vertačnik, G. (2014), The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. Int. J. Climatol., 34: 1657–1675. doi: 10.1002/joc.3794
Schmocker-Fackel, P. and Naef, F.: Changes in flood frequencies in Switzerland since 1500, Hydrol. Earth Syst. Sci., 14, 1581-1594, doi:10.5194/hess-14-1581-2010, 2010.
Birsan MV, Molnar P, Pfaundler M, Burlando P (2005) Streamflow trends in Switzerland. J Hydrol 314: 312–329, DOI: 10.1016/j.jhydrol.2005.06.008
Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R.,Clark, M., and Andréassian, V.: Large-sample hydrology: a need to balance depth with breadth, Hydrol. Earth Syst. Sci., 18, 463–477, doi:10.5194/hess-18-463-2014, 2014.
Wever, N., Jonas, T., Fierz, C., and Lehning, M.: Model simulations of the modulating effect of the snow cover in a rain-on-snow event, Hydrol. Earth Syst. Sci., 18, 4657-4669, doi:10.5194/hess-18-4657-2014, 2014.

References:
Norbiato, D., Borga, M. and Dinale, R. (2009), Flash flood warning in ungauged basins by use of the flash flood guidance and model-based runoff thresholds. Met. Apps, 16: 65–75. doi: 10.1002/met.126