Author's response: We deeply thank Margherita Maggioni for her meaningful suggestions that greatly helped us to improve the paper. In what follows, we provide a point-by-point answer to her comments, questions and suggestions.

N. Eckert, on belief of the co-authors

Review of the paper **Development and validation using ground truth of a method to identify potential release areas of snow avalanches based on watershed delineation** by Cécile Duvillier, Nicolas Eckert, Guillaume Evin and Michael Deschâtres

General comments:

The paper addresses an issue which is well-known in avalanche science. Potential avalanche release areas (PRAs) are one of the most important parameters to be identified for avalanche hazard assessment (for ex. hazard maps, design of defense measures, etc.).

The Introduction is very rich and refers to an exhaustive literature about the PRA definition topic. Good point! And it explains clearly which are the strong and weak points of existing methods, in particular about the validation method. At the end of the Introduction, then, it clearly states the aim of the paper, which is, beside the PRA definition method, also the development of a test and validation method (see my comments in the following).

Author's response: We thank again Margherita Maggioni for her positive feedback of our work. Let us just stress here that the main objective of the paper is to develop a PRA detection method that grounds on already existing developments (notably by herself, Yves Buhler and their co-authors) and works reasonably well in the context of the French Alps. This is shown by performing an evaluation/validation exercise relying on an excellent data source regarding past avalanches, the CLPA. However, in addition to the benefit for avalanche hazard assessment in the French context, there are also some slight methodological outcomes of the paper that may be of broad relevance for the topic:

i) The determination of individual PRAs using a watershed delineation algorithm;
ii) A validation approach on the basis of accuracy scores computed using two metrics, PRA numbers and area;
iii) Broader findings and reflexions about how to validate a PRA detection method, notably how can a validation data sample be defined, and which scores can be interpreted.

None of these points are completely new in the community, but we find that they have not been fully answered so far in the literature, and we humbly hope that our paper will therefore bring some new useful elements to the debate. The questions raised by both reviews (and notably by referee two) however indicate that these objectives were not clear enough in the first version of the paper. The paper will therefore be largely reworked to better introduce the research questions and discusses the findings and the approach with regards to these questions.

See below our specific response about the validation.

Eventually, let us note here that to make the validation and parametric study more convincing, they will be largely expanded in the revised version of the paper i) in terms of potential set of parameters by conducting a much more comprehensive parametric study, ii) by considering into the analysis the questions of the DEM resolution, iii) by performing the study all over 3 entire massifs and not only over a small area, and iv) by investigating the relation between the determination of the validation sample and the accuracy scores (see additional results / figures below and in our response to referee 2). An additional small area within the massif of Chartreuse (Chartreuse study area / Dent de Crolles) will also be considered to better highlight/illustrate some results.

Concerning the PRA delineation method: I think it is a good method which takes what already exists and adds a GIS watershed routine to define single PRAs. This part is more a GIS technical issue than an avalanche science issue… but it seems that the authors found a good solution to a challenging delineation problem.

Author's response: We thank again Margherita Maggioni for her positive judgment on our work.

Concerning the PRA validation method, I think that confusion matrix and evaluation scores are a good proposal… but the weak point is the pre-processing of the CLPA data. Finally, the authors apply the same procedure developed for the PRA identification to the CLPA database to extract the release areas from the polygons of the entire avalanche extensions… Therefore, they use the same method to produce the PRAs and, in a way, to extract a dataset to validate it. I mean, *we do what we can with what we have*, but then I would not stress so much the importance of the validation method - it is even in the title and using the word "ground truth". For me "ground truth" used in relation to potential release areas should mean observed release areas… Actually, this is the common problem of all the validation methods for PRA delineation, as not many release area datasets exist. For example, also Harvey et al. (2018) made something similar in the analysis of release areas of ski triggered avalanches. In the discussion section, this problem is actually well considered and truly presented. Though, I would not use "ground-truth"...

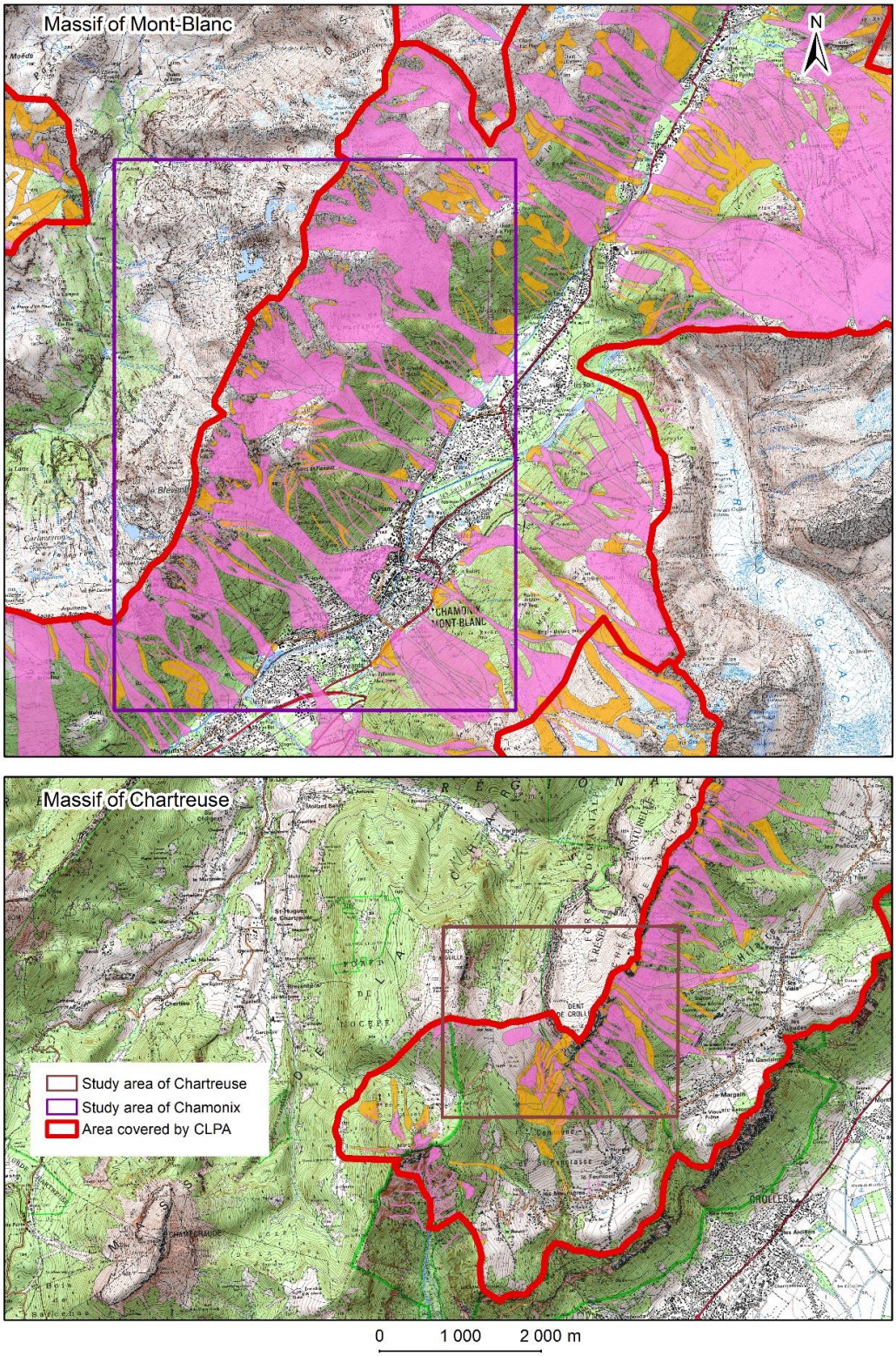
Stephan Harvey, Gunter Schmudlach, Yves Buhler, Lukas Durr, Andreas Stoffel, Marc Christen, AVALANCHE TERRAIN MAPS FOR BACKCOUNTRY SKIING IN SWITZERLAND, Proceedings, International Snow Science Workshop, Innsbruck, Austria, 2018 and presentation for the CSAW 2020 (https://www.oegsl.at/automatic-high-resolution-mapping-and-classification-of-avalanche-terrain-regarding-potential-release-triggering-and-run-out-zones/?lang=en time 3:16 - 3:22)

Author's response: We fully agree that the validation is the crucial issue. Even if, with the CLPA, we have a very valuable data support, it was for us the main source of questions and concern during the work. As said in our main comment, it is also the point on which, even if clearly we do not pretend to solve the problem, we may bring some methodological/generic outcomes/thoughts for the community. This is why we choose to focus on the validation at several points in the paper, and even in the title with the "ground truth" words. But we agree that our rationale was not clear enough and the review process helped us to formalize our thoughts as follows:

- Despite drawbacks inherent to any avalanche cadastre, the CLPA is an excellent source of information regarding past avalanches, possibly one of the finest worldwide, due to its old history, its extremely regular update by devoted technicians, continuous financial support from the French ministry of the environment and because it includes various complementary sources of information (testimonies, landscape footprints, etc.). This makes it over the years closer and closer to the true maximal avalanche prone terrain. From that perspective, it is perfectly suited to evaluate a method that aims at automatically identify the maximal avalanche prone terrain as we aim at. Notably, as CLPA extension polygons are concatenations/unions of all observed avalanche extensions on a given avalanche path, CLPA is more likely to provide an accurate estimate of the entire "ground truth" than any observation of single avalanche events. See the CLPA extracts below that will be inserted in the paper as an additional supplementary figure (response Figure 1).
- CLPA does not distinguish release areas from flow paths and runout zones, which implies that a pre-processing is required to isolate individual release areas within CLPA extensions that can be compared with our PRAs. Hence, for us, the issue is not that the CLPA validation data is not "ground truth" but that indeed the predicted PRAs and validation data are not independent (they are initially, but the pre-processing of the CLPA with the slope, forest, etc. filters introduces some dependency). However, let us say boldly that we are almost sure that obtaining a fully independent sample of "ground truth" PRA is simply not possible. Indeed, even "live", one never really observes a release area, but the full extension of an avalanche, and delineating the release area always involves some subjectivity (except, maybe, with high speed camera and films that can be watched in slow motion to see the avalanche at its earliest stage…). Also, assuming one is able to observe a "true" release area, there is little chance that the entire PRA is observed. Consequently, the definition of any validation sample will always involve some partially subjective and more or less explicit choices, with possible use of some filters (slope, etc.) similar to those we use. Our choice was to do it and to say it explicitly in a transparent manner.

- Even with the best validation sample at hand, one will never be sure that i) all potential PRAs have been spotted, ii) the maximal potential extension that can be released under the most extreme conditions has been spotted for all PRAs. As a consequence, only “true positives” can be validated, as it is never sure that a complete PRA or a part of a PRA automatically identified but not present in the validation sample is not simply missing from the validation sample. This is why we focus in our validation approach on accuracy scores/ true positives only.
- Accuracy scores (or other quantities related to confusion matrixes) were seldom used so far to evaluate PRA detection methods, to our knowledge, only in Bühler et al (2018) and with one single metric. Even if this is far from nothing, this is not much. Especially, we stress that one metric is not enough to judge the accuracy of a PRA detection method, as, e.g. the right number of PRAs can be identified but with wrong extents, and vice-versa. As a first step, we propose to evaluate accuracy scores both for PRA numbers and areas, which may cover the two most critical dimensions of the problem, but additional complementary metrics should probably be used as well in the future (focusing e.g. on the shape of PRAs, their elevation, etc.).

In the reworked paper, we will reinforce these points in the discussion and justification of the research objectives and approach to highlight the limits and outcomes of our work on PRA validation for the community. Yet, we will remove “ground truth” from the title in order to avoid any misinterpretation.
We will also remove unnecessary numbers in confusion matrixes and Tables, focusing only on accuracy scores on numbers and areas. These sum up all information related to true positives that we can decently evaluate (see response Table 1 with the complete parametric study below).
We will also further discuss the obtained scores and how they should be interpreted (see response to referee 2), with the support of an additional in-depth large scale parametric/sensitivity study that includes the effect of change on the validation sample (response Table 1, response figures 2-3). Eventually, results/discussion will include better specific consideration on how the CLPA peculiarities affect the results (supplementary Figure 3).



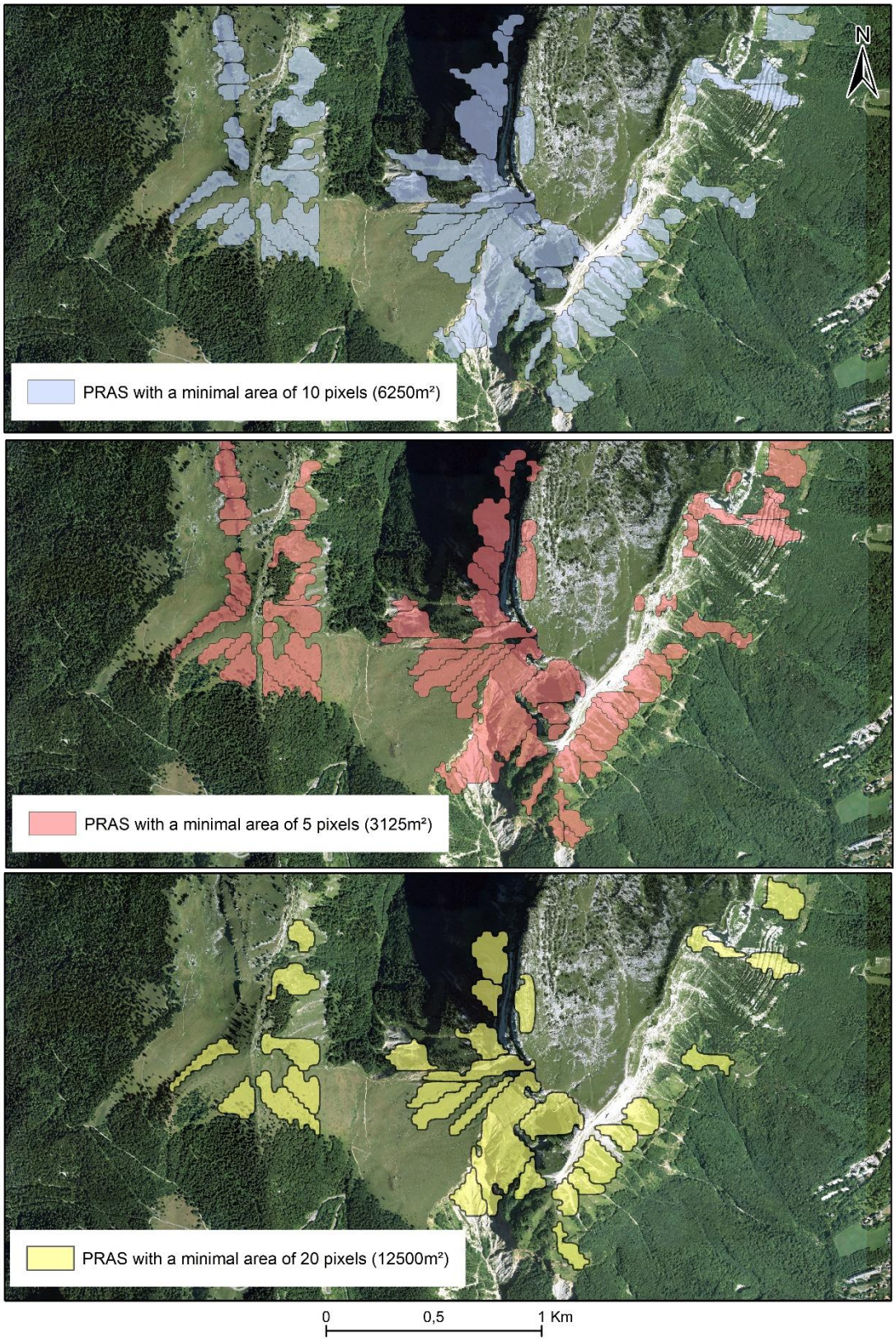
Response Figure 1: *Extracts of the Official French avalanche cadastre "CLPA" (March 2022 edition). Magenta end orange polygons correspond to the extent of past avalanches from i) testimonies and documentary sources and ii) photo-interpretation of landscape footprints, respectively. Full legend at ;https://www.avalanches.fr/static/1public/epaclpa/CLPA_feuilles_carte/CLPA_legende_carte.pdf). Small study areas of Chamonix and Chartreuse/Dent de Crolles are located, as well as the limits of the areas covered by CLPA in both massif.*

		With default values	Minimal area (m2)			Minimal elevation (m)			Slope range (°)				Maximal distance to ridge (m)			
			3125	9375	12500	1200	1600	1800	[26-60]	[30-60]	[32-60]	[34-60]	400	500	700	800
Total area of detected PRAs [km2]		90.80	93.8	88.2	85.2	90.7	88.9	88.7	90.8	64.5	58.9	51.8	81.0	88.5	94.3	95.0
Delta area with regards to default values [km2]		/	2.99	-2.57	-5.64	-0.05	-1.89	-2.14	0.00	-26.25	-31.89	-39.03	-9.80	-2.33	3.52	4.20
Delta area with regards to default values (%)		/	3.3%	-2.8%	-6.2%	-0.1%	-2.1%	-2.4%	0.0%	-28.9%	-35.1%	-43.0%	-10.8%	-2.6%	3.9%	4.6%
Total number of detected PRAs		2003	2632	1654	1369	2000	1979	1941	2002	1598	1582	1505	1877	2008	2088	2104
Delta numbers with regards to default values		/	629	-349	-634	-3	-24	-62	-1	-405	-421	-498	-126	5	85	101
Delta numbers with regards to default values (%)		/	31.4%	-17.4%	-31.7%	-0.1%	-1.2%	-3.1%	0.0%	-20.2%	-21.0%	-24.9%	-6.3%	0.2%	4.2%	5.0%
Total area of detected PRAs within CLPA extensions [km2]		84.9	85.7	83.3	81.2	84.7	83.1	76.6	84.8	61.9	56.3	49.6	75.3	82.3	87.5	87.5
Total number of detected PRAs within CLPA extensions		1601	1768	1391	1201	1590	1589	1520	1597	1409	1406	1349	1468	1576	1622	1621
Accuracy rates	In numbers	90	83.6	89.9	89.3	89.6	89.3	84.4	89.9	86.9	84.2	81.0	85.2	88.1	83.6	88.5
	In areas	96.8	95.7	96.8	96.7	96.7	96.5	91.3	96.7	96.0	94.7	92.6	95.0	96.2	92.6	96.1
Delta accuracy with regards to default values	In numbers	/	-6.4	-0.1	-0.7	-0.4	-0.7	-5.6	-0.1	-3.1	-5.8	-9.0	-4.8	-1.9	-6.4	-1.5
	In areas	/	-1.1	0.0	-0.1	-0.1	-0.3	-5.5	-0.1	-0.8	-2.1	-4.2	-1.8	-0.6	-4.2	-0.7

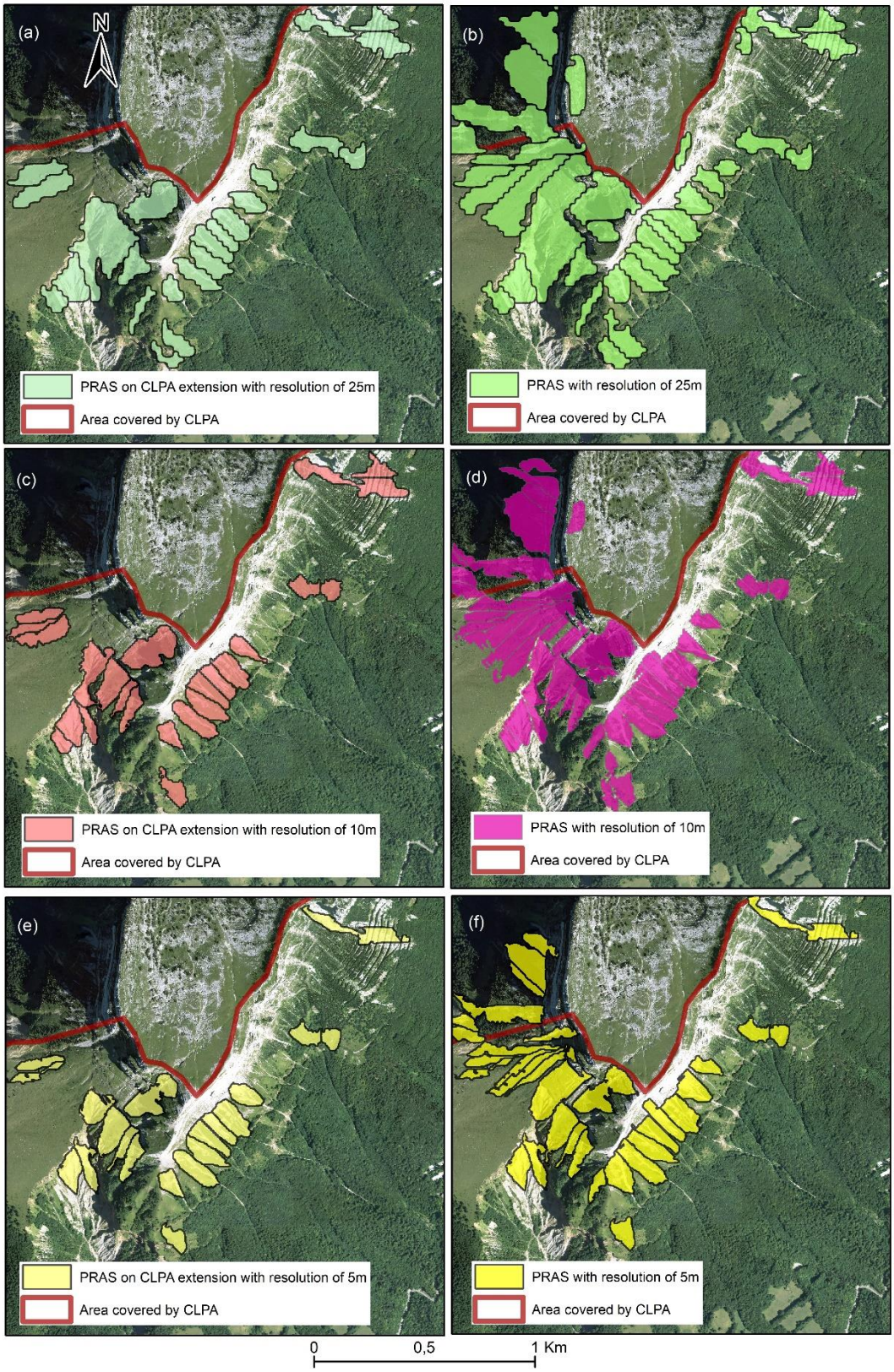
***Response Table 1**: Parametric study performed all over the Mont Blanc Massif. Total area of detected PRAs and total number of detected PRAs are those of the part of the massif covered by CLPA. The table expands information that was provided previously in Table 5 and 7 for the small Chamonix area only.*

		Chamonix area	**Chartreuse area (Dent de Crolles)**	**Chartreuse Massif**	**Mont-Blanc Massif**	**Maurienne Massif**
Accuracy rate (Eq. 3)	In numbers	92.1	93.6	93.5	90	91.4
	In areas	98.3	90.2	96.2	96.8	97

***Response Table 2:** Summary of accuracy scores for the different massifs and study areas (updates Table 3).*



***Response Figure 2**: Effect on PRA detection of the minimal area. Chartreuse/Dent de Crolles study area.*



***Response Figure 3**: Effect of DEM resolution on PRA detection (right) and selection of the validation sample (left). Chartreuse/Dent de Crolles study area. The absence of CLPA in the upper left corner is clearly visible.*

The Results and Discussion sections might be shortened… but, actually, the reader can easily follow the *file rouge* of the whole story, therefore for me it is ok like they are (but see later my comment to lines 447-463).
Author's response: The results and discussion sections of the revised paper will be reorganised in order to avoid any redundancies and incorporate the new results and discussion regarding the validation, parametric study and main findings (see previous responses).

Specific comments:
Title: maybe it could be made simpler as follows: "Development and validation of a method to identify potential release areas of snow avalanches based on watershed delineation". This title doesn't stress too much the validation method, which is not really based on ground truth data, and instead stresses the watershed delineation which is something new in the PRA definition method. Or, even, the author could give a geographical information: "Development of a method to identify potential release areas of snow avalanches based on watershed delineation and validation in the French Alps". From the abstract (very clear!) It seems that the authors develop a validation method which might be applicable also for existing PRA methods. This gives expectations to the readers…
Author's response: We agree that "Development and validation of a method to identify potential release areas of snow avalanches based on watershed delineation" may be a fair title and we will go for it in the revised version of the manuscript. See also our previous responses concerning the overall paper scope/objectives of the work and the specific issue of the methodological contribution of the paper in terms of validation of detected PRAs.

About the Introduction: I think it should finish with the description of the aim of the paper, without the rest, which seems a bit like an abstract… giving already information on which has been found after the analyses (but this section is only the Introduction!). For example, I would move lines 109-116 ("Following … individual PRAs.") to the Results or Discussion sections. And I think it is not necessary to explain how the paper is structured… It comes naturally while reading. Eventually the authors can move the sentences at the beginning of the corresponding sections.
Author's response: We will rework the end of the introduction to even better state the different research objectives of our work (see our first answers). We will also delate the material unnecessary at this stage in the revised manuscript, but we think it is good for the reader to have a hint already at this stage of the paper of the workflow and the main outcomes with regards to these objectives.

Line 28: ok for the reference to Amman and Bebi (2000) which is a general overview paper, while I would not refer to Braun et al (2020) which is a very specific one. Better, maybe, to refer to another general paper, or even a book (for ex. McClung and Shaerer, The avalanche handbook, 1993).
Author's response: We will replace the reference to Braun et al. by the reference to McClung and Shearer in the revised version of the paper.

Line 28-29: I would cancel this sentence "No countermeasure… one minute." and "therefore" at line 31. I understood the message but I think it is not necessary here.
Author's response: This sentence will be removed from the revised version of the manuscript.

It is very good how the authors state the reasons behind the choice of the different inputs (for ex. lines 143-169… and also make a kind of sensitivity analysis about this (Sect. 4.2).
Author's response: We thank Margherita Maggioni for this encouraging comment. However, following our first responses and our response to referee 2, we stress that in the revised version of the paper we will largely expand the parametric study (see our previous responses).

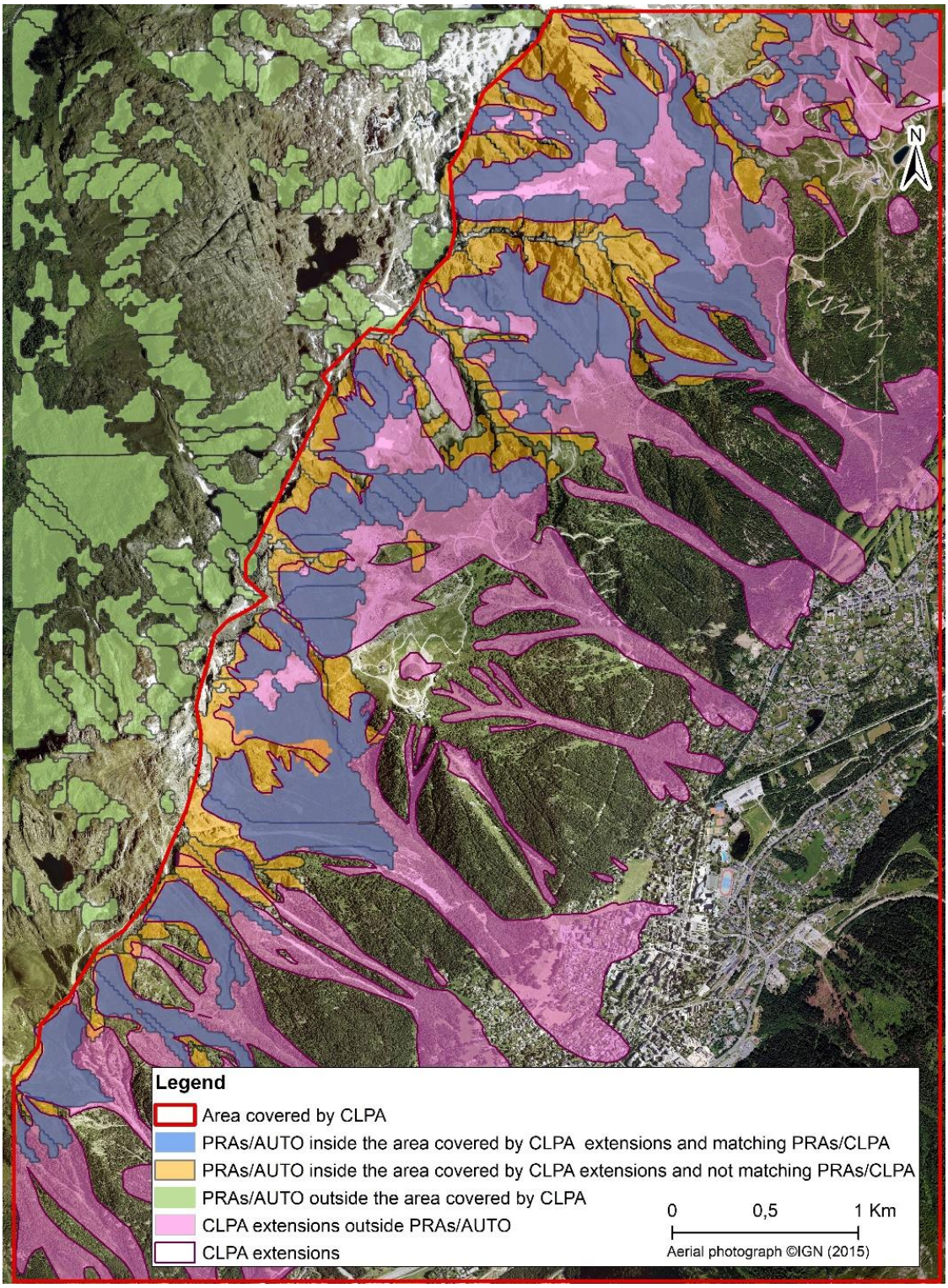
Line 181: I do not understand the reference to Figure 3… the sentence tells about areas without CLPA but the main outputs in Figure 3 are the resulting PRAs… Moreover, here we are still in Section 2. Data, therefore results should not be presented yet. I would cancel Figure 3. Instead, the authors might put

a figure showing the areas covered by CLPA within the three test areas (Mont-Blanc, Chartreuse, Maurienne).
Author's response: We will reorganise the figures and text as suggested in the revised version of the manuscript. Also we will integrate in the paper core the figures providing the full results over the three massifs (currently figures 7, 2 and S3) and add the supplementary figure providing examples of the CLPA in the small test areas of Chamonix and Chartreuse (response figure 1).

Line 232: Here it is not clear if the identification of individual watersheds is made automatically.
Author's response: It is indeed, using the existing algorithm from ARCGIS that follows the references in text and the workflow of Figure 4. This will be precised in the revised version of the paper. Note also that in the revised version of the paper, following suggestions of referee 2, we will detail the principle of the algorithm a bit more in text and move current Figure 4 to the supplements as is does not really belong to our results.

Line 259: I would specify that it is the **planar** area; also at line 265.
Author's response: These will be added in the revised version of the manuscript

Line 316: I would move here Figure 3, which in fact shows the results of the PRA definition for the area of Chamonix. I think it is not necessary to highlight the pink area as "CLPA extension outside PRAs/AUTO"; "CLPA extensions" is enough. The figure would result less messed up (see Figure 7, S2 and S3, which are clearer).
Author's response: We will reorganise the figures as suggested. The figure has been reworked as follows:



***Response Figure 4**: updated figure of the paper.*

Line 336: I would try to find a way to add a box (or a second figure) with a zoom on the area where CLPA exists (region in the lower-right corner) in order to better show the difference between matching (blue) and not matching (light blue) PRAs. Ok to put the other two figures (Mont Blanc and Maurienne) in the supplementary materials.

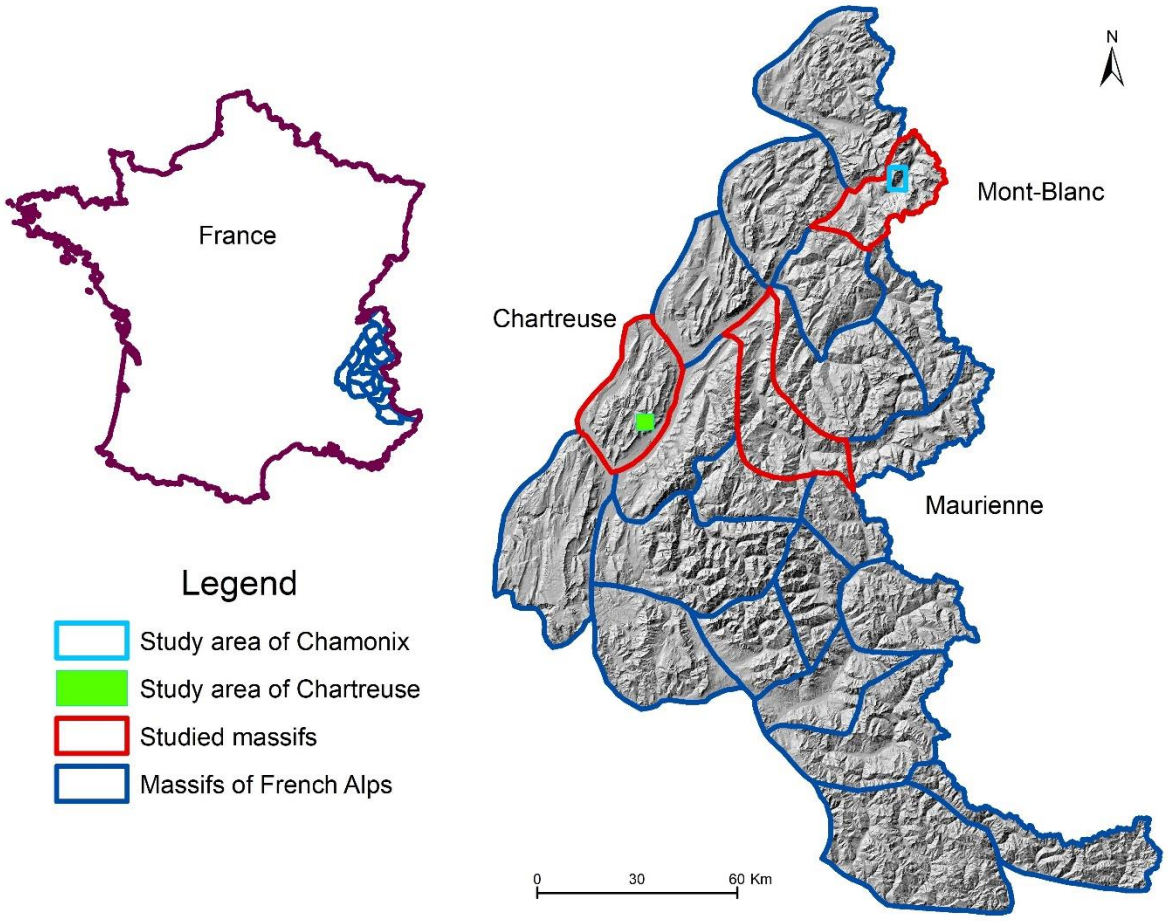
Author's response: We tried to improve the presentation of our results by adding the Chartreuse/Dent de Crolles study area (response figures 2-3 and others in our response to referee 2). Also, the figure related to Chamonix study area has been improved (response figure 4). However we want to stress

that the matching between detected PRAs and validation sample should be checked first in terms of massif-scale scores, which we will provide in a much more systematic way in the revised paper. This provides a much more fair assessment free of local effect and "cherry picking".

Line 447-463… these lines are a repetition of the Introduction… here it is time for discussion! :o)
Author's response: The whole discussion will be deeply reworked in the revised version of the paper. Regarding this specific point, we will delete unnecessary repetitions, but we think it is important before starting the discussion to remember briefly the objectives of the work and what has been achieved.

Line 520-21… I do not understand the sentence… a verb is missing?
Author's response: Indeed the sentence was "The fact that we apply the same filters to the CLPA extension and to the whole terrain on which PRAs are detected also plays a role, sorry. This will be corrected within the reworked discussion section.

Technical suggestions:
Figure 1. I would use a transparency for the violet and light blue colors to show the different study areas, so that the topography from the shaded DEM can be appreciated. **This is actually a general comment valid also for the other figures**.
Author's response: We have worked to improve the readability of the figures following the suggestions. See an example below with the reworked Fig 1 (that now also highlights the small specific study area within the massif of Chartreuse).



***Response Figure 5**: updated figure of the paper that includes the new Chartreuse/Dent de Crolles study area (response Figures 2-3).*

Figure 5: point (3) should be in *italic* and I would substitute parts with **areas**. In the second blue rectangle I would simply write **forest** instead of forest parts.
Line 43: "Wider benefits can also arise **FROM** the systematic …"
Line 327: **PRAs** instead of PRAS
Author's response: These edits will be corrected in the revised version of the manuscript.

Table 2: At the end I would add a reference to Table 1 as a legend for the Confusion matrix.
Author's response: This will be done in the revised version of the manuscript.

Concerning the English: I am not the best person to judge the quality of the English… I would probably let the paper be revised by an English native speaker.
Author's response: We agree that the English of the paper was largely improvable. In addition to changes in structures and tone of some sentences (see answer to referee 2), will the revised version of the paper will be proof-read by a native English speaker.