Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-400-AC2, 2016

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Interactive comment on "Using crowdsourced web content for informing water systems operations in snow-dominated catchments" *by* Matteo Giuliani et al.

Matteo Giuliani et al.

matteo.giuliani@polimi.it

Received and published: 10 November 2016

We would like to thank the referee for his comments. A detailed response is reported below, where authors' replies are in blue.

This paper presents an approach to supplement in situ and satellite data in snow dominated watersheds by using publicly available webcam images and flickr photographs. The authors describe a complete procedure from the crawling of the images to the application of the extracted information on the regulation policy of a reservoir lake. I enjoyed reading this paper and I concur with reviewer 1 that it deserves publication.

Printer-friendly version

Discussion paper

We thank the referee for the positive comment.

I am also left with the feeling that the authors may have somehow eluded the limitations of their approach. The discussion should provide a more balanced analysis, e.g. by discussing the computation cost and data storage issues, the minimal amount or frequency of images to reach a stable solution in the VSI, and most importantly the steps that require human intervention (see specific comments marked (A) and (B) below). I spent some time to play around with this type of data so I can imagine the tedious work and the challenges to automatically filter, align and classify webcams or photos.
Following the referee's suggestion, which was also pointed out by the first referee, we will add a more balanced discussion about requirements and limitations of the proposed approach. As far as the human intervention is concerned, it is worth noting that the requirements of our method are very low. Human intervention is indeed required only for the skyline annotation and the for setting up the experiment on Lake Como basin (e.g., select the webcam to use, ensuring it has enough information). In the revised manuscript, we will discuss in detail the main factors currently limiting our approach, especially in terms of its applicability to the entire web media content.

I encourage the authors to distribute an open source implementation of their processing to foster the development of similar applications in other regions.
We are going to release our algorithms as open source implementation. Furthermore, our intent is to transform the web platform into a unique mountain-related media repository, that would provide computer science and environmental researches not only with input data and algorithms, but also with intermediate step results (e.g., somebody interested in testing a new snow pixelwise classification method could start from already aligned and weather-filtered images).

I provided below a list of points that should be clarified. I hope that the authors

will find my comments useful and look forward to reading an updated version. (NB. the line numbering of the manuscript is awkward, maybe an issue with the Copernicus LaTeX style file)
Specific comments:
P02-L12: AMSR-E derived SWE is generally not considered as "accurate" in mountain regions. Please modify or provide a reference to justify.
The sentence will be modified as suggested by the reviewer: *Space-board passive microwave radiometers (e.g., AMSR-E) penetrate clouds but have coarse spatial resolution (25 km).*

P03-L20: I disagree that the assessment of the VSI through the Lake Como experiment is the "only viable evaluation method". There are other validation approaches, including more direct approaches like a comparison with terrestrial time lapse cameras, comparison with high resolution satellite snow maps, etc. Please clarify or remove this sentence.
The sentence was modified as suggested by the reviewer: *This form of assessment provides an indirect validation of the utility of web and crowdsourced information as the VSI extracted from general-purpose mountain images and the traditional observational data collected with dedicated tools are not comparable directly due to the difference in their physical interpretation and spatio-temporal resolution.*

P05-L19: the skyline is manually defined for a first image. Do you mean that a skyline was manually digitalized on 2000 images (see P05-L09)? If yes this should be more clearly acknowledged. (A)
We are currently running a crowdsourcing experiment for annotating all 2000 skylines as part of our effort to release a public dataset. The experiment described in the paper, instead, relies on a single webcam and required a single skyline annotation. We will emphasize this aspect in the revised version.

P05-Eq1: symbols p' and τ are not defined.
In the equation, p' is a pixel different from p and τ is a threshold on the Euclidean norm || p – p' ||. We will add the definition of both variables in the revised manuscript.

P05-L26: specify what is the edge detection algorithm.
We used the Compass algorithm (Ruzon et al., 2001), an advanced edge detector that uses color distributions. We will add this information in the revised manuscript.
Ruzon, Mark A., and Carlo Tomasi. "Edge, junction, and corner detection using color distributions." IEEE Transactions on Pattern Analysis and Machine Intelligence 23.11 (2001): 1281-1295.

P06-L09: why "cross" correlation? I would say correlation only.
We are measuring cross-correlation because we want to quantify not only the similarity between the two edge maps, but the entire set of similarities at every possible position of one w.r.t. another. Correlation alone in this case would be a mere measurement of non-causality of the two edge maps.

P06-L11: do you define a maximum offset to reduce the computation time, and if yes, how?
We do use a maximum offset of 10 pixels to reduce the computation time (and also to reduce the possible error, since the webcam trembling shifts the image not more than few pixels). The threshold was defined through a trial and error method. We will clarify this point in the revised version.

P08-L21: this is unclear to me: from the edge images, how do you extract the skyline? If this algorithm works, why was it not applied to the webcam images as well? I foresee many obstacles at this step, like the confusion of cloud edges or snow patches edges with skyline edges.
The skyline is extracted from the edge map with a modified version of the multi-stage

graph algorithm by Lie et al. (2005). This was not applied to the webcams as a single annotation was sufficient for obtaining a precise skyline extraction. As the referee correctly pointed out, the algorithm suffered from clouds and challenging meteorological conditions when applied to the user-generated photographs. To overcome this issue, we are currently working on a Convolutional Neural Network model trained on large sets of images to extract a more robust skyline. We will discuss this issue in the revised manuscript. *Lie, Wen-Nung, et al. "A robust dynamic programming algorithm to extract skyline in images for navigation." Pattern recognition letters 26.2 (2005): 221-230.*

P09-L05: what does "local refinement" mean? do you mean a locally varying transformation of the image? If yes specify the method.
The local refinement step is the application of the same edge-alignment procedure, which is first performed during the global step, with a small max radius (50 pixel) and for each mountain peak independently. This allows the peaks to slightly move in their neighborhood to better adapt to the edges. We will clarify this local refinement step in the revised version of the paper.

P09-L05 (sect 2.3): here I understand that you have used a supervised classification to get the snow mask. Then I suggest to explicit the number of samples and the method to define them. (B)
Yes, we used a supervised classifier trained on a dataset that includes 59 images manually segmented in snow/non-snow areas, ending up with more than 7 million annotated pixels. We will clarify this point in the revised manuscript.

P12-L07 (at the end of the page...): please indicate the number of webcam images and the number of flickr photos that were used for this experiment.
The experiment described in the paper was performed by using the images of a single webcam in Livigno, which ensures a continuous time series of daily images over the time horizon 2013-2014 (see Experiment Setting section). We do expect

to obtain better, and more valuable, information by using more webcams along with Flickr photos, where webcams produce a temporally dense series of images of the same view, while crowdsourced photos have better spatial distributions but lower time coverage. Yet, we did not have such data over the period 2013-2014. We will mention this analysis as a possible future research, which, hopefully, will be possible thanks to the continuous acquisition of new web content through our portal.

P14-Eq9: define r.
In the equation, r is the daily release from the lake. We will add the definition of this variable in the revised manuscript.

P16-L32: did you try to use the freezing level as an input to the regulation model?
We did not use the freezing level as argument of the operating policy because, in a previous analysis, we run an automatic selection procedure with the Input Variable Selection techniques for identifying which variables are more valuable for informing the lake operations (see the Information Selection and Assessment framework in Giuliani et al. (2015)). The results of this analysis showed that snow-information is more valuable than the freezing level: SWE was always selected as the most informative variable to be considered for improving the baseline solution, while the IVS algorithm never selected the freezing level. This result can be explained by two reasons: 1) the dynamics of freezing level is highly correlated with the seasonality and, therefore, it does not add too much information to the day of the year, which is one of the argument of the baseline policy; 2) the freezing level is independent from the amount of snow stored in the mountains and, therefore, similar values of freezing levels may be associated to the beginning of the lake inflow peak due to large snow melt as well as to lower inflow if a limited amount of snow was accumulated in the previous months. As a consequence, the freezing level is not able to provide the kind of long lead-time prediction of the volume of water that will be available in the future, which is instead captured by snow-related information.

P18-L05: I created an account and logged in to this website to give it a try but the alignment tool was not really working. The page was not responding when I clicked "continue". It might be a browser issue (I used Firefox 49 on MacOS).
We apologize for this, the problem has been fixed and we invite the referee to try it again.

P19-L09: I am not convinced with the potential of this method in the Atlas mountains because there are few operating webcams and probably a much lower amount of wintertime public photos than in the Alps.
The point is well taken. We will remove the reference to the Atlas Mountains and will better outline in the conclusions the potential limitations of the approach in catchments with few operating webcams and lower number of photos (like Atlas).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-400, 2016.

Printer-friendly version

Discussion paper