Interactive comment on "ESOLIP – estimate of solid and liquid precipitation at sub-daily time resolution by combining snow height and rain gauge measurements" by E. Mair, G. Bertoldi, G. Leitinger, S. Della Chiesa, G. Niedrist, and U. Tappeiner

General response

We thank the three reviewers for their detailed and thoughtful comments, which provide useful indications on how to improve the paper, both from a formal and from a conceptual point of view.

We appreciate the fact that all three reviewers recognized that providing new and effective methods to improve precipitation estimation at sub-daily time scales in cold and mountain areas is a very important topic for many practical and scientific applications.
This encourages us to go on with our research and to submit a revised version of the paper.

However, the three reviewer report several questions, mainly related to:

- quantify the uncertainty associated with data and instruments used for the validation of the approach (snow pillow and micro meteorological instruments) (R1 and R2);
- better evaluation of the effects on wind distribution and snow settling on the snow height data used as input in our approach (R1 R2, and R3);
- an improved analysis of the scaling issues associated with the proposed filtering of the snow height data (R1 R2,and R3);
- a better quantification of the uncertainties associated with the assumptions made in the algorithm (R1 and R2);
- a more accurate analysis of the impact of the mixed rain and rain on snow events on our method (R1 and R3);
- a reorganization of the paper structure and a reduction of its length (R1 R2 and R3).

We discuss those main issues in the following answer as well as in the in the revised paper that we aim to submit shortly.

Anonymous Referee #1

Reviewer 1 - general comments

The paper by Mair et al., proposes an approach to quantify solid and liquid precipitation from snow depth, rain gauge and other micro-meteorological observations on a subdaily time scale. Sub-daily estimates of solid and liquid precipitation are very important for hydrological purposes and any advances in this area have potential in improving hydrological prediction and process understanding. While the approach proposed is theoretically valid some assumptions and approaches taken either oversimplify (precipitation phase separation) or ignore relevant factors (wind redistribution of snow over snow pillows and under snow depth sensors). Each of the steps of the ESOLIP algorithm makes assumptions that introduce uncertainty which need to be quantified. The largest deficiency of this paper is that it does not acknowledge the use of weighing gauges (e.g. Geonor T-200B) which are standard sensors to measure total (solid and liquid) precipitation; the same parameter which this paper is trying to estimate in a much more complex manner. The uncertainty of ESOLIP makes it irrelevant at sites where weighing gauges are utilized and thus the case for using ESOLIP elsewhere

must justify its added uncertainty relative to the higher costs of a weighing gauge. No attempt at this was made. As the method was compared to improper measurements of total precipitation (a combination of rain gauge and snow pillow measurements) the paper is presenting a model (hypothesis) without proper testing. The scientific method requires proper testing which is not done in this article. A second major deficiency is in the model itself in the filtering of snow height measurements. This filtering which is crucial to the performance of ESOLIP, proposes the use of a moving–average, which has significant temporal scaling issues that are not fully addressed. Either a discussion of the scaling issues is needed or a new filtering method is needed, as hourly estimates of precipitation are invalid when using a 5 hr. moving average. Implementation of ESOLIP elsewhere may be problematic as ESOLIP thresholds may change at other locations and the large amounts of data from other sensors or observers needed will limit possible sites where this can be used. From a writing and organization perspective the paper is more of a technical rather than scientific contribution to the field and is long relative to the significance of its technical contribution. Reorganization is needed to reduce overlap as well as length (need to clearly define intro/lit review and methodology sections).

In summary I do not think that this article is currently ready for publication, recommend rejection, as ESOLIP is not tested against a storage gauge, the filtering method is inappropriate for the stated objective and the assumption that snow depth (even on the hourly interval) is solely a function of precipitation is fundamentally flawed.

Reviewer 1 - answers to general comments

R1-GC1 Oversimplification precipitation phase separation

In fact, the transition from solid to liquid precipitation occurs not at a predefined single value, but is along a range as also Harder and Pomeroy (2013) stated. By working on hourly datasets (instead of daily values) and using the wet bulb temperature (instead of only air temperature) we reduce the uncertainty considerably (see Harder and Pomeroy, 2013) compared to other approaches. Unfortunately we do not have accurate rain/snow data at hourly basis from the observer and wanted to keep simple our method. Therefore we chose the predefined value of 1°C wet bulb temperature.
The method of Harder and Pomeroy (2013) was not yet available when we analyzed our dataset. We will consider his calculation method when resubmitting the manuscript.
Regarding the missed mixed rain-snow events the Reviewer 1 wrote in his specific comments "*one option would be to use an appropriate probability of phase parameterization to calculate a mixed precipitation phase and apply that to observed precipitation*". We will consider this option when resubmitting the manuscript.

Reference:
Harder, P. and Pomeroy, JW. (2013). Estimating precipitation phase using a psychometric energy balance method. Hydological Processes DOI: 10.1002/hyp.9799

R1-GC2 Issues related to wind redistribution of snow over snow pillows and under snow depth sensors

Wind influence on the snow cover is a well-known issue in snow hydrology. In the revised paper we will better acknowledge this problem.
However, the site where the snow pillow is installed shows low wind influence throughout the year. We validated total precipitation (accumulative rain gauge / ombrometer) and snow height (snow yardstick) of the snow pillow in the winter period 2009/2010 to corroborate reliability of snow pillow

data. Moreover, problems of wind distribution would affect all devices to measure precipitation and extremely windy sites should be avoided for spatially representative measurements.
With our meteorological filtering (GSR and RH) we limit the possibility of misclassify wind-induced snow accumulation as solid precipitation events during dry weather periods.
One possibility to address this question that we will discuss in the revised paper is to introduce an additional filtering on wind speed, in order to identify potential wind-accumulation events and quantify the associated errors.

R1-GC3 Acknowledge the use of weighing gauges
"Weighing gauges (e.g. Geonor T-200B) which are standard sensors to measure total (solid and liquid) precipitation."
"Must justify its added uncertainty relative to the higher costs of a weighing gauge"

We acknowledge that an instrument as a weighing rain gauge could provide better estimation of total precipitation and we will introduce in the revised version proper citation (Geonor T-200B). When such an instrument is available and properly working, we agree that our method could be of little use. However, at least in the Italian Alps, weighing rain gauges are not such a common device, while having tipping-bucket rain gauges in the proximity of a meteorological station is a quite common situation, so we believe that an approach as ESOLIP oriented to improve precipitation estimation could be useful. Moreover, heated weighing gauges – which are highly recommended to prevent snow accumulation (Savina et al. 2012) - require considerable power supply which can hardly be covered by solar power at remote sites during winter. They cannot separate snow and rain as we are able by using ESOLIP.
Furthermore, total precipitation and snow height measurements by the snow pillow was validated by manual weighting raingauge measurements in the winter period 2009/2010. We realize that this information was missing in the article and we will add this validation data to legitimate the snow pillow as reference to measure solid precipitation (which is the main target parameter in our study).
The equipment needed by ESOLIP (a snow depth automatic sensor and a full micro-meteorological station with unheated rain gauge, air temperature and humidity, wind speed, solar radiation), on the contrary is easier to maintain and with little power consumption with respect to a weighing rain gauge. So the main aim of our approach is to make better use of the proximity of a micrometeorological station to a snow depth sensor to have better insight on the snow water equivalent. We recognize that our method cannot reduce all the uncertainties. In the revised version we will provide a more accurate quantification of the uncertainties of ESOLIP, in order to better clarify its limits of applicability.

Reference:
Savina, M., Schappi, B., Molnar, P., Burlando, P. and Sevruk, B. (2012). Comparison of a tipping-bucket and electronic weighing precipitation gage for snowfall. Atmos Res 103, 45-51. DOI: 10.1016/j.atmosres.2011.06.010.

R1-GC4 Proper test the accuracy of snow pillow observations
"A combination of rain gauge and snow pillow measurements is an improper measurement of total precipitation"

As mentioned before the snow pillow we used has been validated against manual weighing rain gauges giving good performance for a whole snow season (Leitinger et al. 2010). Moreover, trough the combined used of snow pillow and heated rain gauge we are able to separate estimation of liquid and solid precipitation.

Reference:

Leitinger, G., Obojes, N. and Tappeiner, U. (2010). Accuracy of winter precipitation measurements in alpine areas: snow pillow versus heated tipping bucket rain gauge versus accumulative rain gauge. Geophysical Research Abstracts, Volume 12.

R1-GC5 Discuss the scaling issues of the snow height filtering approach

"Either a discussion of the scaling issues is needed or a new filtering method is needed as hourly estimates of precipitation are invalid when using a 5 hr. moving average"

The main reason of choosing a simple moving average is that we want to keep the method simple for operational applications. Of course, other more complex signal filtering techniques could be applied. The challenge it to choose a method that while preserving discontinuities (i.e. snowfall events) is smoothing spurious oscillations based on spectral or variational (Vitti, 2012) methods. In the revised version we will discuss this opportunity.

When we use a moving average, of course a scaling issue is introduced, also as noticed by Reviewer 3, in terms of (1) snow height smoothing of snow fall events which are followed by rapid snow settling or snow melt, affecting snow fall totals (see Table 3) and (2) in terms of temporal smoothing (see Figures 5 and 7).

The impact of snow height total smoothing becomes clear when Table 3 is analyzed. We agree that the length of the smoothing window is a critical parameter of the procedure, and in the revised version we will make this point more clear. The purpose of this paper is to present the approach and discuss the critical parameter. Of course, an operational application of the method will require testing a more extensive dataset. Unfortunately we had not available manual snowfall observation on hourly time-scale.

Temporal smoothing is evident in Figures 5 and 6. In the revised version we will make clear to the reader that this method, since it used a 5 hour mobile averaging, is able to reproduce precipitation only up to a 6 hour time scale.

R1-GC6 Issues related to the possibility of using ESOLIP in other sites

"thresholds may change at other locations"

We recognize that thresholds may change in other locations. The aim of this paper is to propose the methodology, and discuss the threshold used and provide a validation strategy. We recognize that before a systematic application a more extensive testing of the method in other sites would be needed.

In the revised version, we will insert a table to resume the most relevant thresholds that should be validated.

"large amounts of data from other sensors or observers needed will limit possible sites where this can be used"

As explained also before, the equipment needed by ESOLIP (a snow depth automatic sensor and a full micro-meteorological station with unheated rain gauge, air temperature and humidity, wind speed, solar radiation), is the standard equipment of most of the snow survey stations of the snow and avalanche service in the Alps. So we think that there can be quite numerous opportunities to apply it. For many snow events ESOLIP can give much more accurate estimation than to use rough snow depth data or a nearby station down in the valley.

A manual observer is an additional information input, which can further reduce the uncertainty of the method but is not strictly needed.

R-GC7 Reorganization of the paper structure

"Reorganization is needed to reduce overlap as well as length (need to clearly define intro/lit review and methodology sections)"

We will consider the reviewer's suggestions (in the specific comments) to shift parts from one section into the other to reorganize and shorten the manuscript.

Reviewer 1 - specific comments

We thank the reviver for the detailed specific comments which will help us for the revised version of the paper.

If not mentioned in the general comments (if, we add "see above") we will answer here to the more substantial ones. Formal comments will be addressed only in the revised version of the manuscript"

R1 - Pg 8685 line 11: *What about weighing gauges (e.g. Geonor T-200b)? These gauges are commonly utilized and are designed for measuring both solid and liquid precipitation in all conditions (oil covered anti-freeze solution) at sub-daily times scales. There was no attempt at using such observations in the article for developing or testing the algorithm. Was there no gauge available? Are there other sites where one would be?*

Validation of proposed method to such observations (current standard) would be of importance to potential users of this algorithm.

See answer R1-GC3. We do not have such instruments at the investigated test sites. We acknowledge that for a broad validation of the method such instrument would be helpful.

R1 - Pg 8685 line 16: *Wind undercatch (and its correction) of solid precipitation needs to be addressed more completely (see Thériault et al., 2012 for an in-depth discussion). Explanation in article is overly simplified.*

See answer R1-GC2.

R1 - Pg 8685 line 19: *Wind influence is also very critical at the surface. What about blowing snow redistribution on the ground (affects snow pillows but not rain gauge). Needs to be considered here and elsewhere in the article.*

See answer R1-GC2.

R1 - Pg. 8686 line 2: *What about weighing gauges, which are designed for this??*

See answer R1-GC3.

R1 - Pg. 8686 line 7: *Why the reference to snow line and flood magnitudes. Describe this relationship in context of your method or remove.*

We will reformulate this sentence and remove this part.

R1 - Pg. 8686 line 11: *Precipitation does not equal SWE on seasonal scale (sublimation, mid winter melts modify snow cover (SWE)). This needs to be incorporated.*

We will consider this in the revised version of the manuscript

R1 - Pg. 8686 line 14-19: *Need to transition more clearly between paragraphs. A sentence to clarify that SWE from snow height is not appropriate (as snowpack has many processes acting on it over the course of the winter) for this method therefore need to use snow density instead.*

We will consider this in the revised version of the manuscript

R1 - Pg 8687 line 3: *Should move discussion of phase transition literature from 3.3 to here.*
We will consider this in the revised version of the manuscript

R1 - Pg 8687 line 8: *If you are not distributing this method across a basin then why refer to snowline elevation?*
We will reformulate this sentence in order to highlight the benefits of a better estimation of the precipitation phase (solid or liquid) for mountain hydrology.

R1 - Pg. 8687 line 20*: what about the bias/ disadvantages this method may introduce. Temperature is only one factor for the observed fluctuations. What about the sensor bouncing around in wind?*
In the revised manuscript we will deal with the bias / disadvantages this method may introduce. However, compared to fluctuations attributed to temperature change the sensor bouncing in wind is definitely of low importance due to generally low wind speed in our area and proper installation of the sensor.

R1 - Pg. 8687 line 26: *So ESOLIP requires 5 parameters thus one needs: a rain gauge, a snow depth sensor, temperature and humidity (for wet bulb calculation) sensor and anemometer to estimate total precipitation. Each sensor has uncertainty associated with it that is introduced into the method. For this method to be adopted one needs to justify that the added uncertainty is worth it versus the added cost (and lower uncertainty) of installing a weighing gauge?*
See comments to the R1-GC3

R1 - Pg. 8689 line 1: *Can you get hourly estimates when you are averaging over multiple hours? Either modify your objective or your method.*
See answer R1-G5.

R1 - Pg 8689 line 20: *What are considered outliers or incorrect values (more details)?*
Are considered here outliers not-a-numbers, negative values and very high snow increments (more than 50 cm/hour).

R1 - Pg 8690 line 9: *Have serious problems with the use of a moving average to calculate hourly values. 5 hr. moving average will not give you hourly observations (calculates a number for every hour but that number represents conditions of the 5 hour span not that hour). If one wants to calculate total precipitation on the hourly scale one needs a different method to filter data or if using moving averages to rescale the data to 6 hr. (or other length) periods. Extracting interval precipitation from smoothed observations of snow depth can lead to introduction of compounding errors (just look at the change in precip. amounts from original to 7hr moving average in table 2 and 3). 7h moving*
average precip. can be as low as half of observed hourly precip. This method has significant scaling issues that need to be addressed or changed.
See comments to the R1-GC5

R1 - Pg 8690 lines 23-26: *Sentences not clear. Is the goal 2.5 or 1%? What is the significance/ why did you pick these thresholds?*
We took those thresholds to have 97,5% (from both if only single) and 99% (if combined used) correct data.

R1 - Pg 8691 lines 8 – 15: *Very large range in values from average to max and thresholds. How effective/ sensitive is this parameter to variations then? Needs more justification.*
Table 1 shows the full range of the dataset. In the revised version of the paper we will also include the 5% quintile values to illustrate give the "frequency" of very large valued in the dataset.

R1 - Pg 8691 line 21: *How much precip was assumed to be melting snow in the rain gauge? Such a value will give an indication of how effective these filters are. Not all of this will be melting snow as filters don't take into account all possible conditions.*
This amount is roughly less than 20% of the total precipitation measured by the unheated rain gauge, which means less than 5% of the total precipitation. In the revised version of the paper we will calculate exactly this amount.

R1 - Pg 8692: line 5: *Why would this work better, explain.*
We will consider this in the revised version of the manuscript

R1 - Pg 8692 line 6: *Marks et al. 2013 found otherwise.*
This publication came out the 1st May 2013, at this date the manuscript was already in his final co-author-check and under language edition. We will consider the citation in the revised version of the manuscript

R1 - Pg 8692 line 7: *fixed threshold even when taking to account humidity (with Tw) separate phase poorly. Not a fixed point (transition range function of temporal scale) see Harder and Pomeroy 2013.*
See comments to the R1-GC1

R1 - Pg 8692 line 9-10: *with the noted issues and fluctuations in snow height sensor is 1mm (very small) an appropriate threshold. Need to justify the use of this threshold in light of the sensor uncertainty.*
Using a higher threshold would mean to lose many small events, since we use hourly calculation time step. The moving average removes the small fluctuations from the snow height signal. In the revised version of the manuscript we will evaluate the impact of using larger threshold.

R1 - Pg 8692 line 13-14: *Units*
We will consider this in the revised version of the manuscript

R1 - Pg 8692 line 17-18: *I agree that there is no easy way to consider mixed phase but mixed phase events must be incorporated (can constitute a large amount of precipitation). One option would be to use an appropriate probability of phase parameterization to calculate a mixed precipitation phase and apply that to observed precipitation.*
See comments to the R1-GC1

R1 - Pg 8692 line 21-23: *Needs more explanation. Is this saying that changing the value of the Tw threshold did not change the method performance? What range in Tw values was used to assess sensitivity? Changing the amount of rain and snow should change the performance of the method dramatically. How was performance assessed? Cumulative precipitation may not change as much but will change precipitation on the individual event timescale. Both are important if using this precipitation data for hydrological modelling.*

We use Tw = 1°C as fixed threshold as found in literature (Steinacker, 1983). In the revised version we will add a sensitivity study with respect to this threshold. See also comments to R1-GC1.

R1 - Pg 8693: *Much of the first section of results and discussion should be place in the methodology section*

We will consider this in the revised version of the manuscript

R1 - Pg 8694 line 18-20: *ESOLIP as it is using snow height measurements needs to take this into account in the filtering of snow height. Justify if not important otherwise.*

We work on snow height increments at hourly data (5h moving avg) and have thus higher snow amounts then the daily observer. However, if the effects of snow settling are considered, this value becomes realistic. See also specific comment of Reviewer 2 ***Page 8695, line 18-19***. We can reformulate this part in the revised version of the manuscript.

R1 - Pg 8695 lines 1-3: *sentence not clear.*

We can reformulate and will consider this in the revised version of the manuscript

R1 - Pg 8695 lines 1-19: *Section needs to clarified (difficult to follow). What is being taken as true precipitation (observer or hourly readout)? I am assuming that observed precipitation is correct (more correct than sensor) therefore filtering methods overestimates snowfall with respect observer. Need better justification for 5 hour averaging interval rather than balancing the errors between the hourly readout and observer. Both datasets cannot be right, need to pick one (observer) and develop a method that can replicate it well.*

We will reformulate and will consider this in the revised version of the manuscript. However Observers values should also be taken with their measurement error, which is biased toward underestimation. See also the previous comment.

R1 - Pg 8696 line 23*: If methods don't meet the requirements and you don't use them why include them in your discussion?*

We will move this part in in the Methodology.

R1 - Pg 8697 line 9-12: *sentence not clear.*

We will consider this in the revised version of the manuscript***,*** see also R1-GC4

R1 - Pg 8697 line 14: *snowfall or precipitation*

We will consider this in the revised version of the manuscript

R1 - Pg 8697 line 17: *how does a different method perform better than the snow pillow (the reference you are evaluating your method with). Need to clarify this.*
We wanted to say, that the variable density corresponds better to the snow pillow than the constant one. We will reformulate this sentence.

R1 - Pg 8697 line 24: *remove "with little snow". Irrelevant.*
We will consider this in the revised version of the manuscript

R1 - Pg 8697 line 25*: for -> of*
We will consider this in the revised version of the manuscript

R1 - Pg 8698 line 14: *what does this smoothing mean? Is there precipitation calculated when it shouldn't be or not?*
We will consider this in the revised version of the manuscript

R1 - Pg 8698 line 19-21: *Confusing. ESOLIP uses rain gauge data so why would it matter if the snow pillow does not observe the rain?*
ESOLIP estimates solid and liquid precipitation. We wanted to say that this is more informative than having only a snow pillow.

R1 - Pg 8698 line 24: *why calculating SWE and total precipitation. Is this SWE of precipitation? SWE typically is used to quantify the snowpack (it accumulation and subsequent melt along with all of the processes, compaction blowing snow etc... which modify it) rather than quantify snowfall. Objective of paper is to quantify total precipitation so why included SWE?*
We will use more consistent terminology in the revised version of the manuscript, using the terms "solid precipitation", "liquid precipitation" and "total precipitation". See also Reviewers #2 comments, which raised a similar question.

R1 - Pg 8699 line 10-11: *why can one only compare when snow pillow has no snow? Rainfall is measured with rain gauge (can remove rainfall signal (from rain gauge) from snow pillow signal) and rainfall on snow will not increase snow height (so no snow identified)*
The problem is that the snow pillow can measured also part of the rainfall, if the water stored in the snowpack, giving a delayed response to rainfall events. A clean comparison can be made only in periods with no snow over the snow pillow.

R1 - Pg. 8700 line 4: *"giving better information on snow line". Results do not substantiate this statement- not focus of paper.*
We will remove this part in the revised version of the manuscript, since it is not necessary.

R1 - Pg. 8700 line 7: *so with "loss of time resolution" what is the timescale of this method?*
The method can be applied to have precipitation information up to the 6h time-scale.

R1 - Fig 6: *Include unheated rain gauge data*
We will consider this in the revised version of the manuscript

R1 - Fig 7. *a) plot labels y axis as snow precipitation (mm) but caption refers to SWE. Clarify as snow precipitation could be confused as snow depth.*

We will consider this in the revised version of the manuscript

References: *Harder P, Pomeroy JW. 2013. Estimating precipitation phase using a psychrometric energy balance method. Hydological Processes DOI: 10.1002/hyp.9799*

Marks D, Winstral A, Reba M, Pomeroy J, Kumar M. 2013. An Evaluation of Methods for Determining During-Storm Precipitation Phase and the Rain/Snow Transition Elevation at the Surface in a Mountain Basin. Advances in Water Resources http://dx.doi.org/10.1016/j. advwatres.2012.11.012.

Thériault JM, Rasmussen R, Ikeda K, Landolt S. 2012. Dependence of Snow Gauge Collection Efficiency on Snowflake Characteristics. Journal of Applied Meteorology and Climatology 51:745-762.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 8683, 2013.

Vitti A., "Sigseg: a tool for the detection of position and velocity discontinuities in geodetic time-series" in GPS SOLUTIONS, v. 2012, (2012). - DOI: 10.1007/s10291-012-0257-9

Anonymous Referee #2

General comments:

This paper propose a method for calculating hourly total precipitation in mountain areas from unheated rain gauge measurements, snow height measurements by ultrasonic rangers and micrometeorological data gathered by automatic weather stations, based on the increasing number of such recording devices operated at high altitudes for various purposes. The measurement or estimation of total precipitation in mountain areas is a key issue for many practical and scientific applications, therefore any contribution aimed at better understanding the processes and/or at improving precipitation estimates in these areas could have significant impacts. The ESOLIP procedure proposed by the authors could be valid, in principle, for sites equipped with similar instrumentation. However, in my opinion, the manuscripts has many deficiencies, both conceptual and formal.

From the conceptual point of view, the authors do not provide adequate description of the instrument characteristics and, most importantly, they do not discuss error sources and uncertainties in measurements (e.g. the effect of wind redistribution on the readings of snow pillows and ultrasonic rangers), which are relevant because the ESOLIP approach makes use of many meteorological sensors (temperature, relative humidity, solar radiation, wind speed, snow height, rainfall), each of which has its own uncertainty. In addition, the choice of case studies looks unfortunate, because they are peculiar (e.g. too small precipitation events, not representative of typical winter or spring precipitation), affected by likely instrumentation failure (e.g. the heated rain gauge undercatch reaches 78%, which is much larger than errors reported in literature for such instruments), and use of inappropriate measurements (e.g. the snow pillow for estimating total precipitation) for testing the hypotheses. The authors do not adequately discuss the problems stemming from these choices, which significantly impact the obtained results. In many cases it is unclear how the authors optimize the threshold values, filtering methods and calculation methods for fresh snow density, to be used in ESOLIP. Insufficient elements are presented in support of the drawn conclusions. For example it is not obvious how have been established the relative humidity and global radiation thresholds, to be used for assessing the precipitation possibility, because it is not clear how did the authors assessed into which hours there was precipitation (circular reasoning), and how were these threshold combined, for example during night when the global radiation is zero.

From a formal point of view, the paper is rather long and in some cases its reading is difficult. The use of synonyms does not help the reader and a higher consistency throughout the paper would be desirable. Many concepts are implicit and should be clarified (see details in the specific comments). Moreover, it is not always understandable what has been done, with which chronological sequence and into which of the two presented study areas. Many concepts are repeated in various parts of the paper and in some cases they are not in the right part of the manuscript (e.g. methods in the results or vice versa) and need to be re-organized. Finally, an English proofreading is recommended.

To conclude I would say that this paper is not acceptable as is, and that it requires a major revision before being reconsidered for publication in this journal.

Answer to Reviewer 2 - general comments

R2-GC1 Provide an adequate description of the instrument characteristics

We will add adequate description with technical specifications in the method section of the revised

manuscript.

R2-GC2 Discuss error sources and uncertainties in measurements (temperature, relative humidity, solar radiation, wind speed, snow height, rainfall), each of which has its own uncertainty.
We will add a discussion of error sources and uncertainties in measurements in the revised manuscript.

R2-GC3 Discuss the effect of wind redistribution on the readings of snow pillows and ultrasonic rangers
This issue has been raised already by Reviewer1. See answer **R1-GC2** (wind redistribution) and R1-GC4 (snow pillow accuracy).

R2-GC4 The choice of case studies looks unfortunate
"*peculiar event have been chosen or periods when the heated rain gauge had instruments failure*"
In the revised version of the paper also the season 2012/2013 will be included.

R2-GC5 The snow pillow for estimating total precipitation is inappropriate.
This issue has bee raised already by Reviewer1. See answer **R1-GC4** (snow pillow accuracy). Furthermore we will add the information about the validation of the snow pillow against the totalizator.

R2-GC6 In many cases it is unclear how the authors optimize the threshold values:
This issue has been raised also by reviewer 1 and discussed in answer R1-GC6.
More specifically:
- **Filtering methods issues** have been discussed in the answer of Reviewer 1 (**R1-GC5**);
- **Calculation methods for fresh snow density**;
We used the methods found in literature (Brazenec 2005 and Jordan et al. 1999). Regarding the method of Hedstrom and Pomeroy (1998), which we found to be the best for our test site, the author confirmed the applicability for hourly data sets.
- **Clarify how have been established the relative humidity and global radiation thresholds**;
We took those thresholds to have 97,5% (from both if only single) and 99% (if combined used) correct data. See also answer to specific comment ***R1 - Pg 8690 lines 23-26.***

R2-GC7 The paper is rather long, there are synonyms and need to be re-organized
This issue has bee raised already by Reviewer1. See answer **R1-GC7** (paper structure).

<u>Reviewer 2 - specific comments</u>
We thank the reviver for the detailed specific comments which will help us for the revised version of the paper.
If not mentioned in the general comments (if, we add "see above") we will answer here to the more substantial ones. Formal comments are answered with "We will consider this in the revised version of the manuscript"

R2 - Page 8684, line 3*: consider substituting 'themes' (maybe too generic) with 'applications' or 'research themes'*
We will consider this in the revised version of the manuscript

R2 - Page 8684, line 8*: I think that the separation of precipitation into solid and liquid phases would be more clear than speaking about a generic winter precipitation. Try also to be consistent throughout the paper*

See also answer to Reviewer 1 ***Page 8698 line 24****.* We will use more consistent terminology in the revised version of the manuscript, using the terms "solid precipitation", "liquid precipitation" and "total precipitation".

R2 - Page 8684, line 9*: there 'is' an increasing number...*

We will consider this in the revised version of the manuscript

R2 - Page 8684, line 12 and following*: in my opinion here the author should better clarify how the ESOLIP approach works, e.g. by stating more clearly that it is based on unheated precipitation data, for example just after 'is proposed', at line 12.*

We will consider this in the revised version of the manuscript

R2 - Page 8684, line 13*: careful filtering of snow eight 'data' (or measurements)*

We will consider this in the revised version of the manuscript

R2 - Page 8684, line 16*: comparing 'the' results...*

We will consider this in the revised version of the manuscript

R2 - Page 8684, line 16*: do you mean 'solid', by using 'winter'? Or maybe solid plus liquid? Try to be more specific, also in the following of the paper*

See before ***Page 8684, line 8****.*

R2 - Page 8685, line 6: *also the increased frequency of solid precipitation is a limit at higher altitudes*

We will consider this in the revised version of the manuscript

R2 - Page 8685, line 15: *which are usually installed*

We will consider this in the revised version of the manuscript

R2 - Page 8685, line 18: *40% is not a correction factor, I guess you meant 1.40. However I think that you could write 'because underestimates may exceed 40-50%', after 'empirical factors'. Add also further references, e.g. WMO intercomparisons*

We will consider this in the revised version of the manuscript

R2 - Page 8685, line 19*: add some references in support to this statement*

We will consider this in the revised version of the manuscript

R2 - Page 8685, line 22*: why are these data more available than in the past? Maybe you can also mention here the higher time frequency of these newly available data*

In the Italian Alps, the availability of hourly or sub-hourly automatic meteorological stations data is frequent only after year 2000. We will explain this in the revised version of the manuscript

R2 - Page 8685, line 20 and following*: in my opinion this period should be re-arranged. Firstly, I would talk about the limitations (undercatch, costs, time etc...), then I would talk about the newly available technologies and higher frequency data, which provide new opportunities and so on...*
We will consider this in the revised version of the manuscript

R2 - Page 8685, line 27*: consider substituting 'less detailed' with 'lower'*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 1: *snow pillows also have technical problems, e.g. they tend to misread when ice layers form inside the snow pack. And what about snow redistribution by wind?*
We will consider this in the revised version of the manuscript; see also answers R1-GC2 and R1-GC3 on snow pillow validation.

R2 - Page 8686, line 3: *consider replacing 'correspondingly' with 'proportionally'*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 7-10: *please, try to improve this sentence, e.g. by moving the water budget concept at the beginning of the sentence. Try also to avoid the excessive repetition of 'especially', and add few references at the end of the period.*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 10*: to calculate 'the' snow....*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 11*: and therefore 'to' derive...*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 13*: these approaches 'were' not validated*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 14*: across the 'entire' hydrological year (remove yet)*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 16*: I think that this period should be completed explaining why these approaches are applied at a seasonal resolution, rather than at event- or day-scale*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 17 *(and in the following of the paper): In order to calculate 'the' SWE; an accurate estimate of 'the' snow density...*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 18: *is 'generally' assumed*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 21*: that the new snow density may vary*
We will consider this in the revised version of the manuscript

R2 - Page 8686, line 23-26: *this sentence overloads the Introduction and in my opinion it should be moved in the Methods section*
We will consider this in the revised version of the manuscript

R2 - Page 8687, line 1-12*: this part also overloads the Introduction, try to synthesize*
We will consider this in the revised version of the manuscript

R2 - Page 8687, line 13*: for measuring, or to measure*
We will consider this in the revised version of the manuscript

R2 - Page 8687, line 14-15: *snow redistribution by wind is another important source of error for ultrasonic snow height measurements*
See answer **R1-GC2.**

R2 - Page 8687, line 18: *what do you mean with 'vice versa'? Snow removal by wind and/or melt?*
Both. Increasing and decreasing of snow height. We will consider this in the revised version of the manuscript.

R2 - Page 8687, line 22-26: *please, rewrite to improve English and clarity. Maybe better also to use 'solid' and 'liquid', rather that snowfall and rainfall, and to be consistent throughout the paper.*
This is a good suggestion. We will adopt this terminology. See answer to Reviewer 1 ***Page 8698 line 24***.

R2 - Page 8687, line 27: *I think that you should explain here, even shortly, in what consists the ESOLIP procedure. Are the following methods (precipitation occurrence, density, filtering etc...) part of ESOLIP? Try to better link the sentences and the different concepts*
We will consider this in the revised version of the manuscript

R2 - Page 8687, line 28-29: *what is 'precipitation possibility'? Maybe you meant precipitation occurrence? What's the difference between meteorological data and meteorological information?*
Possibility that precipitation occurs. Meteorological data are all data series from measurement devices, Meteorological Information are noted by the observer.

R2 - Page 8688, line 3-5: *do the tested snow height filtering routine and the indicators for possible precipitation make part of the ESOLIP? If so, why not saying (generally) that the ESOLIP was tested in two different locations?*
A test of ESOLIP in two different locations is not possible, because we do not have both observer and snow pillow on both sites. Thus we choose this formulation.

R2 - Page 8688, line 9*: I would avoid starting the sentence with 'in addition'. Otherwise, I would write 'The experimental site is equipped with a meteorological station, which includes the following sensors... A snow pillow also operates at Kaserstattalm....'*
We will consider this in the revised version of the manuscript

R2 - Page 8688, line 14: *it is unclear if the mentioned working period is referred to the snow pillow or to the entire instrumental setup (meteorological station and snow pillow)*
We will consider this in the revised version of the manuscript

R2 - Page 8688, line 14-15: *I would skip mentioning here the filtering routine and the indicators for possible precipitation (maybe better indicators of precipitation occurrence, also in the following of the paper). You should simply introduce the other experimental site of Matsch Valley. And add a figure with the geographical setting of both sites!*
We will consider this in the revised version of the manuscript.

R2 - Page 8688, line 20: *at a distance of 3 km*
We will consider this in the revised version of the manuscript

R2 - Page 8688, line 22: *what are the purposes of these manual measurements? Are they operational observations e.g. in the framework of hydropower production? Provide some details on how these measurements were carried out (e.g. measurements include snow height and fresh snow or only snow height?)*
Both snow height and fresh snow observations are operated daily by an observer of the Hydrographical Office of the Province of Bolzano in the framework of their institutional activities.

R2 - Page 8688, line 24-25*: where they considered or used? Here you restart speaking about the automatic weather stations, why not completing their description before talking about manual observations? Do you refer to all the sensors or to a specific (i.e. snow height) sensor?*
It is referred to all sensors. We will reorganize this part in the revised version of the manuscript

R2 - Page 8689, line 4*: add 'as' after 'data'*
We will consider this in the revised version of the manuscript

R2 - Page 8689, line 9*: add 'actual' after 'related to'*
We will consider this in the revised version of the manuscript

R2 - Page 8689, line 13*: here 'additional data' is not so clear. It obliges a re-reading of the preceding sentence. Try to clarify what is additional data in the preceding sentence*
We will consider this in the revised version of the manuscript

R2 - Page 8689, line 17*: please be more specific when you say 'meteorological data'*
We will consider this in the revised version of the manuscript

R2 - Page 8689, line 18*: avoid reference to a figure in the heading of a section*

We will consider this in the revised version of the manuscript

R2 - Page 8689, line 20-21*: how was this carried out? Did you compare the rain gauge data with nearby weather stations? If not, how did you establish which data were outliers and which were not?*
Yes, we checked the data with two nearby stations, part of the same experiment. Unfortunately the station of Mazia (Hydrographic Office) provided not always data. See also the specific comment from Reviewer 1 ***Pg 8689 line 20.***

R2 - Page 8690, line 4*: but 'it' does not*
We will consider this in the revised version of the manuscript

R2 - Page 8690, line 5*: try to explain why there is this residual error in snow height measurements. Is it dependent on the solar radiation? Additional errors affect the snow height measurements by sonic rangers, such as sudden spikes during precipitation events and wind redistribution/preferential deposition. Did you correct also these errors?*
We corrected temperature-related error (as in the manual of the sensor described) and outliers. See also the specific comment from Reviewer 1 ***Pg 8689 line 20.***

R2 - Page 8690, line 8*: in my opinion one should not apply moving averages when aiming at obtaining hourly data*
See the comment to the scaling R1-GC5

R2 - Page 8690, line 15: *remove the parentheses and 'which was'*
We will consider this in the revised version of the manuscript

R2 - Page 8690, line 16-19: *reading the following sentences, it seems that a two-step procedure is implemented: i) check of metadata to exclude periods without precipitation, and ii) automatic exclusion of erroneous precipitation data, based on hourly measurements of relative humidity and solar radiation. If this is true, then clarify (also in Figure 1)*
Both can be done automated if observer data are prepared. But we can make 2 steps out of it. We will clarify those steps in the revised version of the manuscript

R2 - Page 8690, line 19*: remove 'from further analysis'*
We will consider this in the revised version of the manuscript

R2 - Page 8690, line 19-20*: these are results rather than methods*
We will consider this in the revised version of the manuscript

R2 - Page 8690, line 20-22*: what do you mean with 'describe the general weather conditions'? Be more specific. And what do you mean with 'allow further reduction of periods to investigate'? Maybe additional filtering, or exclusion of erroneous precipitation data from the final data series? Clarify and try to be consistent throughout the manuscript*

The main advantage of using also Observer's data or metadata, is to exclude from the analysis longer periods without precipitation, reducing therefore the uncertainties of the method in quantify snowfall events.

R2 - Page 8690, line 23-27*: how did you establish these maximum % errors?*
See comments above **Page 8688, line 24-25.**

R2 - Page 8691, line 2-3*: do these authors use meteorological data with the same time steps, i.e. hourly data (also for global radiation)?*
They used half hourly time-steps.

R2 - Page 8691, line 3-5: *how did you assess 'real' precipitation?*
By checking it with the described RH threshold.

R2 - Page 8691, line 8*: 'snowpack' in capital letters*
We will consider this in the revised version of the manuscript

R2 - Page 8691, line 6-15: *here the authors present a fixed threshold value for global radiation, to be used for excluding erroneous precipitation recordings. However the clear-sky global radiation varies during the day, and it also varies during the year. Therefore, in my opinion, the authors should explain why did they not used a normalized threshold, e.g. a given fraction of the clear-sky global radiation. And what about the night hours? Maybe they used the relative humidity data alone? It is not clear how they used the radiation data in combination with the relative humidity data, which is however reported in Table 1*
During the night only RH was used. Using a fraction of the clear-sky radiation is an improvement that could be mad to the method. This would require to add a to the procedure a calculator of the incoming solar radiation, which is a quite easy task, but it would add unnecessary complexity in the method. We will discuss also this option in the revised version of the manuscript

R2 - Page 8691, line 21-23: *here the authors introduce the use of the air temperature as an additional variable to distinguish between presence/absence of precipitation, but it is not clear how do they used it, i.e. in a qualitative or quantitative way, for example using fixed thresholds and/or in combination with global radiation. Moreover, how was managed the case of actual rainfall melting the snow accumulated at the rain gauge orifice?*
Air temperature was used in a quantitative way. If the "rainfall" was suspicious, we checked air temperature. The case of actual rainfall melting the snow accumulated at the rain gauge orifice cannot be represented by our method, which will give an overestimation of the rainfall intensity.

R2 - Page 8692, line 2: *snowline has a precise meaning in glaciology (i.e. a set of points forming the lower limit of a snow-covered area. see e.g. Cogley et al., 2010). Use snowfall limit instead, for example, or give a definition. However, in my opinion, the air temperature, or wet bulb temperature or dew point temperature say little about the position of the snowfall limit (i.e. they only could say if it is above, at, or below the elevation of the automatic weather station). Therefore I would prefer to use the term 'precipitation phase' (solid or liquid), or 'precipitation status', in this case.*

Cogley, J.G., Hock, R., Rasmussen, L.A., Arendt, A.A., Bauder, A., Braithwaite, R.J., Jansson, P., Kaser, G., Möller, M., Nicholson, L. and Zemp, M., 2010, Glossary of Glacier Mass Balance and Related Terms. IHP-VII Technical Documents in Hydrology No. 86, IACS Contribution No. 2, UNESCO-IHP, Paris. 114 pp.

Thank you very much for this suggestion. We will consider it in the revised version of the manuscript

R2 - Page 8692, line 5*: why does it work better than other methods? Give a short explanation*

The use of Tw makes the transition "window" smaller between possible rain and snow (Steinacker, (1983)

R2 - Page 8692, line 6: *why is it considered less suitable?*

The results when using Td were not promising. Td tends to be a worst indicator of the snowfall limit with respect to Tw (Steinacker, 1983).

R2 - Page 8692, line 8-9*: this is a quite low limit. Does it derive from the technical specifications of the snow height sensors? Or maybe from your tests? And what about short-term fluctuations and reading errors?*

See above in the specific comment to R1 ***Pg 8692 line 9-10.***

R2 - Page 8692, line 22*: variation of what?*

See above in the specific comment to R1 ***Pg 8692 line 9-10.***

R2 - Page 8692, line 23*: the threshold is = 1°C, not < 1°C*

We will consider this in the revised version of the manuscript

R2 - Page 8693, line 5-26*: try to improve the description of the whole section 4, in particular how it is structured and divided in sub-sections, to improve readability*

We will consider this in the revised version of the manuscript

R2 - Page 8693, line 7-14*: in my opinion this part could be improved by better specifying what has be done and where, and looking after the time sequence of calculations/assessments. This could be better specified also in the methods*

We will consider this in the revised version of the manuscript

R2 - Page 8693, line 19*: remove comma after 'allowed'*

We will consider this in the revised version of the manuscript

R2 - Page 8693, line 19-20*: the accuracy and errors of snow pillow readings should be better analyzed. Why do the authors assume that the evaluations from relative humidity and global radiation are better indicators of precipitation occurrence than the snow pillows reading? Did the authors account for possible errors in the snow pillow reading coming from snow redistribution by wind?*

See R1-GC2

R2 - Page 8693, line 23*: are you sure of this? Theoretically, heated rain gauges should melt precipitation during the events, not after them. Errors of heated rain gauges mainly stem from wind-induced undercatch (there is much literature on that), and more rarely they get occluded from intense snowfalls. Therefore the authors should better justify why the did not used heated rainfall data.*
See R1-GC2. However, we noticed the delayed melting. To not use heated rainfall data was due to literature suggestions on their high errors.

R2 - Page 8694, line 8*: according to previous statements, ultrasonic range disturbances mainly stem from air temperature fluctuations and solar radiation. However, both are low during precipitation events. In addition, there is not enough information on the magnitude of these errors and on the frequency of their occurrence and/or relationship with air temperature and solar radiation. Here it seems that the authors account for these errors using simple moving averages, but this choice should be better justified providing adequate information on sensor errors magnitude and occurrence*
The issues related to the choice of a moving average are discussed in answer **R1-GC5**.

R2 - Page 8694, line 9: Figure 3
We will consider this in the revised version of the manuscript

R2 - Page 8694, line 12*: do you mean 'notable', 'remarkable'?*
We will consider this in the revised version of the manuscript

R2 - Page 8694, line 14*: remove 'however' and start a new paragraph*
We will consider this in the revised version of the manuscript

R2 - Page 8694, line 16-18*: please rephrase, unclear and difficult to read*
We will consider this in the revised version of the manuscript

R2 - Page 8694, line 18*: add 'for these two events' after 'were found' to improve clarity*
We will consider this in the revised version of the manuscript

R2 - Page 8694, line 21*: snow height or fresh snow height? New snow height? Be consistent*
Fresh snow height automatically (difference of time steps) and daily fresh snow height manually by the observer.

R2 - Page 8694, line 21-23: *is this magnitude estimate based on authors experience, literature, data of this study, or what?*
In the revised version of the manuscript we will add here – if available – a proper citation, at the moment it is based on author experience. We are currently looking for literature on snow settling on daily basis.

R2 - Page 8694, line 23-24*: how was this carried out? Was the above-mentioned underestimation taken into account in some way?*
Yes, this is the reason why in evaluating the method we considered as target the Observer snow height data increased of about 20%.

R2 - Page 8694, line 26*: define 'close to reality'*
We will better quantify this in the revised version of the manuscript

R2 - Page 8694, line 20-28*: is this part concerning the two-events analysis or the entire period of observations? Try to clarify*
This part concerns the entire period of observations. We will clarify this in the revised version of the manuscript

R2 - Page 8695, line 10*: what means 'resp.'?*
Respectively

R2 - Page 8695, line 17-18: *did you mean 'a reduction of the same magnitude'?*
Yes, this will be reformulated in the revised version of the manuscript.

R2 - Page 8695, line 18-19: *which reasons (better 'these')? The correspondence between the two sites is totally insufficient in supporting the choice of the 5 h moving average*
The reason of the choice of the 5 h moving average the best compromise between the capability to quite effectively remove spurious oscillation in snow height data and preserve correct snowfall amount. Considering the snow-settling problem, this choice gives the best performances with respect to the observer. The fact that this is working fine also in the second site is a further confirmation.

R2 - Page 8696, line 22-23*: which are 'the requirements for SWE calculation of single events'. Not clear (also at lines 1-2 and 6-7, page 8697).*
We refer here to what written at page 8686 line 21. We will add a reminder in the revised text.

R2 - Page 8697, line 1-2*: please, provide additional information (e.g. statistics, or a figure, or a reference to results presented in the following of the paper) because Figure 4 does not support this assessment*
We refer here to what written at page 8686 line 21. We will add a reminder and more details in the revised text.

R2 - Page 8697, line 4*: provide a more complete description of the test sites, also concerning the local topography and the presence of nearby obstacles, in Section 2. Additionally, discuss the low importance of wind in density calculations, based on site characteristics*
We will consider this in the revised version of the manuscript.

R2 - Page 8697, line 9-19*: consider to complete remove this part*
We will remove this paragraph and include the essential information needed to understand the next part.

R2 - Page 8697, line 24*: why did you select a winter event with such a small precipitation amount? It's not representative of typical precipitation events during winter, I guess*

In the relatively dry climate of the Matsch Valley small snow events are quite common. Moreover, for our methodology this is a more challenging test.

R2 - Page 8698, line 1-3*: these considerations should be moved at the end of the sub-section, after the discussion of both events, taken individually*
We will consider this in the revised version of the manuscript

R2 - Page 8698, line 8*: and 'later' end?*
We will consider this in the revised version of the manuscript

R2 - Page 8698, line 19-21*: this sentence is unclear, maybe 'approach' at line 20 has to be replaced with 'snow pillow'?*
This is a mistake in the paper. Instead of approach should be written ESOLIP

R2 - Page 8698, line 21-22*: this sentence contradicts the previous one, because you don't have enough data to assess the error of the heated rain gauge (i.e. you say that the initial rain was not registered by the snow pillow). Line 22, do you mean 'air' temperatures? 2 m air temperature should be added in Figures 5 and 6. Overall, the analyzed events seem to be unfortunate choices and the conclusions which were drawn seem not to be supported by the data. In the first event, the authors should report also the precipitation recorded by the heated rain gauge after 14:00 of 28/01/2010, because it's clear that the heating system of the rain gauge was not working properly (this issue should be anticipated in the Section 2 and more properly discussed in Section 4). In the second event ESOLIP estimates cannot be compared to snow pillow recordings because, as the authors say, rainfall at the start of the event was not registered as the snow pillow was not covered by snow. Also in this case it is not clear if the heated rain gauge recorded precipitation data after the sunrise of 16/01/2010. If this event should be characteristic of spring conditions, why the authors did not present a spring event?*
In the revised version of the manuscript we will analyze also other years which can provide a greater number of more fortunate precipitation events. We had also the opportunity to check the reliability of the heated rain gauge.

R2 - Page 8698, line 27-28: *also underestimates, because the final values are quite similar*
We will consider this in the revised version of the manuscript.

R2 - Page 8699, line 1: *also in this case, both overestimates and underestimates are visible in the diagram of Figure 7*
We will improve the figure discussion and we will take in consideration also other snowfall events.

R2 - Page 8699, line 4: *I would not talk about 'underestimation' in this case, because it is clear that the heated rain gauge was malfunctioning (see e.g. the completely missed events during the winter). Underestimation here reach about 75%, which would require a correction factor of about four, hardly found in the literature for average seasonal or annual precipitation amounts.*

We will consider this in the revised version of the manuscript, looking at different years. We found for that year some problems for the heated rain gauge, which have been solved in the 2013-'13 snow season.

R2 - Page 8699, line 5: *the authors have not discussed one of the main issues in comparing 'total' precipitation estimates by means of ESOLIP with snow pillows measurements. The latest, indeed, do not record liquid precipitation occurring without snow cover, or percolating through the snow. Maybe the authors in this case included only solid precipitation events, but this is not clearly explained. In my opinion this should be detailed well upfront here and in the methods section.*
Before using the snow pillow for validation, we did a pre-processing of the snow pillow data and we used only the periods without the previous mentioned problems for validation. This will be explained better in the revised version of the manuscript.

R2 - Page 8699, line 14-15*: please, discuss the magnitude of the instrument error and compare it to the errors reported in the literature for rain gauges with similar characteristics. There is also a lack of a discussion of the difference among different seasons, which is large (Table 4)*
We will look in the description of the instruments and then report this, and also we will include more snow seasons in the analysis.

R2 - Page 8699, line 15-18*: I think that a more careful analysis of data would confirm this hypothesis, thus the word 'probably' should be avoided*
We will consider this in the revised version of the manuscript

R2 - Page 8699, line 17: *considering the snow pillow, as the reference in all season is a bit problematic, because it cannot measure liquid precipitation. How did you face with this issue?*
We measure liquid precipitation with the rain gauge, need to state this clear in the revised paper, see also R1-GC2.

R2 - Page 8699, line 17-24*: these conclusions are questionable, because of the above-mentioned problems of using the snow pillow as a reference. A certain overestimation of total precipitation by ESOLIP is desirable, indeed, to account for the underestimations of total precipitation obtained by only snow pillow readings*
See comment before!

R2 - Page 8700, line 2-4*: the determination of precipitation possibility has been assessed using relative humidity and global radiation data. Therefore, what is the link with the differentiation between rainfall and snowfall?*
We should here more clearly point out that the link is trough the use of Tw, as also said in the specific comment of R1.

R2 - Page 8700, line 8-11*: in my opinion this conclusion is not adequately supported by the data and the results provided in the paper*
We will provide better support for this conclusion in the revised version of the manuscript

R2 - Table 1: *explain e.g. in the caption the reason of the different values presented in the table (i.e. taken in the literature and compared to the present work)*
We will consider this in the revised version of the manuscript**.**

R2 - Table 2: *two precipitation events are presented, not just one (modify in the caption). Try to avoid repetition of 'data' in the second line of the caption. The second column does not report snow height increments (which are reported in columns 3 to 6, instead)*
We will consider this in the revised version of the manuscript

R2 - Figure 1*: step 2, there is also a third case, which is precipitation possible but not recorded (e.g. rain falling over a rain gauge which is obstructed by snow, or rain gauge malfunction). Did you accounted also for this possibility?*
This was not accounted for in ESOLIP, we should this mention already at the first pages of the papers.

R2 - Figures 2 and 3: *original data are poorly visible in the diagrams. Give an explanation of why moving averages are interrupted (in the captions)*
We will consider this in the revised version of the manuscript

R2 - Figure 4*: replace 'tested' with 'calculated' in the caption. Clarify the duration of the analyzed precipitation events (hourly data?)*
We will consider this in the revised version of the manuscript

R2 - Figure 5 and 6*: add air temperature data*
We will consider this in the revised version of the manuscript

Anonymous Referee #3

This paper tackles the very important and interesting issue of sub-daily precipitation measurements in mountainous terrain. The approach that has been chosen relies on a combination of various measurements in order to reconstruct precipitation time series.
This could prove quite useful, unfortunately this paper fails in several aspects. First, the choice of cumulated snow fall as a validation parameter for some of the validation periods is not appropriate. Because such measurements highly depend on the measurement protocol (bringing in the effects of snow settling), they do not provide any kind of a reference. The authors acknowledge that a large uncertainty is expected and guess a large correction factor, making all validation pointless. Instead, since Snow Water Equivalent measurements are available, all validation should have been performed on SWE and estimating the impact of filtering should have been performed on SWE instead of the cumulated snow fall.
Then, the measured snow height is at the very core of the method. Unfortunately, these measurements need to be filtered in order to be usable and some very important aspects of the filtering are not exposed. Did the authors had to remove some obvious outliers, and how? The filtering relies on a Window Moving Average, but such a filtering introduces a phase (i.e. the signal is shifted in time). How was this phase removed? What was the centering of the time window? A centered window would not introduce any phase but would not handle data points too close to data gaps. The authors state that a five hours window gives the best results, but don't discuss the smoothing that happens on the signal, further reducing the time correlation between the filtered snow height signal and other measurements. Also, snow settling seems not to be taken into account, since in this model, only increases of snow height produce precipitation.

Finally, the paper lacks clarity and should be further proof-read and reorganized: the data pre-processing is spread over numerous sections, making it hard to follow what really happens. Quite a few sentences should be rephrased and the overall length should be reduced while some keys elements should be more detailed: as previously mentioned, the Window Moving Average should be better explained, the snow settling should be covered as well as rain on snow events. A few elements are also surprising:
When the authors mention on page 8697, line 17 that using a variable density in their model performs better than using a snow pillow for estimating SWE or when in table 1 using a more stringent criteria leads to less points being filtered out. Some citations also appear as not really necessary and should be removed.

Overall, I recommend this paper to be rejected and encourage the authors to address the identified issues, making sure that their claims are backed by robust measurements, before eventually resubmitting.

Answer to Reviewer 3 - general comments

R3-GC1 The approach that has been chosen could prove quite useful, unfortunately this paper fails in several aspects.
We thank the reviver for considering the approach proposed useful. We will address the

shortcomings in the revised paper.

R3-GC2 The choice of cumulated snowfall as a validation parameter for some of the validation periods is not appropriate. Because such measurements highly depend on the measurement protocol (bringing in the effects of snow settling), they do not provide any kind of a reference.
For validation purpose we used SWE – when such observations were available from snow pillow. On the validity of the snow pillow data, see also R1-GC1 and R1-GC4. However, for deciding the optimal filtering interval width in order to capture snowfall events from automatic snow depth observations, the choice to use Observer's snow height measurement could be appropriate, when the effect of the snow settling has been taken in account. On this last point see also answer to Reviewer 2 comment ***Page 8694, line 21-23.***

R3-GC3 Issues in filtering the measured snow height.
This issue has been raised also by Reviewer 1 and discussed in answer **R1-GC5**.
More specifically:
- Did the authors had to remove some obvious outliers, and how?
Besides negative values here we can mention very high snow increments (more than 50 cm/hour), -999 recordings, it was done with an "if-Formula" in excel.
- How was the phase change of filtering removed?
We used a centered moving average.
- What was the centering of the time window?
A centering time window was used. To get data for one hour, the value of the hour itself, of two hours before and two hours after were averaged.
- Discuss the smoothing that happens on the signal
By using this moving average – of course – maxima were reduced depending on the range of these 5 values. See answer **R1-GC5.**
- Snow settling seems not to be taken into account
Snow settling into one hour was not taken into account, but it was shown how big the settling influence is on a day (when compared the hourly accumulated increments to daily values of the observer). See answer to Reviewer 2 comment ***Page 8694, line 21-23.***

R3-GC4 Rain on snow events should be covered
This issue has bee raised already by Reviewer1. See answer **R1-GC1.**

R3-GC5 The authors mention on page 8697, line 17 that using a variable density in their model performs better than using a snow pillow for estimating SWE
This was not written clearly in the paper. We wanted to say, that the variable density corresponds better to the snow pillow than the constant one. We will consider this in the revised version of the manuscript

R3-GC6 In table 1 we mention using a more stringent criteria leads to less points being filtered out.
This was misunderstood. We will reformulate. However, the percentage of missed precipitation is decreasing when the criteria becomes more stringent.

R3-GC7 the paper lacks clarity and should be further proof-read and reorganized.
This issue has bee raised already by Reviewers 1 and 2. See answer **R1-GC7** (paper structure).
More specifically we will take in account the following suggestions:
- "The data pre-processing is spread over numerous sections, making it hard to follow what really happens.
- Quite a few sentences should be rephrased and the overall length should be reduced while some keys elements should be more detailed.
- Some citations also appear as not really necessary and should be removed"

Reviewer 3 - specific comments

We thank the reviver for the detailed specific comments, which will help us for the revised version of the paper. We answer here to the more substantial comments. Formal comments are answered with "We will consider this in the revised version of the manuscript"

R3 - Abstract, line 1: *measuring precipitation is demanding in itself, not only for hydrological applications!*

We will consider this in the revised version of the manuscript

R3 - abstract, lines 10-11*: please rephrase*

We will consider this in the revised version of the manuscript

R3 - abstract, lines 17-18*: please rephrase*

We will consider this in the revised version of the manuscript

R3 - page 8686, lines 5-6*: it might be better to be more specific: "especially in small, steep catchments"*

We will consider this in the revised version of the manuscript

R3 - page 8688, line 20*: "at a distance of 3km"*

We will consider this in the revised version of the manuscript

R3 - page 8690, lines 20-21: *where is the additional data coming from (i.e. how far from the original station)*

It is at a distance of 3km, see 8690 line 20.

R3 - page 8690, lines 23-24: *please rephrase*

We will consider this in the revised version of the manuscript

R3 - page 8690, lines 25-26*: it is not clear what these 1% refer to. Is it compared to a reference precipitation measurement or the original data set?*

It is compared to the original dataset; only 1% of "true" precipitation should be missed.

R3 - page 8691, lines 21-23: *what happens for points passing one criteria and not passing the second one?*

Only if both conditions did NOT allow precipitation, the precipitation possibility was set to 0.

R3 - page 8692, line 20: *such events will be underestimated!*

We will consider this in the revised version of the manuscript

R3 - page 8693, line 13*: I am uneasy about the usage of "metadata" since the authors actually talk about extra meteorological data*

We will be more specific in the revised version of the manuscript

R3 - page 8693, line 19*: extra comma to be removed*
We will consider this in the revised version of the manuscript

R3 - page 8695, line 3*: the table 3 should be fully redone. Using an unreliable measurement for validation is not right!*
We will consider this in the revised version of the manuscript

R3 - page 8695, line 16*: "underestimates snowfall by 18-21%"*
We will consider this in the revised version of the manuscript

R3 - page 8695, line 21: *"calculated with the four methods"*
We will consider this in the revised version of the manuscript

R3 - page 8696, lines 17-24*: please rephrase*
We will consider this in the revised version of the manuscript

R3 - page 8697, line 1*: please rephrase*
We will consider this in the revised version of the manuscript

R3 - page 8697, line 1*: the fresh snow density parameterizations that have been rejected show very reasonable values. It is not clear why they got rejected (see also line 7), or based on which criteria*
The reason is explained at page to 8686 line 21. We will be more specific in the revised version of the manuscript.

R3 - page 8697, lines 22-23: *please rephrase*
We will consider this in the revised version of the manuscript

R3 - page 8698, line 20: *"of both rain gauge and model is likely"*
We will consider this in the revised version of the manuscript

R3 - page 8699, lines 8-12*: please rephrase*
We will consider this in the revised version of the manuscript

R3 - page 8700, lines 2-3*: please rephrase*
We will consider this in the revised version of the manuscript

R3 - page 8707, table 4*: is there any difference between SWE and total precipitation at the chosen test sites?*
Yes! SWE is only snow, total precipitation is rain and snow!

R3 - page 8711, figure 4*: consider using initials as labels instead of numbers. Moreover, identifying on the plot which parameterizations will be used afterward would be a plus. Finally, it is not clear why the first two parameterizations have been selected, since they show a significant number of outliers events (if I understand correctly the meaning of the red crosses)*

We should point out here that, as explained at page 8686, line 21, we want a formula which can include also high densities of fresh fallen snow. So, in this case, having a wider range of fresh snow values means that those parameterization have a better capability to cope with wet snow events.

R3 - page 8714, figure 7*: The explanations for the right panel are unclear. Is adding rainfall from the unheated rain gauge to snow pillow estimates not double-counting precipitation? And for the heated rain gauge, does it simply means that delayed melt is accounted for?*

We have to be clear here, ESOLIP results have been shown one time only for solid precipitation (SWE) and one time with both liquid and solid (total precipitation). The unheated rain gauge data is NOT added to the snow pillow. On the left panel only possible snowfall is considered (Tw < 1°C), on the right panel all possible precipitation events are considered.