

## ***Interactive comment on “The evaluation of the potential of global data products for snow hydrological modelling in ungauged high alpine catchments” by Michael Weber et al.***

**Michael Weber et al.**

franziska.koch@boku.ac.at

Received and published: 6 November 2020

Response to Anonymous Referee #3

Comments on the manuscript: “The evaluation of the potential of global data products for snow hydrological modeling in ungauged high alpine catchments” by Weber et al. This manuscript examines the sensitivity of the simulated snow characteristics in a small alpine basin related to meteorological input data and DEMs using a hydrological model with a sophisticated snow module. Studies like this are of a great importance as high elevation snowpack critically affects alpine ecosystems and water resources in a number of regions in the world. The experiment is comprehensively designed and well

C1

executed. The paper is informative and interesting.

Author’s answer: We would like to thank the reviewer for thoroughly reading the manuscript and for providing us with very valuable suggestions for improvement. Moreover, we thank him/her for mentioning the importance of this work regarding improvements in the investigation of high alpine water resources which is relevant for numerous (ungauged) catchments worldwide. We have carefully considered all of his/her comments and address them point by point in the following.

General comments:

(1) The analysis of the experimental results and the organization as well as the presentation of the analysis output need improvements,

Author’s answer: In general, we will reorganize the paper structure, which was also mentioned as an issue by reviewer 2 and we will better streamline and improve the presentation and analysis of the results. The paper will also be improved with an eye on brevity. As you make some clear suggestions on these points in the specific comments below, please also consider our answers to those points in the following.

(2) Insufficient evaluation of the reference simulation across entire HRUs is also problematic. Because of the large elevation range within the test basin (RCZ), the snow field within the basin is expected to vary widely. Without the evaluation of the elevation dependent model performance with the reference data, the accuracy of the reference run cannot be well established.

Author’s answer: We agree with you that the snow cover in the basin can vary widely as it is the case in any high alpine region with complex topography and that it would be valuable if it was possible to evaluate the accuracy of the elevation dependent snow cover variation in more detail. We will discuss this point in more detail in the updated version. For the study time period 2000 to 2010, however, we unfortunately have no LiDAR data to compare the modelled with the measured snow depth distribution. How-

C2

ever, as we found out in our former study for the RCZ (Weber et al., 2020) and which is also known from other studies (e.g. Grunewald et al. 2013), the spatial distribution in various years and also during the most of the season does not change much (except at the very beginning and the very end of the season). In the study of Weber et al. (2020), we were able to include information of several LiDAR snow depth data sets derived in the years 2015 and 2016 which were conducted for the RCZ to delineate HRUs that guarantee for optimal representation of the snow cover distribution in RCZ. This HRU delineation was, as mentioned in the manuscript, also used for this study. We are thereafter confident that the spatial distribution of the snow cover is quite well simulated in the reference setup. Moreover, we think that the availability of two snow depth gauges in such a small catchment is already a lot and we are quite lucky with the situation in this well gauged high alpine catchment. Of course, it would be nice if each of the ten HRUs was represented with a snow depth gauge; however, this is not available and will most probably not be found in such a potential station density anywhere in the world. Thereafter, the only possible accuracy analysis for this study period is the investigation of snow depth with data measured at the DWD and LWD station which are situated at least at different altitudes. Moreover the locations of these stations were chosen from the two operating institutions to be as representative as possible for the respective altitude, e.g. no shadowing against wind but also no particular exposure, relatively flat terrain, etc.

(3) The lack of the analysis and evaluation on monthly time scales (i.e., annual-cycle resolving analysis). This is important because the forcing and snow fields undergo strong seasonal cycle, hence, seasonal cycle-resolving analysis & evaluation can be useful in assessing key sources of the simulation errors. This will also help to link the seasonal cycle of the runoff to monthly snow ablation.

Author's answer: We agree with you and will integrate an analysis of monthly ablation in the revised version. This will also complete the monthly analysis we already performed for precipitation and runoff. Please also consider our answers to your specific

C3

comments on this point.

(4) The terminology "topography parameterization" is confusing. The true meaning of "topography parameterization" in the manuscript is "characteristics of topographic parameters" such as slope and azimuth that vary according to the DEM resolutions. "Parameterization" typically means representing unresolved properties using resolved values or representing a property using a related variable(s).

Author's answer: You are right, this was misleading. We will change the wording in the text where necessary.

(5) The writing is OK, but I found occasional awkward/unfamiliar sentences. As a speaker of American English as the second language, I don't like to suggest any specific changes in grammar and writing. I strongly recommend the authors to consult native English speakers to go through the entire writing.

Author's answer: Before submitting the reviewed version of the manuscript, we will give it to a native speaker for prove reading.

Specific comments:

(1) The authors present results from elaborate model runs and analyses, but the key messages are not clearly presented and are sometimes confusing.

Author's answer: We will revise the entire paper and streamline the text so that the main goals of our study and the key messages become clearer. Therefore, in the introduction, we will add our two main research questions as bullet points, which are both regarding the overall question how far it is possible to use globally available input data for snow-hydrologic modelling in ungauged high alpine catchments: "To which extent can we use different global meteorological data products varying in product type and spatial scale to simulate snow depth and further snow hydrological parameters as well as runoff in a high alpine catchment?" "What is the influence of different characteristics of topographic parameters like slope, aspect and altitude due to different

C4

DEM products with different resolutions on snow hydrological simulations in a heterogeneous high alpine catchment? We will provide the answers to these questions more clearly in the relevant discussion and conclusion parts as well in the abstract, so that the key message of our results is better visible. Since we already performed additional model runs with the quite new global products ERA5 and ERA5-Land data as suggested by reviewer 1, our updated key messages are regarding: a. the investigated global meteorological data sets: In total, we investigated 12 different meteorological setups considering a variety of different data products (ERA, GLDAS, CFSR, CHIRPS) covering different scales and versions of the products and an additional specific down-scaling method for one product. However, only five meteorological setups including the reference in situ data setup showed agreeable results regarding the simulations of snow depth and other snow hydrological parameters, like total snow cover and ablation period, and the date and quantity of maximum SWE. However, in the results of these five setups considerable differences in the runoff regimes were found with a delay of peak runoff of up to one month. The applied globally available CHIRPS data performed in total agreeably well and gave the best results regarding all other meteorological setups. However, this setup covers the specific case that only precipitation was substituted and other meteorological driver data are still taken from the measured in situ data. Our investigations on the assumption that the entire data set of meteorological forcing is required, the newly added ERA5 and ERA5-Land products, that have a comparably high spatial resolution regarding all other applied global meteorological products, produced the best results. This especially applies to the results of snow depth simulated at the two meteorological stations as well as the catchment mean regarding the further hydrologic relevant snow cover parameters. In total, ERA5 products performed even better than the data transferred from a similar catchment. The other tested globally available meteorological data showed very weak performance in simulating the snow depth at the two stations and lead to absolutely unrealistic snow cover developments and were thereafter excluded for the more detailed analysis of temporal and quantitative comparisons regarding the catchment mean and on HRU

C5

basis of the further snow hydrological parameters and the runoff. b. the investigated different globally available DEMs: In total, we used four different products with a wide span width of spatial resolutions from 2.5 m (LiDAR derived reference DEM setup) up to 1 km (GTOPO30). Two products we used have the same resolution with 30 m (SRTM and ALOS) to check the potential differences in results, although the spatial resolution is the same. Regarding the three globally available DEMs and the reference setup used for the parameterization of the surface characteristics, all DEM setups reproduced the measured snow depth and the further snow hydrological parameters on HRU and catchment scale quite well. However, they show considerable differences in the runoff regime, which is especially the case for the two DEMs of 30 m resolution in comparison to the reference. Despite the fact that the very coarse GTOPO30 DEM performed relatively well on the catchment mean, we advise against using this product in such heterogeneous high alpine terrain since the small-scale differences cannot be captured. The key messages of our results will be also better connected to our following overall suggestions regarding the applicability of such global products in ungauged basins in general. The formulation of the key messages will also be improved in the abstract (as also mentioned in the next comment).

(2) Please improve Abstract so that the key findings in the experiment are presented more concisely clearly. Separating the sensitivity to DEMs and associated orographic parameters from the sensitivity to the forcing data may help organizing with more clarity.

Author's answer: As mentioned in the answer before, we will define the main questions and key messages much clearer and will also restructure the abstract according to your suggestions.

(3) The statement in Abstract, L23:24, contradicts the statement in Conclusion L625-634.

Author's answer: We will rewrite the statement in the abstract to avoid this contradiction.

C6

(4) L149:151: How the mean snow field (e.g., SWE, SCA) over the entire RCZ is calculated? This is among the key evaluation variables.

Author's answer: CRHM calculates for each HRU and each time step values of SWE and snow depth. For these calculations, CRHM requires a meteorological input for each HRU. Since such data is not automatically available for each HRU, we used the method developed by Liston and Elder (2006) to generate it for each HRU. We use the modelled values of SWE and snow depth to calculate the area weighted mean value for the entire RCZ. We are confident, that the chosen HRU delineation is suitable for realistic snow cover simulations since it is the same as applied in Weber et al. (2020), where we established these HRUs and evaluated them with spatially distributed LIDAR snow depth measurements. We will describe this better in the revised manuscript and hope that this is the answer to your question since L149:151 do not deal with the calculation of the mean snow field over the entire RCZ.

(5) L156:157: The met data at LWD are not used as the forcing. Why the snow precipitation at LWD is corrected for undercatch?

Author's answer: The meteorological data at the LWD station has, as already stated in the manuscript, some longer data gaps (especially at the end of the winter seasons and during summer time in the study period of 2000-2010), which is why we refrained from using it as meteorological forcing data. Since 2014, a snow scale was installed at the LWD station, which enabled us to use these measured SWE for more recent years as a good possibility to investigate the precipitation under catch. Similar to the literature reported under catch of snow precipitation of up to 50% (WMO, 2011; Grossi et al., 2017), we also found out in our former study (Weber et al., 2020) that we can expect an under catch of 50% for the RCZ. Assuming that under catch variations in precipitation over the entire catchment are rather negligible, the DWD snow precipitation has been corrected with the factor derived at the LWD station. We will clarify this in the revised manuscript.

## C7

(6) Spell out HRU at its first appearance.

Author's answer: We will spell it out at its first appearance. We can also offer to provide a list of acronyms.

(7) L227: Check the spatial resolution of the CFSv2 data. The finest resolution I could find is T382 that correspond to approximately 0.313degree.

Author's answer: We checked the spatial resolution again and thank you for the hint! In course of this, we realized that we misunderstood the data description of the database where we downloaded the data. As you wrote, the spatial resolution of CFSv2 is approx. 0.313 degree. However, CFSv2 only exists from 2011 onwards. All data before and thus the data we used is data from the Climate Forecast System Reanalysis (CFSR) which has a 0.2 degree resolution. We will change this throughout the manuscript and rename this setup CFSR.

(8) L265: Section 3 → Section 4

Author's answer: We will change that.

(9) L288-291: I don't understand what "contrary development" indicates. This sentence is ambiguous. Please provide more explanations.

Author's answer: We referred to the wrong Figure and will correct that. It is Figure 3b instead of 3a. This figure shows that precipitation in the reference and at the DWD\_Wendelstein setup is contrary over the years 2005 – 2010.

(10) L293-295: This sentence is ambiguous. Please rewrite.

Author's answer: We will rewrite the sentence.

(11) L288:289: This sentence cannot explain the peak in March.

Author's answer: We guess you mean L.298-298 and thank you for pointing this out. This sentence makes indeed no sense and will be deleted.

## C8

(12) May change the title of Section 4.3. I suggest “Landsurface parameterization on basis of DEMs” → “Sensitivity of the land-surface parameters to DEMs (or DEM resolutions)”

Author’s answer: We will change the section sub title according to the reviewer’s suggestion.

(13) Also suggest a new title for Section 4.4: “Influence of the DEMs and associated land-surface parameters on meteorological conditions”

Author’s answer: We will change the section sub title according the reviewer’s suggestion.

(14) L339: 100 m → 200 m (195 m). (200m elevation difference also corresponds to a lapse rate of 5K/km, slightly more stable than the standard atmosphere which is understandable over a cold surface like snow/ice.)

Author’s answer: The difference in catchment mean altitude is 105 m which is roughly 100 m. We will write the exact value instead of roughly 100 m to avoid confusion and also add the catchment mean values in Table 2.

(15) Section 5: The evaluation based only on NSE and R2 is insufficient. Need more metrics, at least the ‘mean bias’ and RMSE. Also provide additional evaluations for each month (i.e., annual-cycle resolving model evaluations)

Author’s answer: We will additionally provide the mean average error (MAE) and, as previously suggested, we will also include an analysis on a monthly basis of the ablation and will discuss this also in the context of the already provided analysis on seasonal precipitation and runoff.

(16) Section 5.1: If there are problems in evaluating the snow simulations at DWD and DLW as stated in Section 5.2, how can you justify evaluating the daily snowdepth against the observations at these sites?

C9

Author’s answer: In section 5.2 we state that we do not validate the snow cover duration and the number of ablation days with measured values, mainly due to data gaps at the LWD station in spring, which make it impossible to determine the melt out day as well as the exact length of snow cover duration for some years. In addition, at the DWD station, the melt out day may be delayed due to the fact that the snow stake is installed directly on the small glacier ‘Nördlicher Schneeferner’. This might bias the definition of snow cover duration and ablation days. Nonetheless, we can very well quantitatively validate the daily modelled with the measurements snow depth for the entire season when data is available (this is true for most days except the days, where we had data gaps at the LWD station at the end of the season). Of course, there might be years regarding the DWD station, in which the model is not able to capture the exact melt out date of the snow cover due to the mentioned effect (ice is conserving the snow cover). However this can almost be neglected in a day based comparison since this snow cover is usually very thin and hardly affects the total seasonally accumulated snow volume. We will clarify the above mentioned points in more detail in the updated version.

(17) L395: “snow towers pile up” → Is this due to the forcing errors or model errors or combined forcing-model errors?

Author’s answer: This is a well-known effect, which might occur if the meteorological input is not realistic, which is the case for some input setups we showed. The model structure, however, is always the same in each of our setups and such snow towers could never be observed when forcing the model with in situ measured meteorological data. The reason for this snow towers is mentioned in the following sentences: “These simulated snow towers are mainly caused by lower mean annual temperatures and higher amounts of precipitation compared to the meteorological values of the reference data set (Section 4.2).” To clarify this in more detail we also added the following definition: “Snow towers are an effect in snow hydrological modelling that occurs mainly at higher altitudes and describes the unrealistically high accumulation of snow over

C10

several years. Reasons can be the insufficient description of redistribution processes in the model or unrealistic meteorological driver data (Freudiger et al., 2017).”

(18) L395: ‘downscaled temperature’ → ‘temperature downscaling’

Author’s answer: We guess the reviewer means L399. We will change it.

(19) L406:421: This paragraph repeats the statements in the previous paragraph. May be removed and present a summary of this paragraph in Conclusion.

Author’s answer: We will remove this paragraph and we will also revise the manuscript with an eye on deleting repetitions in general.

(20) Section 5.2

A. Statements in this section is too qualitative without solid supports from acceptable level of evaluations.

Author’s answer: We partly agree and partly disagree with you. On the one hand, as suggested, we see that it is necessary to also include an analysis of ablation on a monthly basis to better link the snow cover to the presented runoff simulation results. On the other hand, our statements in this section were already supported by “hard” facts and statistical measures as presented in the Tables 2, 4, and 6. Of course, it would be good to perform further statistical analysis with the presented indices. However, we have only ten years regarding values per index and model run and would need many more years to have enough data for a more in depth statistical analysis. In addition, we do not consider this as necessary since our objective is to demonstrate if it is possible to simulate the snow cover and runoff of our catchment with globally available data and subsequently to show the differences between the model runs.

B. If missing data and site characteristics at LWD and DWD, respectively, prevent using these data for evaluation, how can we trust the reference data are accurate enough for evaluating model data?

C11

Author’s answer: This concern is reasonable, probably for all measurement stations worldwide. DWD station: this data is quality-controlled by the DWD, the German Weather Service. Measurement errors are filtered and there are virtually no gaps. LWD station: The instruments of this station are regularly maintained by the LWD, the Bavarian Avalanche Warning service since they rely on good quality data as well. All, the large gaps at this station used to occur in late spring when the avalanche season was over and the LWD no longer produces an avalanche warning bulletin. So they simply had no need for data in this time of the year. So the data recorded at that location is the most accurate one can get. Regarding the location of the measurement stations, both institutions chose a location which is as representative as possible for the altitude. Further restrictions particular concerning the use of DWD data for snow cover duration evaluations are explained in the answer to comment 16.

C. If DMSWE cannot be validated, how can MSWE can be validated? MSWE is supposed to occur on DMSWE.

Author’s answer: We never validated the MSWE with measurements and never stated that in our manuscript, as we do not have measured SWE during the investigated time period. Regarding measurements, we only compared the measured snow depth at the LWD and DWD to validate our model runs. Regarding all SWE related snow hydrological parameters (MSWE, DMSWE, snow cover duration, ablation period), we compared the simulated results of all investigated setups to the results of the reference setup, which is forced with in situ measured driver data at the DWD station and which was parameterized with the high resolution DEM.

(21) L457:470: How can the data of largest warm bias, precipitation underestimation and insolation overestimation perform best in simulating the amount and timing of runoff? This may indicate major flaw in the model physics and/or forcing combinations. Need further discussions. This also indicates the need for a more rigorous evaluations of the reference data and other model data.

C12

Author's answer: This is a good point and will be discussed in more detail in the updated version. To answer this, we first try to explain a bit more in detail what actually happens during discharge concentration in our model setups and what are the local conditions. In reality, there is no surface runoff in RCZ due to its karstic situation. However, CRHM has no routine to route karst runoff. Moreover, there is not enough knowledge of the hydrological response of the karst system until now. Therefore it is almost impossible to appropriately parameterize the routing in CRHM, which is why the runoff presented in our study can be regarded as direct runoff. This means that the surface structure due to the DEM has almost no influence on runoff routing. So the differences are in the processes of discharge formation. Despite the GTOPO30 DEM is much lower in spatial resolution than the reference DEM, it is still high enough or has the effect that the monthly mean temperature is below freezing until March and around 0°C (0.1°C) for April while it is -0.7°C in April in the reference. However, both, the reference and GTOPO30 setup, are clearly above freezing level in May at 4.2 °C and 5.1 °C, respectively. This means that the main melt, which is also temperature induced in early spring, starts almost at the same time in both setups. As Figure 7 shows, the melt induced runoff starts slightly earlier in the reference setup, because overall higher HRUs in the catchment, receive overall considerably more Qsi in the reference than it is the case in the GTOPO30 setup. The Qsi values of GTOPO30 are due to the coarse resolution rather the same over the catchment. Nonetheless, the major part of the RCZ receives in total more Qsi in the GTOPO30 setup. The longer the days last in later spring, the stronger the radiation induced melt effect becomes, which is why in the GTOPO30 setup, melt induced runoff becomes stronger in June than in the reference setup. Moreover, in summer the lower elevation has a greater effect on the temperature than in winter which leads to higher melt induced runoff in the GTOPO30 setup in June and July. This effect decreases in August since less snow is left to melt. We will add further discussions on this point to the text.

(22) L475: Monthly Qsi needs to be evaluated as it is directly involved in snowmelt.

C13

Author's answer: We agree with you and will take a closer look at the monthly radiation and also at the radiation distribution. As mentioned in the answer to comment 21 there are HRUs, which receive significantly overall more Qsi in the reference than in the GTOPO30 setup, and therefore, show an earlier melt in the reference setup. However, there are overall also larger parts (upper, medium and lower parts) in the GTOPO30 setup, which receive in total more radiation than in the reference setup. This effect gains power with longer days and leads to a pronounced melt in June in the GTOPO30 setup.

(23) L506:507: This may be an overstatement for the CHRPS data. CHRPS yields good results for snow cover and runoff, but not in NSE and R<sup>2</sup> of the daily snow depth.

Author's answer: The NSE and R<sup>2</sup> of the CHIRPS setup for daily snow depth is clearly the best among all presented globally available different meteorological setups. Despite the quality measures are lower than for the reference, they are still considerably high.

(24) 511-530: Can the inter-model difference also be related to their snow models? Do all of these models use the same snow model?

Author's answer: The inter-model difference can also be related to the used snow models. As explained in the paragraph, the studies cannot be compared one by one since they are conducted on different scales, in different (micro-) climatic regions and also with different models. However, these studies show that it is generally possible to obtain reasonable results with global data, if conditions are appropriate, e.g. data is available for downscaling or the investigation of larger scales.

(25) Please discuss the poor NSEs and R<sup>2</sup>s with the transferred and CHRP data.

Author's answer: As stated in the answer to comment 23, we consider the quality measures of the CHIRPS setup to be relatively good. According to Moriasi et al. (2007) and NSE > 0.5 can be considered as satisfactory in hydrological modelling. This threshold

C14

is also used in alpine hydrological modelling (e.g. Rahman et al., 2013). The reason for the good performance of the CHIRPS setup is that the forcing data is the same as in the reference setup except for precipitation. As illustrated in Figure 4, CHIRPS precipitation well reflects the precipitation regime in RCZ. However, it is biased to higher values. Regarding the transferred data from Mt. Wendelstein, in particular the NSE values are worse. This can be traced back to the years 2002, 2005 and 2010 in which the peaks are strongly overestimated regarding the DWD\_Wendelstein run. The NSE is particularly sensitive to large values. We will add a discussion on this.

(26) L550-551: This statement ignores the substantial differences in runoff between GTOPO30 and ALOS/SRTM.

Author's answer: We agree with you and will rewrite the sentence to account for the differences in runoff.

(27) L571: "the choice of the DEM has far less impact" Incorrect. DEMs have large impacts on the simulated streamflow annual cycle

Author's answer: We will clarify this sentence, because in comparison to the setups varying in meteorological driver data, the impact is less.

(28) L600:604: Misleading. The 'borrowed' data performed poorly in terms of NSE and R2.

Author's answer: We will discuss this to be clear that the 'borrowed data' from another station in a similar catchment performs worse than the CHIRPS setup.

(29) L510-512: There is a mystery. How can the forcing data sets of such a wide variation can produce such similar simulations? This needs answers from the authors.

Author's answer: We are not sure if you really mean L510-512 since we did not find a statement there that fits to your comment. We assume you perhaps could have meant L610-612, which is about differences due to DEMs. The answer why the different DEM setups performed quite similar is that on the one hand, the ALOS and SRTM DEM have

C15

the same spatial resolution, which is high enough that the HRU surface characteristics are equally well represented as with the reference DEM. On the other hand, if results at the catchment scales are considered, in the coarser GTOPO30 setup, differences are leveled out. Moreover, as written in the answer to comment 21, the effect of altitude and enhanced radiation input is not as strong in winter, as in summer. Nonetheless, regarding the individual HRUs there are considerable differences as Figure 8 illustrates.

#### References

Freudiger, D., Kohn, I., Seibert, J., Stahl, K., and Weiler, M.: Snow redistribution for the hydrological modeling of alpine catchments, *WIREs Water*, 4, e1232, doi:10.1002/wat2.1232, 2017.

Grossi, G., Lendvai, A., Peretti, G., and Ranzi, R.: Snow Precipitation Measured by Gauges: Systematic Error Estimation and Data Series Correction in the Central Italian Alps, *Water*, 9, 461, doi:10.3390/w9070461, 2017.

Liston, G. E. and Elder, K.: A Meteorological Distribution System for High-Resolution Terrestrial Modeling (MicroMet), *J. Hydrometeorol.*, 7, 217–234, doi:10.1175/JHM486.1, 2006.

Moriassi, D. N., Arnold, J. G., van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations, *Transactions of the ASABE*, 50, 885–900, doi:10.13031/2013.23153, 2007.

Rahman, K., Maringanti, C., Beniston, M., Widmer, F., Abbaspour, K., and Lehmann, A.: Streamflow Modeling in a Highly Managed Mountainous Glacier Watershed Using SWAT: The Upper Rhone River Watershed Case in Switzerland, *Water Resour Manage*, 27, 323–339, doi:10.1007/s11269-012-0188-9, 2013.

Weber, M., Feigl, M., Schulz, K., and Bernhardt, M.: On the Ability of LIDAR Snow Depth Measurements to Determine or Evaluate the HRU Discretization in a Land Sur-

C16



face Model, Hydrology, 7, 20, doi:10.3390/hydrology7020020, 2020.

WMO: Technical regulations: Basic documents no. 2, Volume I – General Meteorological Standards and Recommended Practices, 2010th ed., World Meteorological Organization, Geneva, 2011.

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2020-326>, 2020.