Enhance existing Swiss precipitation products with particular regards to snowfall

1 August 2021 - 31 January 2023

Final report

Michael Schirmer[1,2], Tobias Jonas[2]

[1]Swiss Federal Institute for Forest, Snow and Landscape Research WSL

[2]WSL Institute for Snow and Avalanche Research SLF

with the support of the Federal Office of Meteorology and Climatology MeteoSwiss, in the framework of GCOS Switzerland

Snow monitoring station in the Swiss Alps (© SLF)

1 Summary

It is well known that precipitation grids exhibit biases during snowfall events due to either an undercatch of precipitation gauges or the interpolation when there are too few gauges, especially at high elevations. However, in Switzerland daily snow depth observations are available from a network of 450 monitoring sites with a focus on high-altitude regions. This data set was used to correct the solid precipitation phase from the widely used Swiss daily precipitation grid product RhiresD.

Solid precipitation was determined from RhiresD using a partitioning function based on air temperature at an hourly time step. Hourly COSMO reanalysis data were used to disaggregate daily fields of precipitation and air temperature observations, limiting the analysis period to the last seven years from September 2015 to August 2022. Alternatively, solid precipitation was also estimated with a parametric snow model based on snow depth data from the monitoring sites. This model has been trained on over 10000 snow pit measurements and shows no bias compared to 20 years of bi-weekly observations of snow water equivalent from 45 locations at different altitudes in Switzerland.

These two datasets were combined using a data assimilation method called Optimal Interpolation (OI). OI seeks the best (i.e. minimum error variance for the analysis), bias-free guess of the unknown field (here solid precipitation) by updating prior information on the grid (RhiresD) with observations (monitoring data). These updates result in day-by-day corrections of the grid, but are only available for the analysis period, i.e. the last seven years of the dataset. To obtain long-term corrections for climatological applications, bias statistics were developed that vary monthly. A global correction of 20 % more solid precipitation was used as a benchmark. The monthly and spatially varying corrections fluctuated around this value. A clear trend in elevation could not be identified due to regional differences in the measurement networks for snow and precipitation observations.

The correction products were evaluated at the point scale using cross-validation. This assessment showed an increase in performance from none to global to monthly to daily correction. After daily correction, the lowest bias category contained about 15% more stations compared to uncorrected data and about 6% compared to the benchmark. The mean squared error of snow depth, modelled with an energy balance snow model fed with these corrected and uncorrected input variants for solid precipitation, was reduced by about 25% in the highest altitude category (around 2500 m) compared to the uncorrected input or the benchmark. Below 1500 m, however, where most snowfall occurs close to 0°C, the error reduction was negligible.

Two methods were chosen for a spatial evaluation, firstly a water balance analysis and secondly a validation with lumped hydrological models in about 60 relatively undisturbed catchments in Switzerland. In such spatial applications, it must be considered that flat field observations overestimate snow depth compared to complex surrounding terrain. To compensate for this effect, an additional correction for steep terrain based on LiDAR data used in previous publications was developed. The two errors, gauge underestimation and representativeness error, happened to be of similar magnitude but with opposite sign. This complicates any spatial evaluation, as the combined correction, especially when aggregated at the catchment scale, was comparatively small. Accordingly, the methods were not accurate enough to verify/falsify any benefit of using the corrected precipitation fields in the context of hydrological modelling.

2 Scientific report

2.1 Introduction

2.1.1 Motivation / Background

Snow is an essential component of the water cycle in Switzerland. More than 30% of the total precipitation is snowfall, and about 40% of runoff in Swiss rivers originates from snowmelt. Seasonal snow constitutes a temporary storage of precipitation, and its condensed release during a relative short period of time entails seasonally variable runoff. But the presence of snow cover also has other important implications. It drastically increases the land surface albedo with direct feedbacks on radiative fluxes and surface temperatures, which in turn affect, amongst other meteorological variables, air temperatures and turbulent heat fluxes. These variables do not only affect snowmelt but also set boundary conditions relevant to plant growth and other ecological processes. Further, the presence of snow directly affects wildlife and plant functioning, e.g. by hampering the mobility of animals or determining the light availability for low vegetation. For the above reasons, snow is considered an essential climate variable (ECV).

Measuring snowfall amounts is difficult. Precipitation gauges have a known undercatch for solid precipitation (e.g. Goodison et al. 1998; Rasmussen et al., 2012; SPICE project). At an alpine site in Switzerland (Weissfluhjoch, Davos) Smith et al. (2020) found an undercatch of 45% in comparison to a double fenced automatic reference gauge. Many studies have also shown that the relationships between wind and undercatch have a large scatter. Transfer functions have been developed as part of the SPICE project (e.g. Kochendorfer et al., 2021). However, the magnitude of these unresolved errors decreases significantly when more effective shielding is used (Kochendorfer et al., 2021). This is especially a problem in Switzerland, where most precipitation gauges are not shielded. Other known problems arise from the accumulation of snow in and on the gauge, which either obscures the gauges and/or later releases snow in the gauge.

Switzerland has a worldwide unique station density measuring snow depth. Approximately 100 Intercantonal Measurement and Information System (IMIS) snow stations were built in the late 90ies in flat and sheltered terrain to estimate new snow fall amounts. Snow depth from these and other stations (SwissMetNet, the automatic measurement network of MeteoSwiss) and several manual observers with at least one reading a day provide valuable information for two operational services run by the SLF in Davos, namely the Avalanche Warning Service and the Operational Snowhydrological Service. IMIS stations provide an additional benefit to the existing station network of MeteoSwiss since they are typically not mounted in the valleys, but in elevations between 1200 m and 3000 m (mean: 2200 m). Integrating snow information from IMIS stations, SwissMetNet, and from manual observers into widely used Swiss precipitation products is therefore aimed directly at Pillar 2, Priority 2.1 of of the GCOS Call for Proposals 10/2020.

Snow distribution is highly variable in space and time, and variability across different spatial scales is due to different processes. While larger-scale weather patterns drive regional snow distribution patterns that vary dramatically between years, it is micro topography and vegetation that shape small-scale snow distribution patterns, which are remarkably persistent between years. Obviously, snow is also stratified with elevation, slope, and aspect at intermediate spatial scales due to orographic effects on precipitation and air temperature including precipitation phase, preferential deposition of snowfall, and the redistribution of snow on the ground. These aspects make studying snow distribution dynamics both an interesting and challenging undertaking.

While flat and sheltered conditions at IMIS stations may provide for good snow measurements, they unfortunately do not provide representative values. In fact, true spatial mean values are typically overestimated if assessed based on data from flat measuring sites (Grünewald et al., 2015). This is because on average snow is shallower in sloped terrain, within forests, and in exposed terrain. This project will account for this effect building on published work.

2.1.2 Brief overview of work done

Biases in snowfall estimates for widely used precipitation grids were determined using snow observations, snow modelling and data assimilation routines, which resulted in the development of pixel-wise and both daily and monthly varying correction schemes. The established correction scheme for RhiresD was evaluated at the point scale using cross-validation at 450 stations. At the spatial level, it was examined whether an assessment of the water balance or the performance of the hydrological model would benefit from corrected data.

2.2 Methods

2.2.1 Data

Daily snow depth measurements (Figure 1) from seven years were used, i.e. from September 2015 to August 2022. Stations and manual observations from different sources were integrated, namely IMIS snow stations, SwissMetNet, manual observers, and weather stations from Austria, Germany, and France, resulting in a total of 437 locations. The IMIS stations are in high mountain regions, which provides value for this project (Figure 1b, c).

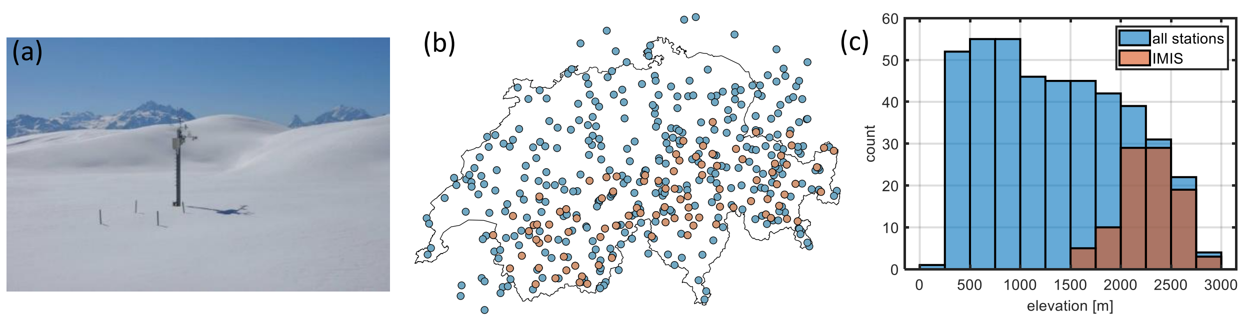

Figure 1. IMIS snow station (a), spatial distribution of daily snow depth observations (b) and elevation distribution (c).

Daily air temperature measurements from 256 weather stations were used for the same period to split the total precipitation into its solid and liquid phases. Hourly COSMO reanalysis data was used to time-disaggregated daily air temperature and precipitation data. For the independent point-by-point evaluation, snow water equivalent (SWE) observations were used, which were measured manually every two weeks at 45 locations in Switzerland. For a spatial evaluation, runoff measurements were used at around 70 relatively undisturbed catchments in Switzerland in the period from 1989 to 2018.

Two precipitation products were used to calculate the correction factors: Firstly, the gridded deterministic daily precipitation network RhiresD v2.0 (MeteoSwiss, 2021) and the probabilistic gridded product RhydchprobD (MeteoSwiss, 2019; Frei and Isotta, 2019). The two precipitation grids to be corrected are referred to in the following only as "RhiresD" for better readability.

2.2.2 Pseudo-observations of solid precipitation

A parametric model (HS2SWE, Magnusson et al., 2014) was used to calculate the daily solid precipitation (HNW). This model was calibrated using the snow depth measurements described in Section 2.2.1 and derives Snow Water Equivalent (SWE) using observed snow depth (HS) as it accumulates, compacts, and melts layer by layer (Magnusson et al., 2014).

2.2.3 Determining the precipitation phase

The daily RhiresD precipitation grid was divided into a solid and a liquid precipitation phase at an hourly time step. The required temperature grids were determined in a similar way as the TabsD product (Frei, 2014). Both the daily precipitation and air temperature grids were time-disaggregated using hourly COSMO data. The hourly COSMO air temperature data were modified maintaining the same daily mean as the daily grid for each individual grid cell. Since the spatial precipitation patterns determined by station interpolation or by a numerical weather prediction model can be very different, an alternative to the pixel-by-pixel approach was chosen. The hourly precipitation grids were gradually blurred with a low-pass filter to ensure that each pixel with precipitation in the daily grid was matched by a pixel with precipitation in the hourly grid. Air temperature and precipitation were then interpolated to station locations. Finally, the precipitation partitioning function derives the phase for each pixel and station location using a logistic smoothing function whose parameters, originally determined by Magnusson et al. (2014), were adapted to the hourly time step used here.

2.2.4 Data assimilation of point pseudo-observations

Optimal interpolation (OI) was applied to update daily solid precipitation (HNW) using the pseudo-observations described above. Figure 1 illustrates this procedure. RhiresD solid precipitation showed considerable differences to pseudo-observations at the stations, which was minimised after correction through Optimal Interpolation, e.g. in the Valais or Plateau. Since OI works best with an unbiased background field, Magnusson et al. (2014) determined a global undercatch correction of 20 % for the RhiresD solid precipitation, so that it best matched the pseudo-observations at all stations. This constant correction factor of 20 % was also used as a benchmark for correction factors developed later, which is more spatially and temporally specific. A mathematical description of the OI and further details can be found in Magnusson et al. (2004). OI was applied to grid and to station locations. As HS measurements and pseudo-observations are subject to errors, the output of OI at the station locations were used to calculate correction factors (next Section). They can be seen as the best compromise between two types of information that are both subject to errors.

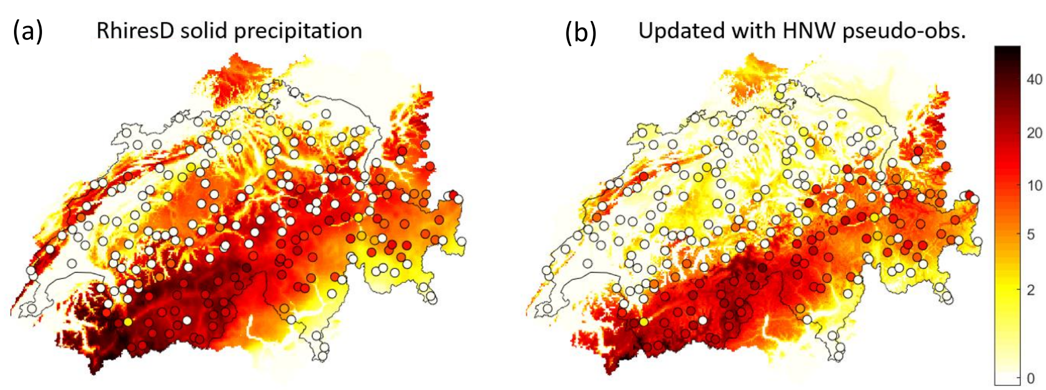

Figure 2. Modelled solid precipitation based on COSMO data, before (a) and after (b) assimilation of snow station data. Circles represent snow observations, the grid represent the interpolated values after Optimal Interpolation (OI).

2.2.5 Correction factors

The daily updated solid precipitation provides the day-by-day correction of RhiresD. This is available for seven years and could be used for now-casting applications in the future. To determine a correction factor suitable for climatological applications, monthly varying correction factors were simply calculated as the temporal sum of updated solid precipitation divided by the sum of RhiresD solid precipitation determined at each pixel and station location. Only pixels/stations receive a correction factor if there were at least five days with more than five millimetres of snow in the data set. This results in data gaps particularly at lower elevation and in the summer months, which were filled with ordinary kriging. RhydchprobD was treated differently from RhiresD: as this product does not take orographic precipitation into account (Frei and Isotta, 2019), the long-term climatology of RhiresD was applied to eight ensemble members of RhydchprobD. For this purpose, the long-term climatology of RhydchprobD was calculated and daily anomalies were calculated, which were finally multiplied by the long-term climatology of RhiresD.

2.2.6 Point evaluation

At the station scale, different variants of solid precipitation were tested: (i) uncorrected RhiresD solid precipitation (see Section 2.2.3), (ii) RhiresD only corrected with a constant undercatch of 20 % (see Section 2.2.4), (iii) monthly varying correction factor (individual for each station, see Section 2.2.5) and (iv) a daily correction (individual for each station, see Section 2.2.5). These solid precipitation variants were compared with the pseudo-observations described in Section 2.2.2). As this is an error-prone and modelled variable, continuously measured snow depth (HS) and bi-weekly measured SWE were also used as a reference. For this purpose, the different variants of solid precipitation were fed into an energy balance snow model FSM (Essery et al., 2015), whose output was analysed by HS and SWE. Cross-validation was used to assess the ability of OI to interpolate to new sites. Cross-validation was also used to evaluate a-priori settings of OI.

2.2.7 Spatial evaluation

For the spatial evaluation it is important to consider that point measurements of solid precipitation on flat fields overestimate the surrounding mean value in complex terrain (see Introduction). To account for this overestimation, a precipitation multiplier was used that is a robust function of slope at a sub-grid scale of 25 m. Slope is largely correlated with small-scale roughness (Lehning et al., 2010). This function was determined based on previous publications (Grünewald et al., 2015). As these data do not cover forested terrain, no correction was made for forested parts of Switzerland (similarly based on a sub-grid scale of 25 m). Daily solid precipitation derived from RhiresD was multiplied by a monthly and pixel-wise varying correction factor and a pixel-wise varying precipitation multiplier. Only the solid phase was altered, while the liquid phase was added to receive total precipitation.

Since there are no widespread LiDAR data available for observing snow depth in Switzerland, runoff measurements in rather undisturbed catchments in Switzerland were chosen to test different total precipitation variants. The selection of catchments is based on previous studies (Griessinger et al., 2019; Brunner et al., 2019). First, the water balance was calculated using runoff measurements and modelled real evapotranspiration with PREVAH (Viviroli et al., 2009). Different variants of total precipitation amounts were tested, e.g. with/without the precipitation multiplier/undercatch correction factor. Note that it is methodologically incorrect to use only the undercatch correction at the spatial scale. We included this to test if our evaluation methods could falsify this obviously wrong input variant. Second, the ability of four lumped conceptual hydrological models (HBV in the TUWmodel R package, Viglione and Parajka, 2020; GR4J to GR4J with CemaNeige snow module in the airGR R package, Coron et al., 2020) to transfer calibrated parameters in time with uncorrected and corrected precipitation input was tested. TabsD (Frei, 2014) was used to produce spatial averages of air temperature. The models were calibrated with the first half of a 30-year data set and evaluated with the second half, and vice versa. For calibration and evaluation, the Kling-Gupta-Efficiency (KGE) criterion (Gupta et al., 2009) was used.

2.3 Results

2.3.1 Pseudo-observations of solid precipitation

A long-term comparison between simulated Snow Water Equivalent (SWE) values and bi-weekly manual observations shows no bias.

2.3.2 Correction factors

Figure 3a shows a static correction factor, i.e. calculated for the whole data set, for RhiresD. The values ranged from 0.4 (overestimation of RhiresD solid precipitation) to 1.7 (underestimation). Note that the values in the Plateau were particularly prone to uncertainties because the dataset was small (not many snowfalls greater than five millimetres in seven years). In addition, snow modelling around the melting point is challenging due to mixed precipitation phase or simultaneous accumulation and melt processes.

The modelling uncertainty at higher elevations is much lower, as the number of days in the dataset increases and the relative number of warm days with snowfall decreases. These areas mainly showed corrections around 1.2 (Figure 3b, red dots) with a slight elevation gradient, but also a large scatter due to regional differences. These regional differences were most obvious in the Glarus and the Vorderrhein, where the corrections were larger, or conversely in the upper Rhone region (Goms), where the corrections were smaller. These regions were already identified by Magnusson et al. (2014) using similar methods.

The generally low correction factor of 1.2 over large areas in the Swiss Alps is remarkable. The undercatch of an unshielded precipitation gauge in the Swiss Alps at the Weissfluhjoch was quantified with 45% (Smith et al., 2020). However, these values were quantified for snow only, while the here presented correction factors also included mixed precipitation. Calculating correction factors based only on data if the air temperature was below -2°C (using the same threshold as in Smith et al. (2000)) resulted in much larger values (Figure 3c) and similar to those reported from SPICE. For the same reasoning, colder winter months have higher correction factors compared to warmer spring months (Figure 3d). Another reason for the conservative estimate could be found in the OI routine: When the correction factors were calculated directly with pseudo-observations instead of using the OI output (Figure 3b, blue dots), they are less conservative as the proposed correction factors using OI (red dots). Furthermore, interpolation and extrapolation routines of RhiresD define how precipitation is determined in high elevation areas with a limited number of gauges. These seasonally and regionally varying routines may overestimate snow fall in higher elevations. Both reasons can also explain observed regional differences.

We conclude that the correction factors presented here were conservative compared to single site estimates. Reasons for this can be found in the snow modelling and the RhiresD data processing chain.

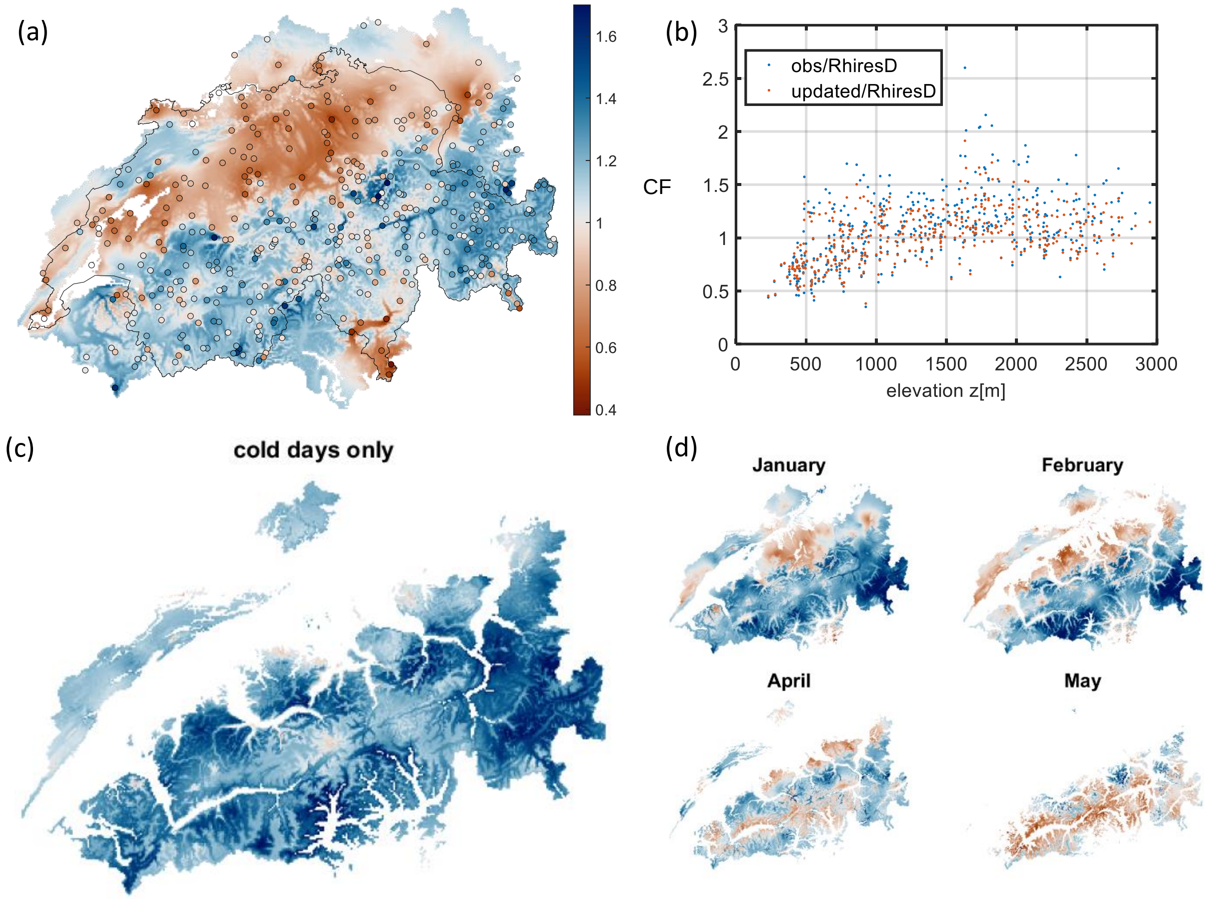

Figure 3. Static correction factor for RhiresD at station locations (circles) and on a 1 km grid (a), and elevation dependency of station locations of CF (red) or pseudo-observations divided by RhiresD solid precipitation (blue) (b), static correction factor excluding warm days (>-2 °C) (c), and seasonal dependency of the correction factor (d).

Correction factors were also calculated for eight ensemble members of RhydchprobD (not shown). Spatial patterns are similar between ensemble members (which should be the case since ensemble members are not temporally dependent on each other) and to RhiresD.

2.3.3 Point evaluation

2.3.3.1 Estimation of OI a-priori settings

The values presented so far were created with settings for OI already used in Magnusson et al. (2014). The expected uncertainty of the presented correction factor grid would favour a smoother interpolation than the one presented in Figure 3a. Therefore, the horizontal and vertical correlation length scales of the background field (i.e. RhiresD solid precipitation debiased with 20 %) were altered stepwise from 300 km to 30 km in the horizontal direction and from 2000 m to 250 m in the vertical direction to obtain potentially smoother fields. These parameter settings were cross-validated at the approximately 450 stations. The default settings of 30 km and 500 m received the best values, so they were retained for further analysis. As the aim of this section to produce smoother correction factor fields could not be achieved by choosing different but similarly good correlation length scale parameters, the use of low-pass filters was tested to remove unwanted small-scale variance (not shown).

Cross-validated (CV) estimates of RMSE and bias were analysed for four different variants of solid precipitation (RHIRESD), a benchmark method which globally corrects with a factor of 1.2 (RHIRESD 1.2), a spatially (pixel/station-based) and monthly varying correction factor (CF) and a spatially varying daily correction (OI) After daily correction, the lowest bias category contained about 15% more stations compared to uncorrected data and about 6% compared to the benchmark (not shown).

The same variants of solid precipitation were fed into the energy balance snow model FSM to compare results with observed snow depth, which is not a pseudo-observation like solid precipitation. Figure 4 shows the cross-validated RMSE and bias in different elevation classes. In the highest class, an expected increase in quality can be seen in both the RMSE and the bias. A reduction of the snow depth RMSE by approx. 30 % and 25 % compared to the uncorrected input and the benchmark, respectively, can be observed. Also in the second highest class, the values for the input variants CF and OI were best. In the lower classes, the differences between the precipitation variants were small. A negative bias, which would indicate that the correction factors applied are too conservative, could not be detected.

In summary, the evaluation of the point correction factors indicates that correcting RhiresD solid precipitation with the proposed correction factors provided an overall advantage, particularly at high elevations. However, this advantage was relatively small compared to a very simple benchmark where solid precipitation was corrected at a constant 20%, and furthermore, this advantage was not present for all quality metrics examined.

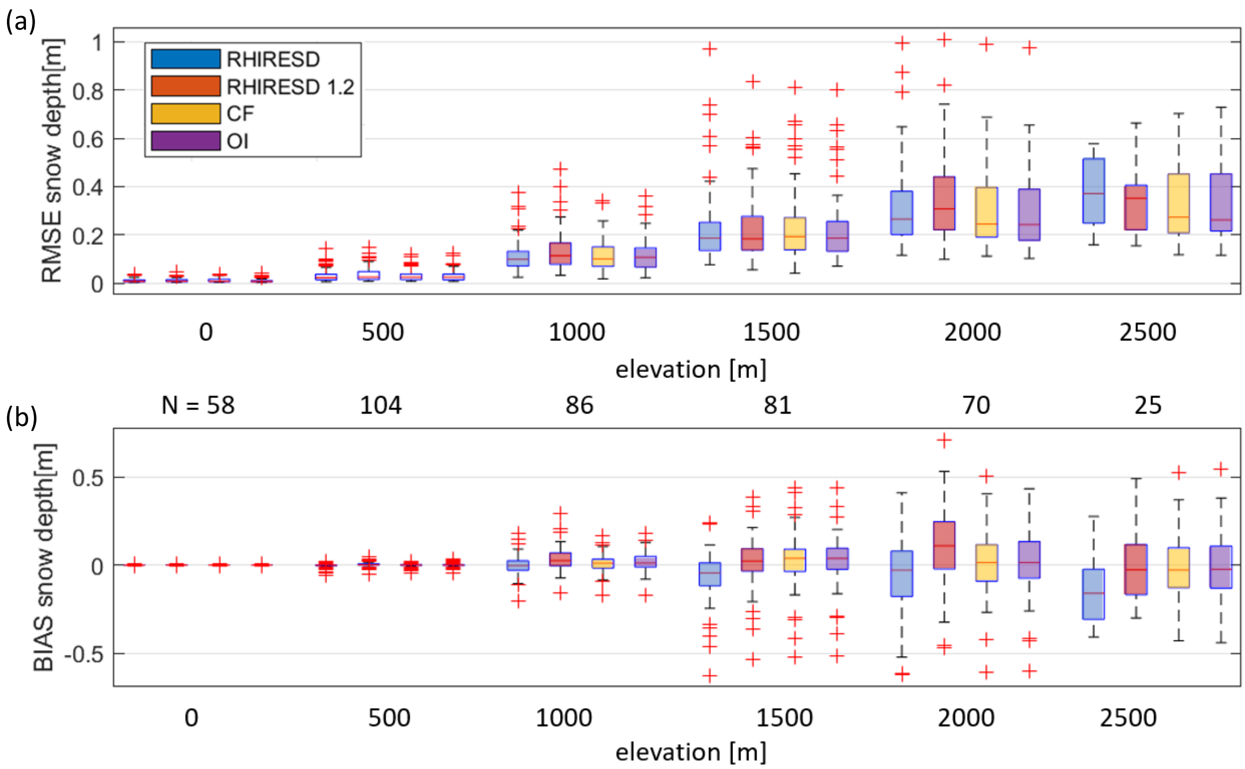

Figure 4. Cross-validated RMSE (a) and bias of modelled snow depth taking different solid precipitation versions as input for energy balance snow modelling.

2.3.4 Spatial evaluation

For the spatial evaluation, the overestimation of snowfall that occurs in typical flat field observations compared to the surrounding complex terrain was considered with the developed precipitation multiplier (prec_multi, Figure 5a). Figure 5b shows the effective correction on total precipitation, i.e. the combined effect of undercatch and representativeness error, averaged on selected catchments. The results show that the two errors almost balance each other out. The representativeness error was often larger than our undercatch error estimates, which resulted in a downward correction for most catchments, especially in central Switzerland.

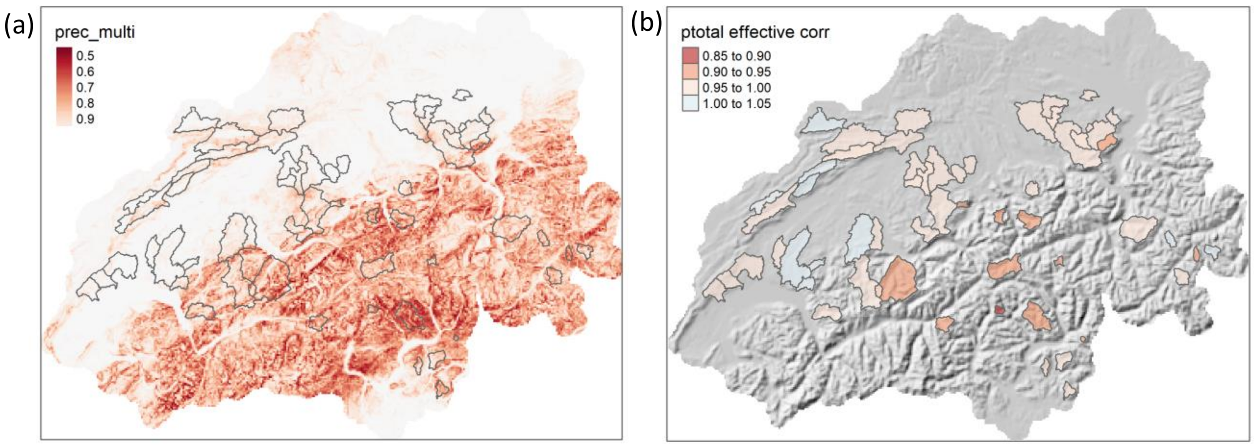

Figure 5. Precipitation multiplier (prec_multi) correcting for overestimation of flat field observations relatively to surrounding complex terrain (a) and the combined effect on total precipitation averaged per catchment (ptotal effective corr) (b). The selected catchments are represented by grey polygons.

2.3.4.1 Water balance assessment

Mean annual precipitation of three different variants was plotted against the observed runoff and modelled real evapotranspiration in selected catchments (Figure 6). No improvement but rather a deterioration of the corrected versions compared to the uncorrected RhiresD variant could be observed, e.g. due to a slope parameter closer to one or a better R^2. However, large uncertainties in both the observed runoff and the modelled evapotranspiration make this type of evaluation problematic.

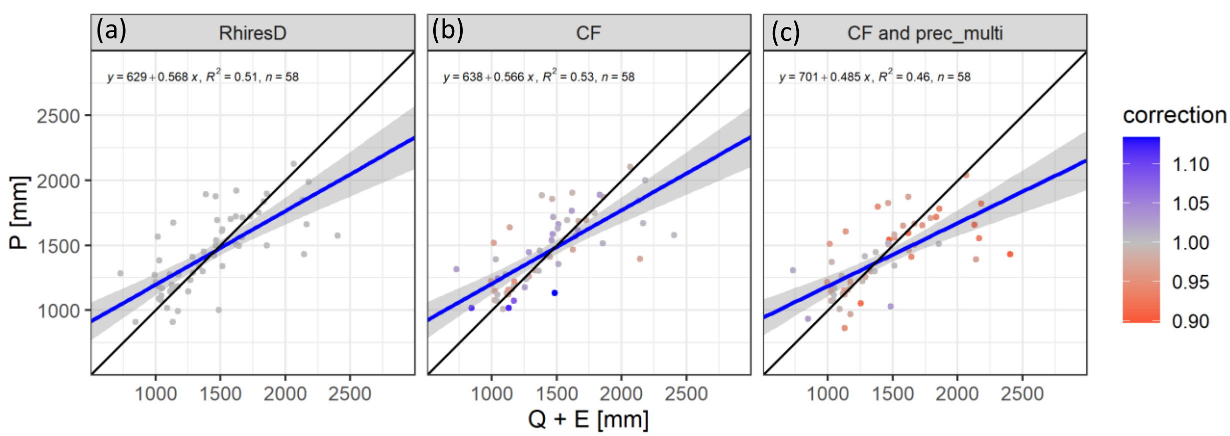

Figure 6. Mean annual values of precipitation variants plotted against observed runoff (Q) plus modelled real evapotranspiration Q, RhiresD uncorrected (a), correction of undercatch (CF) (b), correction of undercatch (CF) and overestimation of flat fields (prec_multi) (c). The colours indicate the correction for each precipitation variant. An equation of a linear regression fitted to the points is shown.

2.3.4.2 Hydrological modelling

The Kling-Gupta efficiency (KGE) of the validation period of four different lumped conceptual hydrological models is shown in Figure 7. Especially in spring, some catchments for the GR6J model seem to show an improvement over the uncorrected version of RhiresD. However, most values were within an approximate sampling uncertainty of the KGE or are less good. A clear advantage of the corrected variants could not be observed in this evaluation. Note, that these simple hydrological models seem to be able to produce quite good KGE values for substantially different precipitation variants outside a calibration period for most catchments, i.e. despite their relatively small number of parameters (6, 7, 8, 15), they were able to compensate for input errors.

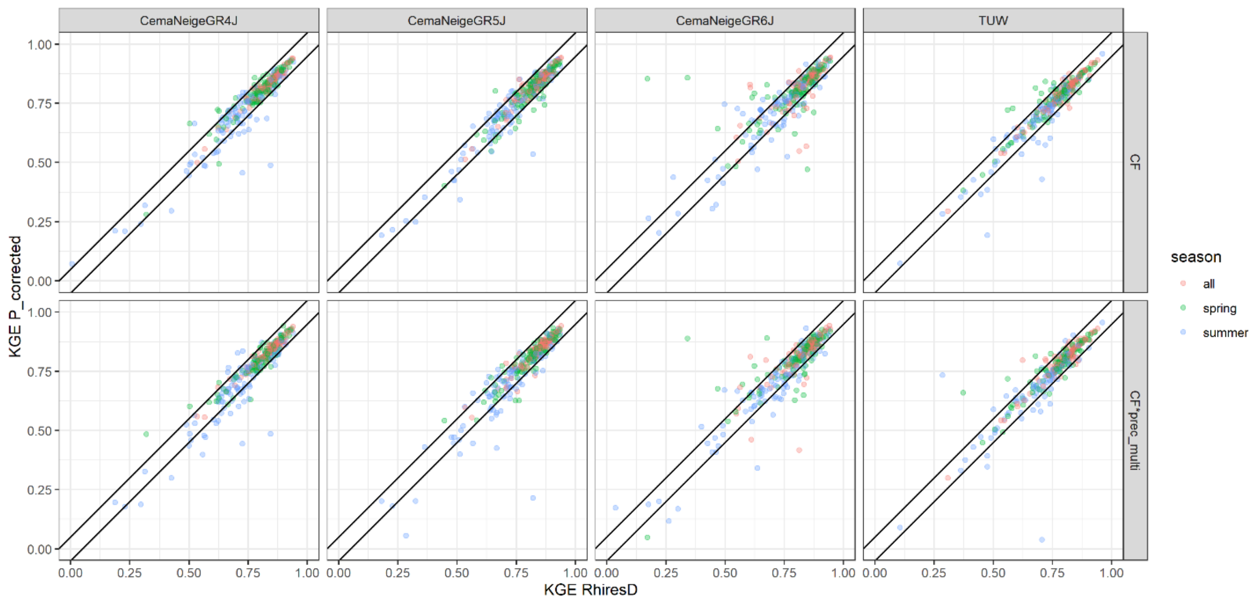

Figure 7. The Kling-Gupta-Efficiency (KGE) criterion of four different lumped conceptual hydrological models using two different corrections of RhiresD as input vs. uncorrected RhiresD. Plotted are values for different seasons. All values are from a 15-year long validation period outside of a 15-year long calibration period. An approximate sample uncertainty range of +-0.05 is indicated with the two dark lines (Clark et al., 2021).

Both evaluation methods, i.e. water balance analysis and the evaluation with lumped hydrological models, could not distinguish between corrected, uncorrected or incorrectly corrected precipitation. And this despite the fact that the incorrectly corrected precipitation differed greatly from the other variants. For the hydrological modelling part, the differences shown between the input sources may also stem from parameter uncertainties in the model, which could be estimated with Markov Chain Monte Carlo methods, or from random perturbations in the precipitation inputs. This suggests, that calibrated hydrological model simulations did not provide a framework suitable to infer information about the quality of the input data.

2.4 Conclusions and limitations

In this project, grid point specific corrections for daily solid precipitation for widely used Swiss precipitation grids were produced and evaluated. The monthly varying corrections can be applied to a long-term climatology of precipitation grids. A day-by-day correction is available for the last seven years. On the point scale, the correction factors for the underestimation of gauges fluctuated around 1.2. Regional differences have prevented the formation of a clearer altitudinal gradient. Due to published results an undercatch of snow (no mixed precipitation) of about 45% in the Swiss Alps at approximately 2500 m elevation was expected. When excluding mixed precipitation (with excluding warm days > -2 °C), corrections in a similar magnitude could be achieved. A limitation is the parametric snow model used for snow accumulation, compaction and melting, which is a rather simple approach compared to energy balance snow modelling. This implies that important processes such as simultaneous accumulation and melting or the varying density of new snow could not be taken into account. However, it was shown that this method was robust for stations at different elevations. We were able to show that at the point scale it was beneficial to correct for solid precipitation using the proposed methods, but the benefits of the more elaborate corrections over the benchmark method (a constant correction of 1.2) were relatively small.

At the spatial scale the spatial averages needed to be corrected for both gauge undercatch and flat field overestimation compared to complex terrain. The correction for the latter is based on published data that does not account for more recent data sets. In addition, the correction was not applicable to forested terrain. The chosen evaluation methods were not able to distinguish between corrected, uncorrected, or falsely corrected precipitation variants, even though falsely corrected precipitation was largely different from the other variants.

2.5 Outreach work, publication of data and results

Preliminary results were presented at the Swiss Geoscience Meeting 2022 in Lausanne. A shortened manuscript version of this report is planned to be published in German in Wasser, Energie, Luft.

2.6 Outlook

The limits of the parametric snow model can be updated using particle filter data assimilation techniques and energy balance snow models. Ongoing work in this direction can be used to update the results presented here. We used published data to estimate the overestimation of the flat field (Grünewald et al., 2015), while in the meantime new LiDAR measurements with new and more precise technology are available at additional sites in Switzerland. It needs to be checked whether the corrections used here are also supported by more recent LiDAR data, especially in forested terrain. In addition, evaluation methods for the spatial scale need to be developed that can distinguish between small differences in precipitation input. One possibility is the frequently used method of calibrating and evaluating hydrological models in nearby catchments or sub-catchments of a larger catchment.

2.7 Acknowledgements

We want to thank Jan Magnusson, Bertrand Cluzet and Rebecca Mott for their assistance with model and evaluation code and in-depth discussions. We would also like to thank Massimiliano Zappa for providing both hydrological data and model access.

2.8 References

Brunner, M. I., Liechti, K., and Zappa, M.: Extremeness of recent drought events in Switzerland: dependence on variable and return period choice, Nat. Hazards Earth Syst. Sci., 19, 2311–2323, https://doi.org/10.5194/nhess-19-2311-2019, 2019.

Clark, M. P., Vogel, R. M., Lamontagne, J. R., Mizukami, N., Knoben, W. J. M., Tang, G., Gharari, S., Freer, J. E., Whitfield, P. H., Shook, K. R., and Papalexiou, S. M.: The Abuse of Popular Performance Metrics in Hydrologic Modeling, Water Resour. Res., 57, e2020WR029001, https://doi.org/10.1029/2020WR029001, 2021.

Coron, L., Delaigue, O., Thirel, G., Perrin, C., and Michel, C.: airGR: Suite of GR Hydrological Models for Precipitation-Runoff Modelling, https://doi.org/10.15454/EX11NA, R package version 1.4.3.65, 2020.

Essery, R.: A factorial snowpack model (FSM 1.0), Geosci. Model Dev., 8, 3867–3876, https://doi.org/10.5194/gmd-8-3867-2015, 2015.

Frei, C.: Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances, Int. J. Climatol., 34, 1585–1605, 2014.

Frei, C. and Isotta, F. A.: Ensemble Spatial Precipitation Analysis From Rain Gauge Data: Methodology and Application in the European Alps, J. Geophys. Res.-Atmos., 124, 5757–5778, https://doi.org/10.1029/2018JD030004, 2019.

Goodison, B. E., Louie, P. Y., and Yang, D.: WMO solid precipitation measurement intercomparison, World Meteorological Organization Geneva, Switzerland, 1998.

Griessinger, N., Schirmer, M., Helbig, N., Winstral, A., Michel, A., and Jonas, T.: Implications of observation-enhanced energy-balance snowmelt simulations for runoff modeling of Alpine catchments, Adv. Water Resour., 133, 103410, https://doi.org/10.1016/j.advwatres.2019.103410, 2019.

Grünewald, T. and Lehning, M.: Are flat-field snow depth measurements representative? A comparison of selected index sites with areal snow depth measurements at the small catchment scale, Hydrol. Process., 29, 1717–1728, https://doi.org/10.1002/hyp.10295, 2015.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Kochendorfer, J., Earle, M., Rasmussen, R., Smith, C., Yang, D., Morin, S., Mekis, E., Buisan, S., Roulet, Y.-A., Landolt, S., Wolff, M., Hoover, J., Thériault, J. M., Lee, G., Baker, B., Nitu, R., Lanza, L., Colli, M., and Meyers, T.: How Well Are We Measuring Snow Post-SPICE?, B. Am. Meteorol. Soc., 103, E370–E388, https://doi.org/10.1175/BAMS-D-20-0228.1, 2022.

Lehning, M., Grünewald, T., and Schirmer, M.: Mountain snow distribution governed by an altitudinal gradient and terrain roughness, Geophys. Res. Lett., 38, 1–5, https://doi.org/10.1029/2011GL048927, 2011.

Magnusson, J., Gustafsson, D., Hüsler, F., and Jonas, T.: Assimilation of point SWE data into a distributed snow cover model comparing two contrasting methods, Water Resour. Res., 50, 7816–7835, https://doi.org/10.1002/2014WR015302, 2014.

MeteoSwiss: Documentation of MeteoSwiss Grid-Data Products Daily Precipitation (final analysis): RhiresD, 2021.

MeteoSwiss: Documentation of MeteoSwiss Grid-Data Products Daily Precipitation Ensemble: RhydchprobD, 2019.

Rasmussen, R. M., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A. P., Black, J., Thériault, J. M., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K., and Gutmann, E.: How well are we measuring snow: The NOAA/FAA/NCAR winter precipitation test bed, B. Am. Meteorol. Soc., 93, 811–829, https://doi.org/10.1175/BAMS-D-11-00052.1, 2012.

Smith, C. D., Ross, A., Kochendorfer, J., Earle, M. E., Wolff, M., Buisán, S., Roulet, Y.-A., and Laine, T.: Evaluation of the WMO Solid Precipitation Intercomparison Experiment (SPICE) transfer functions for adjusting the wind bias in solid precipitation measurements, Hydrol. Earth Syst. Sci., 24, 4025–4043, https://doi.org/10.5194/hess-24-4025-2020, 2020.

Viglione, A. and Parajka, J.: TUWmodel: Lumped Hydrological Model for Education Purposes, available at: https://CRAN.R-project.org/package=TUWmodel (last access: 27 Feb 2023), R package version 1.1-1, 2020.

Viviroli, D., Zappa, M., Gurtz, J., and Weingartner, R.: An introduction to the hydrological modelling system PREVAH and its pre- and post-processing-tools, Environ. Modell. Softw., 24, 1209–1222, https://doi.org/10.1016/j.envsoft.2009.04.001, 2009.